



University  
of Glasgow

Adamson, Walt Eric (2006) *Mapping the proteins of the herpes simplex virus type 1 capsid.*

PhD thesis

<http://theses.gla.ac.uk/3963/>

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

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

This thesis 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

**MAPPING THE PROTEINS OF THE HERPES SIMPLEX VIRUS  
TYPE 1 CAPSID**

**BY**

**WALT ERIC ADAMSON**

**A THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF  
PHILOSOPHY  
IN THE FACULTY OF BIOMEDICAL AND LIFE SCIENCES  
AT THE UNIVERSITY OF GLASGOW**

**MRC VIROLOGY UNIT  
INSTITUTE OF VIROLOGY  
CHURCH STREET  
GLASGOW  
G11 5JR**

**MAY 2006**

ORIGINAL COPY TIGHTLY BOUND

# Table of contents

<b>ACKNOWLEDGEMENTS</b>	8
<b>ABSTRACT</b>	9
<b>ABBREVIATIONS</b>	11
<b>AUTHOR'S DECLARATION</b>	15
<b>1. INTRODUCTION</b>	16
1.1. INTRODUCTION TO HERPESVIRUSES	16
1.1.1. Classification of herpesviruses	16
1.2. VIRION ARCHITECTURE	21
1.2.1. The genome	21
1.2.2. The capsid	24
1.2.3. The tegument	30
1.2.4. The envelope	33
1.3. VIRAL REPLICATION	34
1.3.1. Virus attachment and entry	34
1.3.2. Transport to the cell nucleus	35
1.3.3. Disruption of host cell protein synthesis	37
1.3.4. HSV-1 gene expression	38
1.3.5. Viral DNA replication	40
1.3.6. DNA packaging	40
1.3.7. Acquisition of tegument and egress from the host cell	44
1.3.8. Latency	47
1.4. CAPSID STRUCTURE	49
1.4.1. Principals of capsid structure	49
1.4.2. Overview of capsid structure	52
1.4.3. The asymmetric unit	53
1.4.4. Structural composition of the capsid	54
1.5. CAPSID ASSEMBLY	58
1.5.1. Nuclear assembly of HSV-1 capsids	59
1.6. THE CAPSID IN VIRUS EVOLUTION	64
1.6.1. Three-dimensional reconstructions of herpesvirus capsids	65
1.6.2. Similarities between virus families	66
1.6.3. Similarities between herpesviruses and dsDNA bacteriophages	67
1.7. EXPERIMENTAL AIMS	64
<b>2. MATERIALS AND METHODS</b>	71
2.1. MATERIALS	71
2.1.1. Chemicals	71
2.1.2. Oligonucleotides	71
2.1.3. Enzymes	71
2.1.4. Cell lines	71
2.1.5. Tissue culture medium	72
2.1.6. Viruses	72

2.1.7.	Bacterial culture medium	73
2.1.8.	Bacterial strains	74
2.1.9.	Plasmids	74
2.1.10.	Antibodies	75
2.1.11.	Buffers and solutions	75
2.1.12.	Commercial kits	78
2.1.13.	Miscellaneous reagents	78
2.2.	<b>METHODS</b>	79
2.2.1.	Gel electrophoresis	79
2.2.2.	DNA cloning and manipulation	80
2.2.3.	Bacterial methods	84
2.2.4.	Cell culture methods	86
2.2.5.	Electron microscopy	91
2.2.6.	Semi-dry western blot analysis	92
2.2.7.	Construction of recombinant baculoviruses	93
2.2.8.	FLAsH labelling	95
2.2.9.	Transient DNA packaging assay	96
<b>3.</b>	<b>USING INSERTIONAL MUTAGENESIS TO STUDY VP19C</b>	97
3.1.	INTRODUCTION	97
3.2.	CONSTRUCTION OF INSERTIONAL MUTANTS	98
3.3.	COMPLEMENTATION OF INSERTIONAL MUTANTS	102
3.4.	INTERACTION OF INSERTIONAL MUTANTS WITH VP23 AND VP5	103
3.5.	INSERTIONAL MUTANTS IN DNA PACKAGING	104
<b>4.</b>	<b>MUTATIONAL ANALYSIS OF THE VP19C N-TERMINAL REGION</b>	105
4.1.	THE VP19C N-TERMINAL	105
4.2.	CLONING STRATEGY FOR N-TERMINAL DELETION MUTANTS	105
4.2.1.	Nuclear Localisation of N-terminal Deletion Mutants	106
4.2.2.	'Repairing' the Mutation in pUL38-63FBpCI*	107
4.2.3.	Nuclear Localisation and Complementation of VP19C-NLS Fusions	108
4.3.	MAPPING THE VP19C NLS	109
4.4.	NATURE OF THE GROWTH DEFECT ASSOCIATED WITH VP19C-63NLS	111
4.5.	VP19C-45NLS DIFFERENCE MAPPING	112
<b>5.</b>	<b>MUTATIONAL ANALYSIS OF VP19C – DISCUSSION</b>	115
5.1.	COMPUTATIONAL ANALYSIS OF VP19C	115
5.2.	VP19C INSERTIONAL MUTAGENESIS	116
5.3.	MUTATIONAL ANALYSIS OF THE N-TERMINAL REGION OF VP19C	120
5.4.	THREE-DIMENSIONAL RECONSTRUCTIONS OF VP19C	123
<b>6.</b>	<b>LABELLING VP5 WITH FLASH-EDT<sub>2</sub></b>	125
6.1.	INTRODUCTION	125

6.2.	INSERTION OF THE CCPGCC MOTIF INTO THE UL19 OPEN READING FRAME	125
6.3.	LABELLING WITH FLASH-EDT <sub>2</sub>	128
7.	<b>SITE-DIRECTED MUTAGENESIS OF VP5</b>	130
7.1.	INTRODUCTION	130
7.2.	IDENTIFICATION OF VP5 RESIDUES THAT MIGHT HAVE A ROLE IN VP26 BINDING	130
7.3.	INTRODUCTION OF AN <i>Mlu</i> I SITE INTO THE UL19 OPEN READING FRAME	130
7.4.	CONSTRUCTION OF SITE-DIRECTED MUTANTS	132
7.5.	IDENTIFICATION OF FURTHER VP5 RESIDUES THAT MIGHT HAVE A ROLE IN VP26 BINDING	133
7.6.	ANALYSIS OF SITE-DIRECTED MUTANTS	134
	7.6.1. Complementation	134
	7.6.2. Interaction with VP19C and preVP22a	134
	7.6.3. DNA packaging	135
	7.6.4. Interaction with VP26	135
8.	<b>USING MUTAGENESIS TO STUDY VP5 – DISCUSSION</b>	136
8.1.	LABELLING VP5 WITH FLASH-EDT <sub>2</sub>	136
8.2.	VP5 UPPER DOMAIN SITE-DIRECTED MUTAGENESIS	138
9.	<b>CONCLUSIONS</b>	141

## REFERENCES

**Mutational Analysis of the Herpes Simplex Virus Triplex Protein VP19C. (2006). Adamson, W. E., McNab, D., Preston, V. G. and Rixon, F. J. *Virol.* 80, 1537-1548.**

# List of figures and tables

All figures and tables are located on the page immediately after the listed page numbers.

## CHAPTER 1 – INTRODUCTION

Table 1.1	Human herpesviruses and their primary and associated illnesses	20
Table 1.2	Summary of the HSV-1 major capsid proteins	25
Table 1.3	Summary of the HSV-1 tegument proteins and their known functions	30
Table 1.4	Summary of the HSV-1 glycoproteins and their known functions	33
Table 1.5	Summary of the HSV-1 DNA packaging proteins and their known functions	41
Figure 1.1	Lineage map showing the evolutionary divergence of the herpesviruses	18
Figure 1.2	Alpha, beta, and gammaherpesviruses phylogenetic tree	18
Figure 1.3	Electron micrograph of a frozen hydrated HSV-1 virion	21
Figure 1.4	Herpesvirus genome organisation	22
Figure 1.5	Organisation of the HSV-1 genome	22
Figure 1.6	Cross-section through a cryo-EM reconstruction of a HSV-1 virion	23
Figure 1.7	Structure of the HSV-1 capsid determined to 8.5 Å resolution	24
Figure 1.8	Structural organisation of the HSV-1 scaffolding genes	28
Figure 1.9	Visualisation of icosahedrally ordered tegument proteins associated with the HSV-1 capsid	31
Figure 1.10	Overview of the HSV-1 lytic cycle	34
Figure 1.11	Various views of a 3D reconstruction of the HSV-1 portal at 8.5 Å resolution	41
Figure 1.12	The genetic organisation of the LAT region	48
Figure 1.13	Schematic representation of the relative locations of the various structural components of the HSV-1 capsid	51
Figure 1.14	Structure of the triplex at 8.5 Å resolution	51
Figure 1.15	The asymmetric unit of the HSV-1 capsid	52
Figure 1.16	A ribbon representation of the crystal structure of the upper domain of VP5	53
Figure 1.17	Structural comparisons of the hexon and penton	55
Figure 1.18	The main events in the HSV-1 virus assembly pathway showing the relationship between capsid assembly, maturation, and DNA packaging	58
Figure 1.19	Comparison of P3 and adenovirus hexon	66
Figure 1.20	The secondary structural elements of HSV-1 VP5 and HK97 gp5, and their molecular interactions in capsids	70

## CHAPTER 2 – MATERIALS AND METHODS

Figure 2.1	pGEM-T Easy, pFASTBAC1, and pFASTBAC pC1	74
------------	--	----

### CHAPTER 3 – USING INSERTIONAL MUTAGENESIS TO STUDY VP19C

Table 3.1	Sequencing primers used in DNA sequence analysis of the UL38 insertional mutants	100
Table 3.2	Locations and sequences of five amino acid insertions into VP19C	102
Table 3.3	Intracellular distribution of the severely disabled VP19C insertional mutants	104
Figure 3.1	The Mutation Generation System	97
Figure 3.2	Translation of the 15bp insertions	97
Figure 3.3	Digestion of pooled DNA with <i>HincII</i> and <i>XbaI</i> following insertional mutagenesis	98
Figure 3.4	Estimating positions of insertions into the UL38 ORF	100
Figure 3.5	<i>PstI</i> digests to identify clones with insertions lying in or near the UL38 ORF	101
Figure 3.6	Functional analysis of VP19C insertional mutants	102
Figure 3.7	Influence of VP19C insertional mutants on the distribution of VP23	103
Figure 3.8	Influence of VP19C insertional mutants on the distribution of VP5	103
Figure 3.9	Ability of VP19C insertional mutants to package DNA	104

### CHAPTER 4 – MUTATIONAL ANALYSIS OF THE VP19C N-TERMINAL REGION

Table 4.1	PCR primers used in cloning UL38 N-terminal mutants	105
Figure 4.1	Cloning strategy for VP19C N-terminal deletion mutants	105
Figure 4.2	Cloning of the UL38 N-terminal truncation	105
Figure 4.3	The mutation in pUL38-63FBpCI*, and its subsequent correction	107
Figure 4.4	Roles of the VP19C N-terminal sequences	108
Figure 4.5	Analytical agarose gels showing <i>EcoRI/SalI</i> digests of representative pGEM-T Easy N-terminal VP19C clones	109
Figure 4.6	Mapping the VP19C NLS	110
Figure 4.7	Expression of VP19C, VP19C-45NLS, and VP19C-63NLS in U2OS cells	111
Figure 4.8	Effects of VP19C N-terminal mutations on capsid assembly	111
Figure 4.9	Expression of VP19C, VP19C-45, and VP19C-63 in Sf21 cells	112
Figure 4.10	Cloning strategy for UL38-45NLSrescue	113
Figure 4.11	Protein composition of truncated VP19C capsids, and an electron cryomicroscopic image of vUL38-45 NLS capsids in vitreous ice	113
Figure 4.12	UL38-45NLS capsids in difference mapping	113
Figure 4.13	3D representations of wild-type and VP19C-45NSL triplexes	114

### CHAPTER 5 – MUTATIONAL ANALYSIS OF VP19 – DISCUSSION

Table 5.1	Chi-squared tests showing differences in the three regions of VP19C in terms of ability to tolerate insertional mutagenesis	118
-----------	---	-----

Figure 5.1	Sequence comparisons of VP19C homologues	115
Figure 5.2	Sequence analysis of herpesvirus triplex subunits	116
Figure 5.3	Three-dimensional reconstructions of the HSV-1 triplex proteins	116
Figure 5.4	Nuclear localisation of UL38in38	117
Figure 5.5	Arginine residues in the N-terminal regions of triplex $\alpha$ subunits of alphaherpesviruses	120
Figure 5.6	The similarity between the tops of the two VP23s in a triplex	123
Figure 5.7	Disorder predictions for HSV-1 VP19C and VP23	124

## CHAPTER 6 – LABELLING VP5 WITH FLASH-EDT<sub>2</sub>

Table 6.1	Oligonucleotides used for the insertion of the FIAsh sequence into VP5 insertion mutants	127
Figure 6.1	FIAsh-EDT <sub>2</sub> labelling of proteins	125
Figure 6.2	Positions in the VP5 upper domain at which extra amino acids were added for the introduction of the FIAsh-EDT <sub>2</sub> recognition sequence	127
Figure 6.3	<i>Sma</i> I digests were used to identify clones containing the FIAsh oligonucleotide	127
Figure 6.4	Abilities of VP5 mutants to complement VP5 null mutant virus with and without the insertion of the FIAsh oligonucleotide	128
Figure 6.5	FIAsh-EDT <sub>2</sub> labelling of cells expressing UL19in1235-Flash and preVP22a	129
Figure 6.6	FIAsh-EDT <sub>2</sub> labelling of cells transfected with UL19in1235-Flash	129

## CHAPTER 7 – SITE-DIRECTED MUTAGENESIS OF VP5

Table 7.1	Primers used in cloning UL19 site-directed mutants	131
Figure 7.1	The VP5 upper domain	130
Figure 7.2	PCR method for site-directed mutagenesis	130
Figure 7.3	Gel photographs from cloning steps during construction of VP5 site-directed mutants	131
Figure 7.4	Abilities of VP5 site-directed mutants to complement VP5 null mutant virus	134
Figure 7.5	Influence of VP19C and preVP22a on the distribution of VP5 insertional mutant pWA836	134
Figure 7.6	Ability of VP5 site-directed mutants to package DNA	135
Figure 7.7	Influence of VP5 site-directed mutants on the distribution of VP26	135

## CHAPTER 8 – USING MUTAGENESIS TO STUDY VP5 – DISCUSSION

Figure 8.1	Levels of background fluorescence often seen in cells labelled with FIAsh-EDT <sub>2</sub>	137
------------	--	-----

## Acknowledgements

I would like to thank my supervisor Dr Frazer Rixon for the advice and encouragement that he has given me throughout my PhD and for his thorough and critical reading of this thesis. I would also like to thank David McNab, Joyce Mitchell, Paul Hassan, and Marion McElwee not only for providing their technical expertise so willingly, but also for the friendship, good humour, and willingness to discuss Scottish First Division football that has made Lab 300 an enjoyable place to work over the last few years. Thanks also go to the other occupants of the Quiet Room, whose chat has helped break the thesis-writing tedium over the past few months.

I would like to thank Claire Addison for help with DNA sequencing, Jim Aitken for assistance in preparing samples for electron microscopy, and Wen Jiang, Matthew Baker, and Wah Chiu for production of three-dimensional capsid reconstructions.

I would like to thank my father, Donald Adamson, for his excellent assistance in proofreading this thesis, and along with Riika Adamson for providing support and encouragement throughout my studies.

Thanks must also go to the people who have been a welcome distraction to this PhD, and have helped to make the last few years so enjoyable: Andy, Beccy, Jo, Julia, Morag, and the various other characters from Team Walt, the Glasgow Queens Crew, and the Chippy Wall.

This thesis is dedicated to my mother, Liisi Adamson.

## Abstract

Herpes simplex type-1 (HSV-1) is the prototypical herpesvirus for scientific study. HSV-1 capsids have an icosahedral structure with capsomers formed by the major capsid protein, VP5, linked in groups of three by distinctive structures called triplexes. The aims of the work presented in this thesis were to use a variety of mutagenesis techniques to investigate the proteins of the HSV-1 capsid.

Triplexes are heterotrimers formed by two proteins in a 1:2 stoichiometry. The single-copy protein is called VP19C, and the dimeric protein is VP23. Insertional and deletional mutagenesis was carried out on VP19C and the effects of the mutations on virus growth and capsid assembly were examined. Insertional mutagenesis showed that VP19C can be divided into three regions with respect to their ability to tolerate five amino acid insertions, with two regions of approximately 100 amino acids at the N- and C- terminal regions of the protein being more tolerant of such insertions than a ~350 amino acid central region. The N-terminal ~100 amino acids of the protein, which are particularly insensitive to insertional mutagenesis, correspond to a region that is poorly conserved among herpesviruses. Some, but not all, severely disabled mutants were compromised in their ability to bind VP23 and VP5.

Analysis of deletional mutants revealed the presence of an unusual nuclear localisation signal (NLS) near the N-terminus of VP19C. This was mapped to a 33 amino acid region by fusion of specific sequences to a green fluorescent protein (GFP) marker.

By replacing the endogenous NLS with that from the simian virus 40 (SV40) large T antigen, we were able to show that the first 45 amino acids of VP19C were not essential for assembly of functional capsids and infectious virus particles. However, removing the first 63 amino acids resulted in the formation of aberrant capsids and prevented virus growth, suggesting that the poorly-conserved N-terminal sequences have some as-yet-unidentified function.

In an attempt to locate the N-terminus of VP19C within the triplex, the structure of B-capsids was determined for virus expressing a form of VP19C in which the first 45 amino acids were replaced by the SV40 NLS. Surprisingly, the most obvious differences were at the tops of the hexons, where additional masses appeared to be present in wild-type

capsids, and at the tops of the pentons, where the mutant capsids contained extra mass. The differences were not uniform across capsids, instead forming a gradient with the greatest apparent excess of wild-type capsid mass towards the centres of the faces and the greatest deficiency at the vertices. This pattern of differences can be explained most simply by a change in the overall curvature of the mutant virus capsid, resulting in a more angular capsid with flatter faces and more prominent vertices.

There are many examples of proteins that can be fused to GFP without affecting their properties. However in some cases the large size of GFP can interfere with the function and fate of recombinant proteins. Although GFP tagging has been carried out successfully on the smallest capsid protein, VP26, attempts to tag VP19C with GFP have been unsuccessful and the successful tagging of other capsid proteins with GFP has not been reported. The labelling of proteins with FIAsh-EDT<sub>2</sub> has been suggested as a possible alternative to GFP tagging. It has a significant advantage over GFP tagging in that the only alteration to the protein sequence that is required is the addition of a short (typically 6-12 amino acids) motif. The sequence of VP5 was altered to include this motif, and although FIAsh-EDT<sub>2</sub> was able to label the recombinant protein, levels of background fluorescence were such that without improvements this method was not suitable for studying the HSV-1 capsid.

VP5 interacts with VP26 at the tips of hexons and with tegument at the tips of pentons. Our understanding of these interactions has been aided by the resolution of the crystal structure of the VP5 upper domain. Since the publication of the crystal structure, computational and experimental approaches have been used to identify the particular amino acids involved in these interactions. Computational analysis led to eighteen amino acids in the VP5 upper domain being suggested as having a role in these interactions. Site-directed mutagenesis was carried out on nine of these mutants, and although one mutation prevented the production of infectious virus, none of the mutants examined were unable to bind VP26, suggesting that the interaction between VP5 and VP26 may involve several points of contact.

Taken together, the results in this thesis have helped to increase our levels of understanding of the HSV-1 capsid, and provided new insights into the structure of individual proteins within the capsid.

## Abbreviations

<b>A</b>	
aa	amino acid
Å	Angstrom
Ac	<i>Autographa californica</i>
AIHV	Alcelaphine herpesvirus
Amp	ampicillin
APS	ammonium persulphate
AtHV	Ateline herpesvirus
ATP	adenosine triphosphate
<b>B</b>	
BHK-21	baby hamster kidney-21
BoHV	Bovine herpesvirus
bp	base pair
BSA	Bovine serum albumin
<b>C</b>	
C-	carboxy (terminal)
°C	degrees centigrade
CaHV	Callitrichine herpesvirus
CCV	Channel catfish virus
CeHV	Cercopithecine herpesvirus
cryo-EM	electron cryo-microscopy
<b>D</b>	
3D	three-dimensional
Da	Dalton
dH <sub>2</sub> O	deionised water
dNTPs	2'-deoxynucleoside-5'-triphosphates
DNA	deoxyribonucleic acid
ds	double-stranded
<b>E</b>	
EBV	Epstein-Barr virus
EHV	Equine herpesvirus
<i>E. coli</i>	<i>Escherichia coli</i>
EDT	1,2-ethanedithiol
EDTA	ethlenediaminetetra-acetic acid
EM	electron microscopy
<b>F</b>	
FCS	foetal calf serum
FIAsH	fluorescein arsenical helix binder
FITC	fluorescein isothiocyanate
<b>G</b>	
g	gram
GAG	glycosaminoglycan
GaHV	Gallid herpesvirus
GAM	goat anti-mouse antibody
GAR	goat anti-rabbit antibody
GC	G+C content of DNA

GFP	green fluorescent protein
GMEM	Glasgow modified medium
<b>H</b>	
HCl	hydrochloric acid
HCMV	Human cytomegalovirus
HHV	Human herpesvirus type
HSV	Herpes simplex virus type
<b>I</b>	
ICP	infected cell protein
IE	Immediate Early
IRL	internal repeat long
IRS	internal repeat short
<b>K</b>	
k	kilo ( $10^3$ )
kb	kilobase pair
kDa	kilodalton
KSHV	Kaposi's sarcoma herpesvirus
<b>L</b>	
l	litre
LAT	latency associated transcript
LB	Luria-Bertani medium (L-Broth)
<b>M</b>	
M	molar
m	milli ( $10^{-3}$ )
mAb	monoclonal antibody
McHV	Macaque herpesvirus
MeHV	Meleagrid herpesvirus
mg	milligram
min	minute
ml	millilitre
mm	millimetre
mM	millimolar
m.o.i.	multiplicity of infection
mRNA	messenger RNA
MuHV	Murine herpesvirus
MW	molecular weight
<b>N</b>	
n	nano ( $10^{-9}$ )
N-	amino (terminal)
NBCS	new born calf serum
ng	nanogram
nm	nanometre
<b>O</b>	
ORF	open reading frame
OsHV	Oyster herpesvirus
<b>P</b>	

PaHV	Panid herpesvirus
PBS	phosphate buffered saline
PCMV	Porcine cytomegalovirus
PCR	polymerase chain reaction
pfu	plaque forming units
PRV	Pseudorabies virus
<b>R</b>	
RNA	ribonucleic acid
rpm	revolutions per minute
RRV	Rhesus rhadinovirus
RSB	reticulocyte standard buffer
RSC	rabbit skin cells
RSCUL19	RSC cell line expressing UL19
RSCUL38	RSC cell line expressing UL38
<b>S</b>	
SaHV	Saimirine herpesvirus
SCMV	Simian cytomegalovirus
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sf	<i>Spodoptera frugiperda</i>
ss	single stranded
SuHV	Suid herpesvirus
SV40	Simian virus 40
<b>T</b>	
TEMED	N,N,N',N',-tetramethylethylene
Tris	tris (hydroxymethyl) aminoethane
TRITC	texas red isothiocyanate
ts	temperature sensitive
TuHV	Tupaiaid herpesvirus
<b>U</b>	
μ	micro (10 <sup>-6</sup> )
μg	microgram
μl	microlitre
UL	unique long
US	unique short
UV	ultra violet
<b>V</b>	
V	volt
vhs	virion host shutoff protein
VP5ud	VP5 upper domain
v/v	volume to volume ratio
VZV	Varicella-Zoster virus
<b>W</b>	
WT	wild-type
w/v	weight to volume ratio
<b>X</b>	
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactosidase

**Amino Acid Abbreviations**

<u>Amino Acid</u>	<u>Three letter code</u>	<u>One letter code</u>
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

## **Author's Declaration**

All the results presented in this thesis were obtained by the author's own efforts, unless otherwise stated.

# **CHAPTER 1**

## **INTRODUCTION**

# 1. Introduction

## 1.1. Introduction to herpesviruses

### 1.1.1. Classification of herpesviruses

Herpesviruses are a large, diverse family with a host range that probably comprises all vertebrates and at least one type of invertebrate. In nature, each vertebrate herpesvirus is closely associated with a single host species, and the hosts most extensively studied so far can be infected by several distinct herpesviruses. It is therefore likely that the 124 recognised species described as of 2005 (Davison *et al.*, 2005a) represent only a tiny proportion of the number in existence. The host-specific occurrence of herpesviruses indicates that they have evolved with their hosts over long periods of time and are extremely well adapted to them. (Davison, 2002).

The primary means of identifying members of the herpesvirus family has been according to virion structure. The virions are 200-250 nm in diameter and have four distinct layers. The dsDNA genome is packaged within an icosahedral shell known as the capsid, which is approximately 125 nm in diameter. The capsid is surrounded by an amorphous protein layer known as the tegument, which is in turn wrapped in a lipid membrane containing several viral glycoproteins. (Roizman and Pellett, 2001; Davison, 2002).

As well as having a common virion structure, herpesviruses also have some common characteristic biological properties.

- They all encode a large number of proteins and enzymes involved in nucleic acid metabolism, and DNA syntheses.
- Viral DNA replication and capsid assembly occur in the nucleus of the infected cell.
- The production of infectious progeny virus usually results in the destruction of the host cell.
- All herpesviruses studied have the ability to establish latent infection (the dormant state for viruses, in which only a small proportion of viral genes are expressed and no progeny virus is produced).

Although herpesviruses share a number of biological properties, they exhibit considerable variation, which can be seen in many aspects of herpesvirus biology. Variation occurs in

the cell types that the virus can infect in tissue culture, the length of the replication cycle and the cell types in which latent infection can be established.

While classification as herpesviruses has primarily been based on virion structure, various additional criteria have been used for classifying herpesviruses into lower taxa, within the categories of subfamily, genus and species. The criteria used until recently were mainly biological properties – which were largely successful for classification into subfamilies – and antigenic properties. The majority of currently known herpesviruses infect mammals and birds. These herpesviruses can be further classified as *Alphaherpesvirinae*, *Betaherpesvirinae* and *Gammaherpesvirinae*.

#### **1.1.1.1. Alphaherpesvirinae**

Herpesviruses within this subfamily have the ability to infect a wide range of hosts in tissue culture under experimental conditions. They have short reproductive cycles (<24 hours), efficiently destroy infected cells and grow rapidly in cell culture. They establish latent infection in the ganglia of the sensory nervous system. *Alphaherpesvirinae* can be further classified into four genera: *Simplexviruses*, which include Herpes Simplex Virus type-1 and type-2 (HSV-1 and HSV-2), *Varicelloviruses*, which include Varicella-Zoster Virus (VZV); *Mardiviruses*, which include Marek's disease virus-2 and -3 (GaHV-2 and GaHV-3); and the *Iltoviruses*, which include Infectious laryngo-tracheitis virus (ILTV or GaHV-1). (Roizman and Pellet, 2001).

#### **1.1.1.2. Betaherpesvirinae**

Herpesviruses within this subfamily have a restricted host range in tissue culture. They have a long reproductive cycle and grow slowly in cell culture conditions. Infected cells can become large (cytomegalia), and carrier cultures can be established in which cells survive even after being infected. *Betaherpesvirinae* tend to establish latency in a wide range of tissues, for example in secretory glands and the kidneys. *Betaherpesvirinae* can be further divided into four genera: *Cytomegaloviruses*, which include Human cytomegalovirus (HCMV); *Muromegaloviruses*, which include Murine cytomegaloviruses (MHV-1); *Roseoloviruses*, which include Human herpesviruses type-6 (HHV-6) and type-7 (HHV-7); and *Probosciviruses*, which include Elephant Endothelial Herpesvirus (EIHV-1). (Roizman and Pellet, 2001; McGeoch *et al.*, 2005).

### 1.1.1.3. **Gammaherpesvirinae**

Herpesviruses within this subfamily have a very restricted host range in tissue culture, being limited to the family or order to which the natural host belongs. *Gammaherpesvirinae* can replicate in lymphoblastoid cells and some also exhibit lytic replication in certain types of epithelial and fibroblastic cells. In infections within the natural host, *Gammaherpesvirinae* tend to be specific to B or T lymphocytes. Latent infections are usually established in lymphoid tissue. They can be further subdivided into three genera: *Lymphocryptoviruses*, which include Epstein-Barr virus (EBV); *Rhadinoviruses*, which include Kaposi's Sarcoma Herpesvirus (KSHV); and *Macaviruses*, which include Equine Herpesvirus type-2. (Roizman and Pellet, 2001; McGeoch *et al.*, 2005).

### 1.1.1.4. **Classification of herpesviruses using DNA sequence analysis**

In recent years, classification by biological properties has been superseded by classification based on the relationships between DNA and amino acid sequences. Thus, classification is based primarily on genetic information such as (i) the structure of the genome, (ii) DNA and protein sequence homologies, and (iii) the possession and conservation of certain genes, and their arrangements within gene clusters (Roizman and Pellett, 2001; McGeoch *et al.*, 1995). However, the original classifications have proved to be remarkably accurate, with few viruses requiring to be reclassified. One notable exception has been Marek's disease virus-1, the causative agent of tumour-like growths in the lymphoid tissues of chickens. This was originally classified as a gammaherpesvirus because of its ability to infect lymphoid tissue; however, DNA sequence analysis has shown that it is more akin to *Alphaherpesvirinae* and the virus has been reclassified accordingly (Buckmaster *et al.*, 1988).

More sequence information is available for herpesviruses than for any other large DNA virus family, and our understanding of the evolution of these viruses is consequently well developed. At the time of writing, the complete genome sequences of 39 herpesviruses were available (Andrew Davison, personal communication). Comparisons of amino acid and DNA sequences have allowed the construction of phylogenetic trees illustrating the genetic relationship between herpesviruses, as shown in Figures 1.1 and 1.2.

Sequence data is now available for several closely related pairs of species, such as HSV-1 and HSV-2, Equine herpesvirus type-1 and type-4 (EHV-1 and EHV-4), VZV and Simian

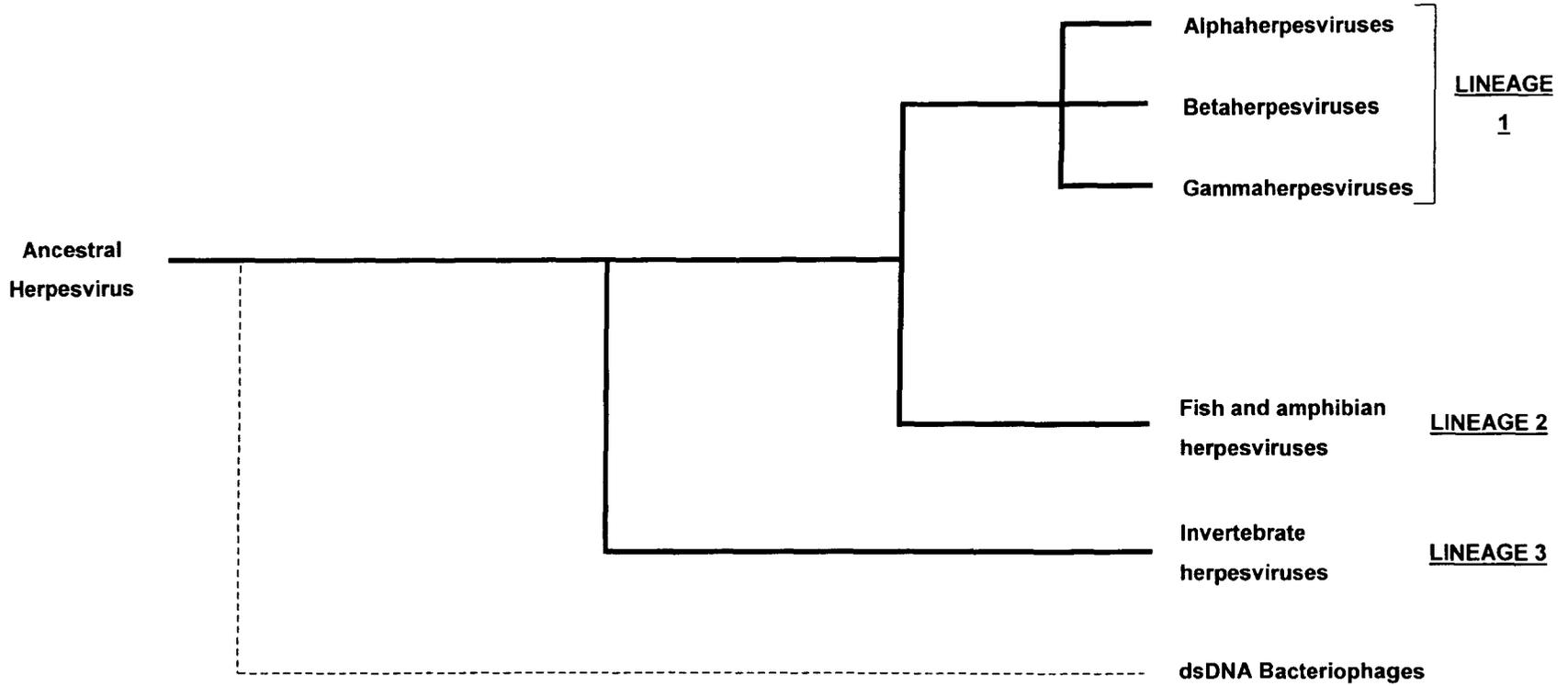


Figure 1.1: Lineage map showing the evolutionary divergence of the herpesviruses (not drawn to scale).

Lineage 1 comprises herpesviruses that infect mammalian and avian hosts, and which contains the alpha, beta, and gamma herpesvirus subfamilies. Lineage 2 comprises herpesviruses that infect fish and amphibians. Lineage 3 comprises herpesviruses that infect invertebrate hosts (of which only one example is known at present). A potential fourth lineage (broken line) is shown to reflect the similarities in structure, capsid assembly, and DNA packaging observed between *herpesviridae* and dsDNA bacteriophages that suggest that both groups are descended from a common ancestral virus.

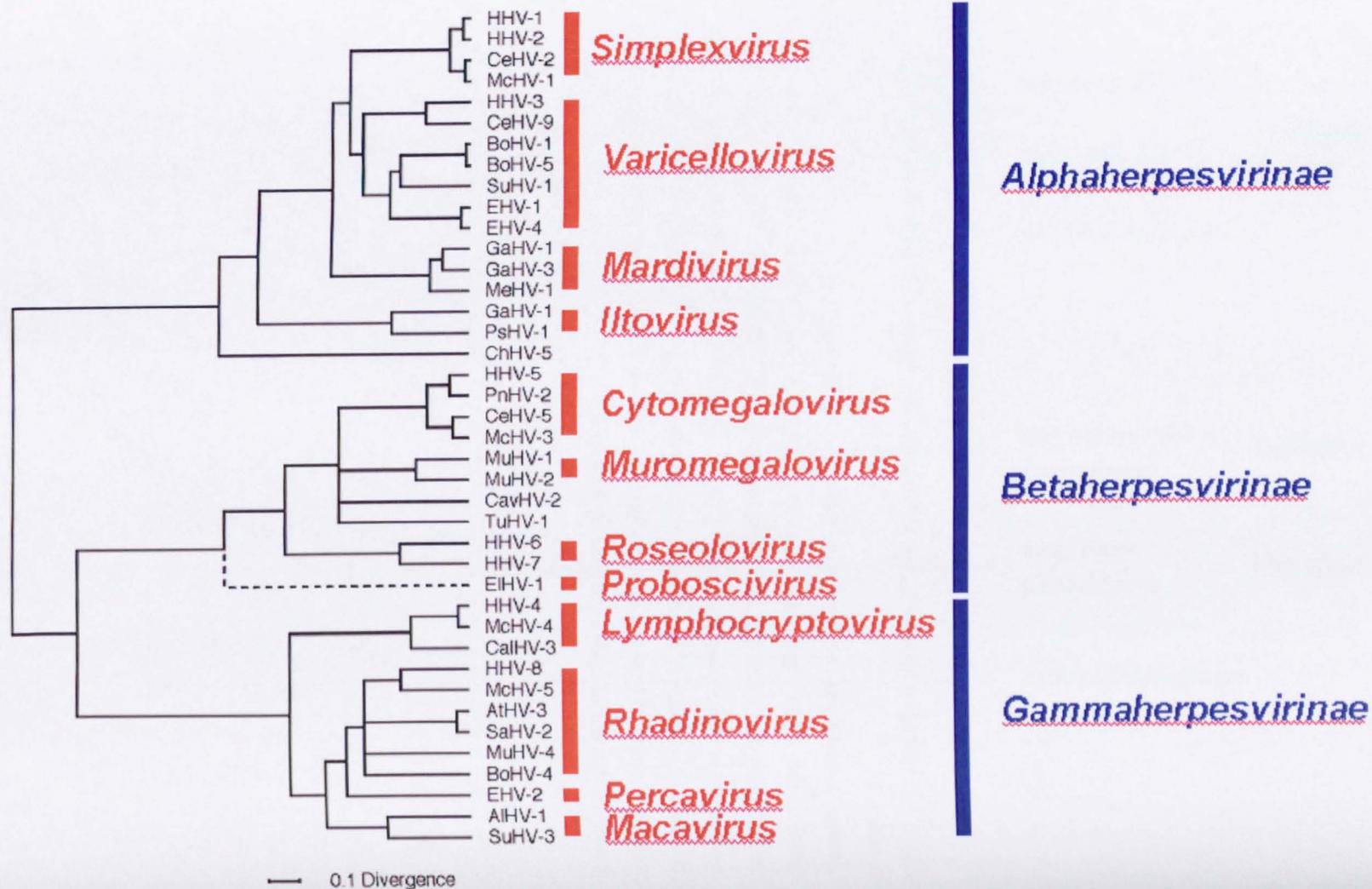


Figure 1.2: Alpha, beta, and gammaherpesvirus phylogenetic tree based on the alignments of six sets of homologous proteins.

Figure provided by Andrew Davison.

varicella virus (SVV), and HHV-8 and Rhesus rhadinovirus (RRV). Members of some pairs have presumably survived by occupying distinct biological niches within the same host. Divergence took place a long time before the present hosts evolved and thus was not necessarily caused by host speciation. For example, HSV-1 and HSV-2 are estimated to have diverged about eight million years ago (McGeoch *et al.*, 1995), implying that related pairs of viruses will also be present in animals that have diverged subsequently (such as other higher primates). However, if we consider the cases of VZV and SVV, and HHV-8 and RRV, we note that these infect different hosts (humans in the former of each pair and Old World monkeys in the latter), indicating that divergence may have begun upon host speciation.

#### **1.1.1.5. Fish, amphibian and invertebrate herpesviruses**

Until the last decade, sequence information was available only for mammalian and avian herpesviruses. These viruses are clearly related and evolved from a common ancestral herpesvirus. In recent years, two other major herpesvirus lineages have been discovered, one leading to fish and amphibian herpesviruses, and the other to invertebrate herpesviruses, of which only one example, Oyster Herpesvirus (OsHV-1), is known at present (Davison, 1992, 1998, 2002; Bernard and Mercier, 1993;). If these viruses evolved with their hosts, the fish viruses would have diverged about 450 million years ago and the invertebrate virus approximately one billion years ago (Kumar and Hedges, 1998; Wray *et al.*, 1996). The fish and frog herpesviruses evolved from a common ancestor as they share a subset of genes. It is not possible to evaluate the inheritance patterns of OsHV-1 genes since no related viruses have been identified. This virus is also unusual in that it is not restricted (at least in shellfish hatcheries) to a single bivalve species or even a single genus (Arzul *et al.*, 2001).

In contrast to the obvious relationships among viruses within the mammalian/avian or fish/amphibian lineages, the extent of divergence between the three lineages is so great that it is difficult to be convinced purely from comparisons of amino acid sequence that the viruses are derived from a common ancestor. Each lineage has a few functions that are ubiquitous among living organisms, such as DNA polymerase, DNA helicase, or protein kinase, but these could have been acquired by independent capture. Genes encoding proteins that are confined to herpesviruses are not detectably conserved. However, analysis of the Channel Catfish virus (CCV) and OsHV-1 has shown that both have a similar capsid morphology to HSV-1 (Davison and Davison, 1995; Davison *et al.*, 2005b) and that CCV and HSV-1 appear to package their DNA in a similar way, by expelling a scaffold protein

from a preformed capsid (Davison and Davison, 1995). On balance, therefore, it seems likely that the three major lineages evolved from a common ancestor. Support for this hypothesis must await the derivation of further information on herpesvirus protein structure, at which level distant evolutionary connections are more likely to be visible. Structural analysis of herpesvirus capsids will be discussed further in Section 1.4.

If the relationship between the three lineages is accepted, it implies that the common ancestor of all modern herpesviruses existed prior to the evolution of the first vertebrates, approximately 400 million years ago and that the morphology of this virus was unmistakably that of a herpesvirus (Davison, 2002). Also, the parallels drawn between capsid structure and the DNA packaging processes of herpesviruses and some bacteriophages suggest that herpesviruses have origins even earlier, before the separation of prokaryotic and eukaryotic cell types (Section 1.6).

#### **1.1.1.6. Herpesviruses that infect humans**

Currently, eight herpesviruses are known to infect humans (Table 1.1). Infection is widespread in both the developed and underdeveloped world. As with all herpesviruses, those that infect humans have the ability to establish latent infection, in which viral DNA is maintained as a circular episome within the nuclei of infected cells. The specific cell type in which latency is established varies between viruses. However, in all human herpesviruses latent infection lasts for the lifespan of the host. During that time the virus can sporadically reactivate and enter lytic replication.

It should be noted that latent infection is asymptomatic, as the virus only expresses a small subset of the viral genome, and it is only reactivation that leads to the recurrence of physical symptoms. For example, in HSV-1 the classical route of infection is through mucosal membranes of the mouth and nose, leading to latent infection in the trigeminal ganglia. The virus remains invisible to the host in the ganglia until reactivation, which can occur after an interval of many years, causing the reappearance of the virus at the site of the original infection where it may form characteristic cold sores. Other human herpesviruses behave in similar fashions, although reactivation does not always lead to the recurrence of the original symptoms. For example in VZV, primary infection (usually in children) causes a generalised infection that produces the distinctive skin lesions of chicken pox. As with HSV-1, latent infection is established in the ganglia of the sensory nervous system. However reactivation is not generalised, being normally restricted to a

Virus	Common Name	Subfamily ( <i>Genus</i> ) <sup>1</sup>	Site of Latency <sup>1</sup>	GC content (%) <sup>1,2</sup>	Genome Size (kb) <sup>2,3</sup>	Primary Associated Illness <sup>3,4,5,6</sup>	Other Associated Illnesses <sup>3,4,5,6</sup>
HHV-1	Herpes simplex virus type 1 (HSV-1)	<i>Alphaherpesvirinae (Simplexvirus)</i>	Sensory nerve ganglia	68.3	152	Cold sores	Genital lesions, conjunctivitis, herpetic whitlow, keratitis, encephalitis
HHV-2	Herpes simplex virus type 2 (HSV-2)	<i>Alphaherpesvirinae (Simplexvirus)</i>	Sensory nerve ganglia	70.4	152	Genital lesions	Oral lesions, conjunctivitis, herpetic whitlow, keratitis, encephalitis
HHV-3	Varicella-zoster virus (VZV)	<i>Alphaherpesvirinae (Varicellovirus)</i>	Sensory nerve ganglia	46	125	Chicken pox/fever	Reactivation can lead to shingles
HHV-4	Epstein-Barr virus (EBV)	<i>Gammaherpesvirinae (Lymphocryptovirus)</i>	B-lymphocytes	60	172	Glandular fever, mononucleosis	Burkitt lymphoma, nasopharyngeal carcinoma
HHV-5	Human cytomegalovirus (HCMV)	<i>Betaherpesvirinae (Cytomegalovirus)</i>	Leukocytes, endothelial cells	57	229	Congenital abnormalities, mononucleosis, hepatitis	Immuno-compromised individuals suffer from gastro-enteritis and retinitis
HHV-6	Exanthem subitum virus	<i>Betaherpesvirinae (Roseolovirus)</i>	T-lymphocytes	42	160	Infant rash ( <i>exanthema subitum</i> )	Associated with certain malignancies: lymphoma, leukaemia, co-factor in cervical and oral carcinoma
HHV-7	HHV-7	<i>Betaherpesvirinae (Roseolovirus)</i>	T-lymphocytes	45	145	Febrile illness	-
HHV-8	Kaposi's sarcoma-associated herpesvirus (KSHV)	<i>Gammaherpesvirinae (Rhadinovirus)</i>	B-lymphocytes	54	170	Associated with Kaposi's sarcoma	-

Table 1.1: Human herpesviruses and their primary and associated illnesses

(1) Roizman, 1992; (2) Subak-Sharpe and Dargan, 1998); (3) Levy, 1997; (4) Gold and Nankervis, 1991; (5) Evans and Niederman, 1991; (6) Nahmias *et al.*, 1991.

region of the skin associated with a single nerve tract (known as a dermatome), where it causes a painful rash known as shingles.

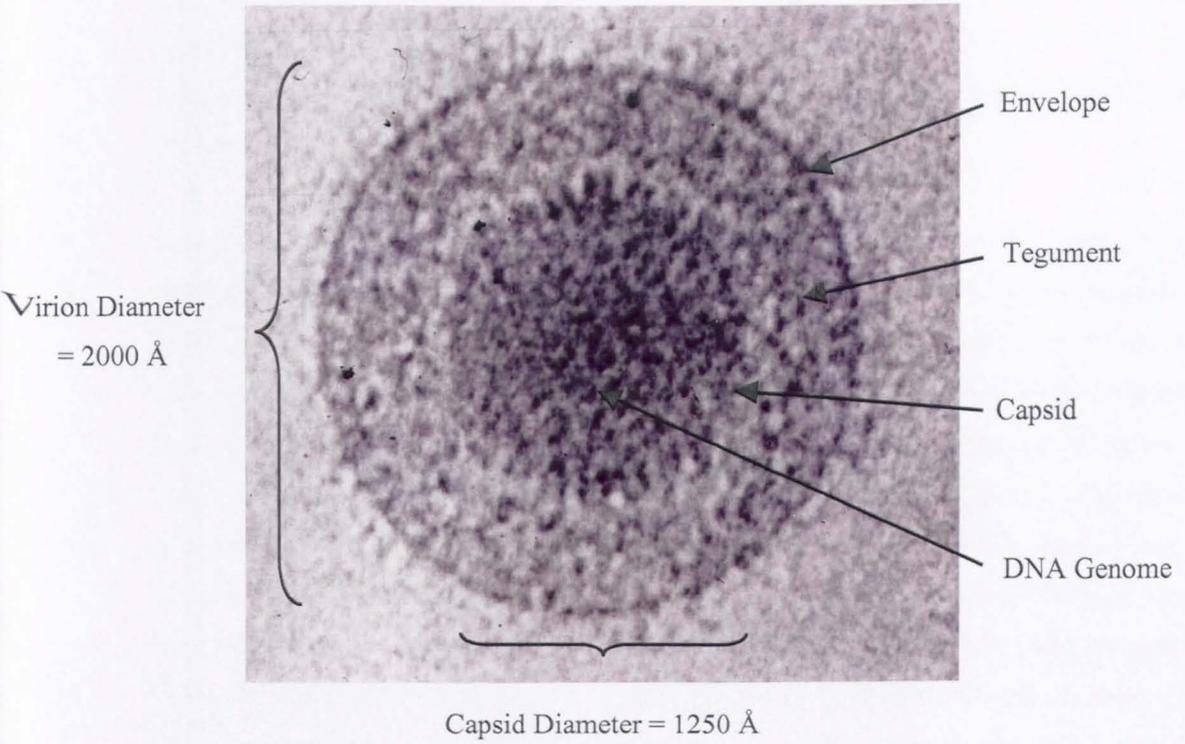
The causes of herpesvirus reactivation are poorly understood, but it often seems to be triggered by some form of stimulus, such as mental or physical stress, hormonal changes, old age, or exposure to UV radiation. One factor of particular importance is the immune state of the host, with particular relevance to individuals infected with the human immunodeficiency virus (HIV). Organ transplant patients or patients using immunosuppressive drugs are known to be at risk to reactivation of HCMV (Sia and Patel, 2000). The ability of herpesviruses to remain latent in the host without causing detectable symptoms is a survival mechanism that allows the virus to persist in a form invisible to the immune system, yet still retain the ability to spread to naive hosts when reactivated. Such a strategy requires that the primary infection does not kill the host, and it is not surprising that infections of humans are not normally fatal except in unusual circumstances. However, this only applies when herpesviruses infect their natural hosts. An example can be observed in Cercopithecine herpesvirus-1, an alphaherpesvirus that normally infects macaque monkeys. In the natural host, infection is benign, resulting in minor or no symptoms. However, in human infection, the result is severe encephalomyelitis, which is fatal in approximately 80% of cases. HSV-1 latency is discussed in further detail in Section 1.3.8.

## **1.2. Virion architecture**

The structure and morphology of the virion is conserved across herpesviruses. The virion consists of four distinct layers: the DNA genome, the capsid, the tegument, and the envelope (Figure 1.3). The structure and composition of each of these layers are discussed in the following section. The discussion will concentrate particularly on the HSV-1 virus capsid, which is the primary subject of the work described in this thesis.

### **1.2.1. The genome**

The HSV-1 (strain 17) genome is a single linear molecule reported as consisting of 152261 base pairs (bp) of double-stranded DNA (dsDNA), with a GC content of 68.3% (McGeoch *et al.*, 1988; Dolan *et al.*, 1998). Comparison between the genome sequences of HSV-1 and HSV-2 has resulted in the identification of 74 open reading frames (ORFs) that are conserved between the two viruses (McGeoch *et al.*, 1988; Dolan *et al.*, 1998).



**Figure 1.3:** Electron micrograph of a frozen hydrated HSV-1 virion.

The virion consists of four distinct layers: the DNA genome, capsid, tegument and envelope. Although not directly visible, the DNA genome is located within the capsid; the densely packaged nature of the DNA genome accounting for the electron-dense appearance of the capsid. The envelope is visible as a thin, electron-dense layer surrounding the virion, while the tegument is a less electron-dense layer separating the envelope and the capsid.

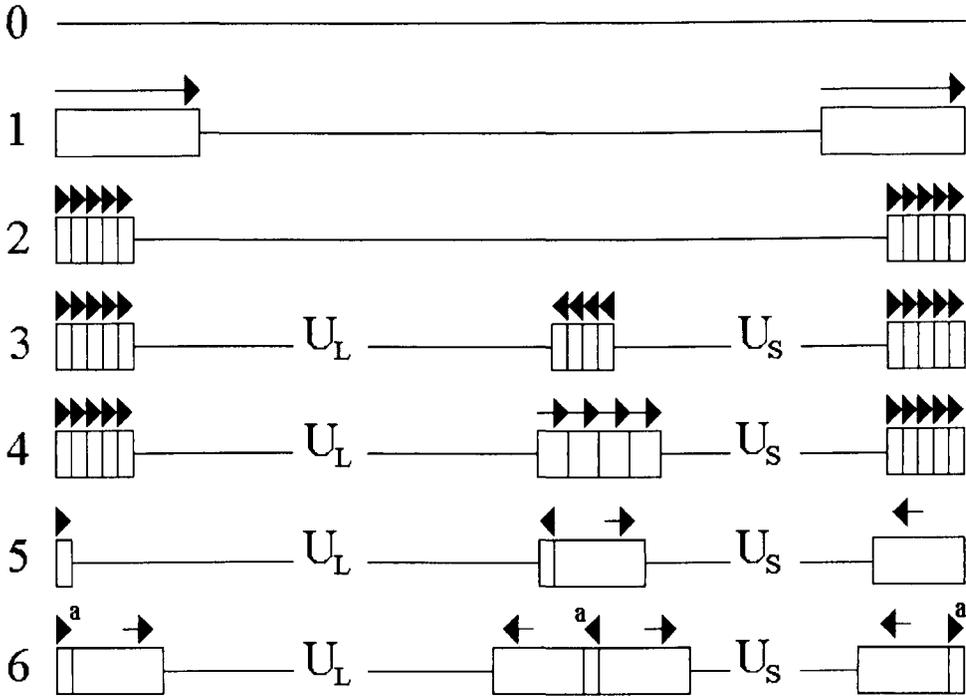
This figure was reproduced with permission from W. Chiu, Baylor College of Medicine, USA.

### 1.2.1.1. Organisation of the genome

Herpesvirus genomes are linear double-stranded genomes of 125-245 kb (Davison, 2002). Their base composition is highly variable, with GC contents ranging from 31% to 75%. Genomes have been classified into seven distinct groups according to their sequence arrangement and the presence of terminal or internal repeats (Roizman *et al.*, 1992; Davison and McGeoch, 1995). Figure 1.4 shows a schematic representation of the genome groups.

Group 0 genomes have a single unique region with no internal or terminal repeats. Tupaia herpesvirus (THV), isolated from the tree shrew, is characteristic of this group (Koch *et al.*, 1985). Group 1 genomes have a unique long region with direct repeats at each terminus. Members of this group include channel catfish virus (CCV), HHV-6 and HHV-7 (Davison, 1992; Gompels *et al.*, 1995; Nicholas, 1996). Group 2 genomes contain a single unique region flanked by multiple copies of a direct repeat at each terminus. Examples of group 2 genomes include Herpesvirus Saimiri (HVS) and KSHV (Albrecht *et al.*, 1992; Moore *et al.*, 1996). Group 3 genomes are similar to group 2, although they contain multiple copies of the same repeat sequence internally, in inverted orientation, which result in the genome being divided into two unique regions flanked by inverted repeats. The two unique regions can invert relative to each other, producing four genetic isomers that are isolated in equimolar amounts from infected cells. A virus isolated in rabbits, cottontail herpesvirus, is an example of this group (Cebrian *et al.*, 1989). Group 4 genomes contain two unique regions flanked by multiple direct repeats at each terminus as well as unrelated internal repeats. EBV has this genome structure (Baer *et al.*, 1984). Group 5 genomes have two unique regions flanked by inverted repeats. The inverted repeats are not related, but are specific to each unique region. VZV belongs to this genome class (Davison and Scott, 1986). Group 6 genomes are similar to group 5 however they contain an additional sequence (the  $\alpha$  sequence), which is present as direct repeats at the termini and in inverted orientation at the junction between the long and short segments. Additionally, the repeats flanking UL are longer than those seen in group 5. The long and short regions are capable of inversion and four genomic isomers can be isolated in equimolar amounts from virions. HSV-1, HSV-2 and HCMV are all members of this group (Roizman, 1979).

The organisation of the HSV-1 genome is illustrated in Figure 1.5. It is made up of two distinct regions – the long region (L) and the short region (S). Each region contains a unique sequence ( $U_L$  and  $U_S$ ) flanked by a pair of inverted repeat sequences known



**Figure 1.4:** Herpesvirus genome organisation

A schematic representation of the genome structures as defined by Roizman (1992), and Davison and McGeoch (1995). Group 0 genomes contain a single unique coding region and no repeat units. Group 1 genomes comprise a single unique region with a single direct repeat sequence at each terminus. Group 2 genomes are similar to group 1, but contain multiple repeats at each terminus. Group 3 genomes have multiple repeats at each terminus and internal repeats in the opposite orientation at the junction between the two unique regions. Group 4 genomes are similar to group 3, but the internal repeats show no similarity to the terminal repeats. Group 5 genomes have inverted repeats surrounding the two unique regions; the  $U_S$  repeats are larger than those flanking the  $U_L$  region. Group 6 genomes are similar to group 5, however they also contain *a* sequence repeats at each terminus and in inverted orientation at the L/S junction. Arrows indicate the orientation of the repeats in each case.

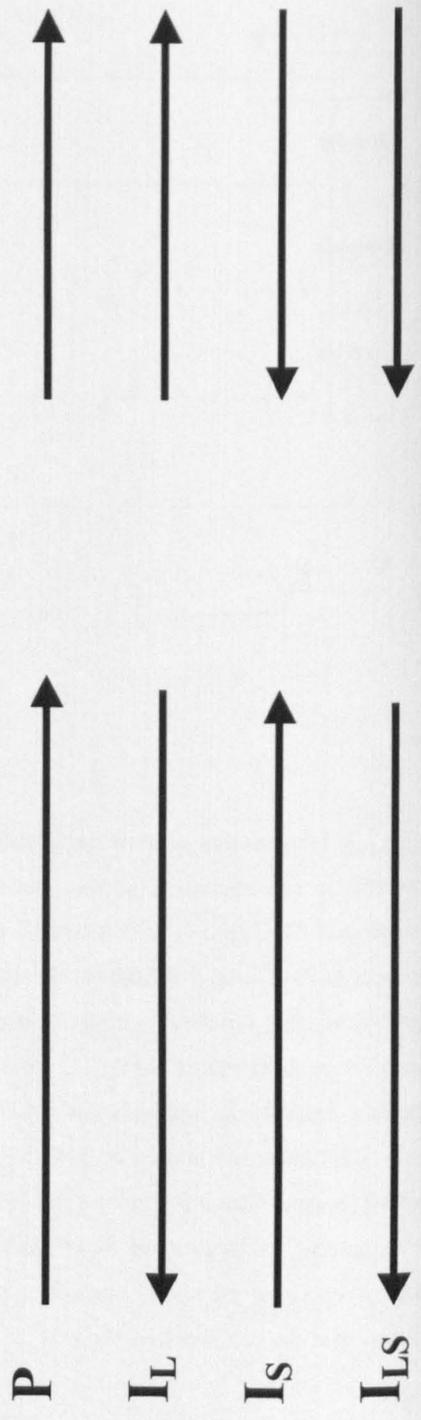


Figure 1.5: Organisation of the HSV-1 genome.

Above: A schematic illustration of the structural organisation of the HSV-1 genome. The genome is made up of two separate regions; L (long) and S (short), with each domain consisting of terminal repeats (TR<sub>L</sub> and TR<sub>S</sub>), internal repeats (IR<sub>L</sub> and IR<sub>S</sub>), unique sequences (UL and US), and the 'a' sequence located at the termini and at the L and S junction. TR<sub>L</sub>/IR<sub>L</sub>, and TR<sub>S</sub>/IR<sub>S</sub> are pairs of inverted repeats, therefore the L and S regions can be inverted relative to each other.

Below: A schematic illustration of the four possible isomers of the HSV-1 genome resulting from L and S inversion. P (prototype – no inversion of either the L or S region), IL (inversion of L), IS (inversion of S), and ILS (inversion of both L and S regions).

respectively as long repeats ( $R_L$ ) and short repeats ( $R_S$ ). Depending on their positions, they are known as terminal repeats ( $TR_L$  and  $TR_S$ ) or internal repeats ( $IR_L$  and  $IR_S$ ). The molecule also contains a terminal redundancy of approximately 400 bp known as the  $a$  sequence. One or more copies of the  $a$  sequence are located internally at the junction of the L and S segments, in the opposite orientation to the terminal  $a$  sequences (McGeoch *et al.*, 1988). The presence of inverted repeats at either side of the  $U_L$  and  $U_S$  regions means that internal recombination can occur, with the L and S regions being inverted relative to each other. This results in the formation of four isomers of the HSV-1 genome (Figure 1.5), known as P (prototype),  $I_L$  (inversion of the L region),  $I_S$  (inversion of the S region) and  $I_{LS}$  (inversion of both the L and S regions). DNA isolated from cells infected with wild-type HSV-1 consists of equimolar concentrations of the four isomers. Viruses in which the genomic DNA has been prevented from recombining due to deletion of the internal inverted repeats retain their ability to replicate in cell culture, indicating that the ability to rearrange the genome is not critical for viability (Roizman and Knipe, 2001). HSV-1 ORFs are distributed throughout the genome, including the inverted repeats. Each gene is named firstly with respect to the region of the genome in which it is located (i.e.  $U_L$ ,  $U_S$  etc.) and secondly by assigning a number relating to the gene order in that region (i.e.  $UL1$  is the gene closest to the start of  $U_L$  and  $UL56$  is the gene closest to the end of  $U_L$  in the P orientation). This system gives each gene a unique reference name. For example, the capsid shell proteins are encoded by ORFs located in the middle of the  $U_L$  region and are designated  $UL18$ ,  $UL19$ ,  $UL35$  and  $UL38$ .

### 1.2.1.2. Packaged state of the genome

In the herpesvirus virion the genome is found tightly packed inside the icosahedral capsid shell, filling all the available space within the capsid (Booy *et al.*, 1991). Together, the genome and capsid are referred to as the nucleocapsid. The DNA inside the virions is organised in concentric layers with a spacing of 26 Å (Figure 1.6) (Zhou *et al.*, 1999), which has been likened to the 'spooling' model of DNA packaging proposed for bacteriophage T7 (Cerritelli *et al.*, 1997). This arrangement suggests that the DNA, which enters through a unique portal, is wrapped around the inner surface of the capsid shell, accumulating one layer at a time until at least ten concentric layers of tightly packed DNA have formed. (It should be noted that although the DNA is observed to be closely associating with the inner surface of the capsid shell, no icosahedral symmetry has been detected in the structure of the packaged DNA, suggesting that no specific DNA-protein interactions are involved (Zhou *et al.*, 1999). The validity of this model was confirmed by Bhella *et al.* (2000), who showed that this model of packaging is also applicable to the

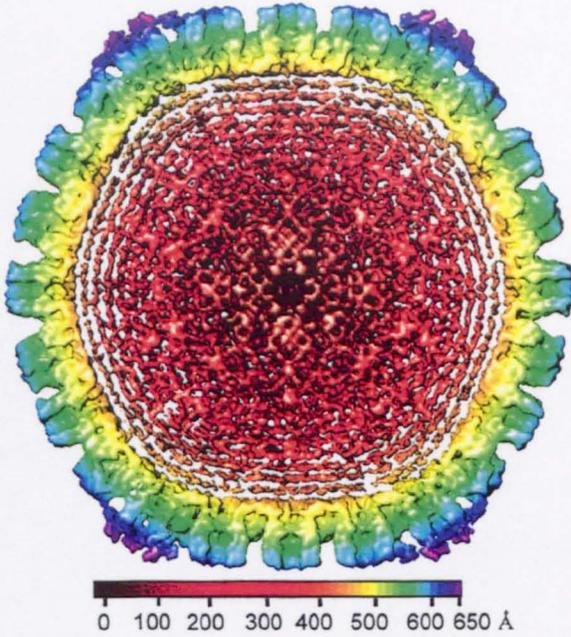


Figure 1.6: Cross-section through a cryo-EM reconstruction of a HSV-1 capsid.

This cross-section is viewed along a two-fold axis of symmetry. Concentric shells of density (coloured red) are attributable to the genomic DNA packaged within the capsid. The spacing between layers of DNA is 26 Å. The resolution of this structure is 20 Å. The figure is coloured radially, using the colour scheme shown beneath the virion. Internal structures (packaged DNA are shown in red, the capsid floor is shown in yellow, the triplexes and middle domains of the capsomers are shown in green, while the outer domains of the capsomers, including the tegument densities associated with the pentons, are shown in blue and purple.

This figure was reproduced from Zhou *et al*, (1999).

packaged state of HCMV genomic DNA, the only difference being that HCMV genomic DNA was found to be packaged at a higher density than that of HSV-1, with an inter-layer spacing of 23 Å. The HCMV genome is 35% larger than that of HSV-1. However the volume of the HCMV capsid is only 17% greater, giving an explanation of why the HCMV genome is packaged at a greater density.

## **1.2.2. The capsid**

The HSV-1 capsid is an icosahedral protein shell with a diameter of 125 nm (Wildy *et al.*, 1960). Its function is to encase and protect the viral genome and, following infection, to deliver the viral genome to the nucleus of the cell. A reconstruction of the HSV-1 capsid determined to 8.5 Å resolution is shown in Figure 1.7.

### **1.2.2.1. Capsid forms found in infected cells**

Three distinct types of capsids have been observed by EM within the nuclei of infected cells. The three forms have been classified as A-capsids, B-capsids and C-capsids, and they can be separated by gradient centrifugation. The structure of the capsid shell is the same for all three forms, but they differ in their internal composition. A-capsids are empty capsid shells with no internal structure. B-capsids contain an internal proteinaceous core composed of the scaffolding proteins VP21 and VP22a, while C-capsids contain the packaged viral genome and lack the protein core of the B-capsid (Gibson and Roizman, 1972; reviewed in Steven and Spear, 1997).

It was originally proposed that B-capsids represent the intermediate in the capsid assembly/DNA packaging process and that the scaffolding protein within the B-capsid is replaced by the viral genome to produce the C-capsid. However, a fourth class of intranuclear particle has now been identified – the procapsid, which is now known to be the precursor of the other three forms. The procapsid was originally identified in cell-free systems but has also been detected in thin sections of infected cells (Newcomb *et al.*, 1996; Trus *et al.*, 1996; Rixon and McNab, 1999; Newcomb *et al.*, 2000). The shell of the procapsid is spherical rather than polyhedral as in the A, B and C capsids, and it is believed that the process of capsid maturation and angularisation occurs simultaneously with DNA packaging (Heymann *et al.*, 2003). If this model is correct then A and B capsids are dead-end products, and only the C-capsid represents a genuine intermediate in the virion assembly process. The procapsid will be discussed further in Section 1.5.1.

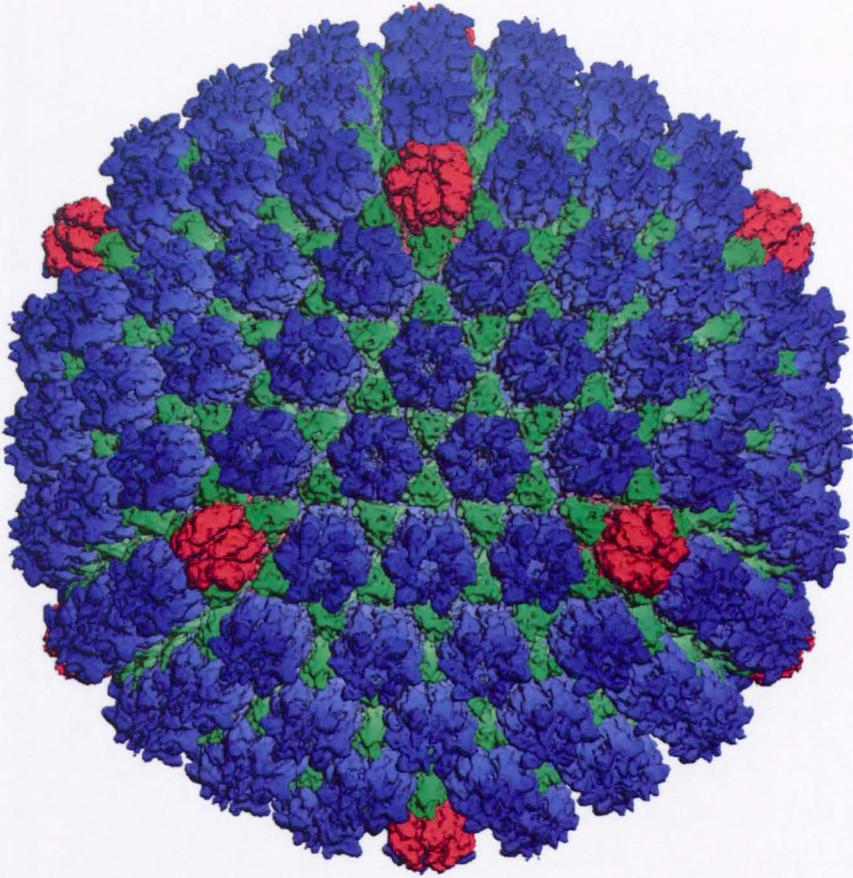


Figure 1.7: Structure of the HSV-1 capsid determined to 8.5 Å resolution.

The capsid shell has a diameter of 1250 Å and a MW of approximately 192 Mda. In this reconstruction it consists of 162 capsomeres: 150 hexons (blue) forming the faces and edges of the capsid and 12 pentons (red) forming the capsid vertices. Between adjacent capsomeres lie the capsid's 320 triplexes (green). The HSV-1 capsid conforms to  $T=16$  icosahedral symmetry. Only icosahedrally-ordered structures are shown in this reconstruction. Non-icosahedrally-ordered structures, such as the UL6 portal complex, are not shown.

This figure was reproduced from Zhou *et al*, (2000).

### 1.2.2.2. Protein composition of HSV-1 capsids

The HSV-1 B-capsid is composed of seven major and several minor proteins. The major proteins of the capsid are summarised in Table 1.2 and as described individually below. The structural arrangements of the proteins within the HSV-1 capsid will be discussed in Section 1.4.

#### VP5

VP5 (also see Section 1.4.4.1) is encoded by the UL19 ORF (Costa *et al.*, 1984; McGeoch *et al.*, 1988) and has a predicted molecular weight (MW) of 149075 Da. This figure correlates well with the observed MW of 150-155k Da, which indicates that VP5 does not undergo any significant post-translational modification (Gibson and Roizman, 1972; Cohen *et al.*, 1980; Rixon *et al.*, 1990). VP5 is also known as the major capsid protein. It forms both the hexon and the penton subunits and comprises approximately 70% by mass of the total capsid shell. Thus, it is not surprising that VP5 is essential for capsid formation. Weller *et al.* (1987) identified temperature-sensitive mutants of the UL19 ORF that prevent capsid formation at the non-permissive temperature. Furthermore, Desai *et al.* (1993) constructed a VP5 null mutant, in which the UL19 ORF was interrupted by the insertion of the *LacZ* gene. This mutant was unable to form capsids unless wild-type VP5 was supplied *in trans* by a complementing cell line. HSV-1 capsid assembly studies using the baculovirus expression system have also shown that VP5 is essential for capsid assembly (Newcomb *et al.*, 1994; Thomsen *et al.*, 1994; Tatman *et al.*, 1994).

VP5 forms extensive interactions with the other capsid proteins. It was Nicholson *et al.* (1994) who observed the migration of VP5 to the nucleus in the presence of preVP22a and first proposed an interaction between the two proteins. This interaction was later confirmed by the yeast two-hybrid system (Desai and Person, 1996). Further investigation of this interaction revealed that VP5 binds to the C-terminal 25 amino acids of scaffolding proteins preVP21 and preVP22a (Kennard *et al.*, 1995; Thomsen *et al.*, 1995; Desai and Person, 1996; Hong *et al.*, 1996) and that residues in the N-terminal 100 amino acids of VP5 are involved in mediating this interaction (Desai and Person, 1999; Warner *et al.*, 2001). VP5 has also been shown to interact with VP19C and VP26 (Nicholson *et al.*, 1994; Rixon *et al.*, 1996). Furthermore, an interaction between VP5 and a tegument protein (tentatively identified as VP1-3) has been revealed by cryo-EM (Zhou *et al.*, 1999).

Gene	Protein	MW (Da)	No. of copies per capsid	Location in capsid	Presence in A-, B-, and C-capsids
UL18	VP23	34268	630 or 640 <sup>†</sup>	Triplexes	A, B, C
UL19	VP5	149075	955 or 960 <sup>†</sup>	Hexons and Pentons	A, B, C
UL26	VP21	39875	100 <sup>†</sup>	Internal scaffold	B
UL26	VP24	26628	100 <sup>†</sup>	Internal scaffold	A, B, C
UL26.5	VP22a	33765	1000 to 1500 <sup>†</sup>	Internal scaffold	B
UL35	VP26	12095	900	Tips of hexons	A, B, C
UL38	VP19c	50260	315 or 320 <sup>†</sup>	Triplexes	A, B, C

Table 1.2: Summary of the HSV-1 major capsid proteins within HSV-1 A-, B-, and C-capsids.

\*. The UL6 portal protein complex may replace a single penton subunit. As a result, the actual copy number of VP23, VP19C and VP5 may be reduced (for example, in VP5, 955 copies per capsid as opposed to 960 copies).

†. Approximate numbers

References: Homa and Brown, 1997; Steven and Spear, 1997

Structurally, VP5 is the best characterised of the HSV-1 capsid proteins. Proteolytic digestion of VP5 has identified a large fragment believed to represent a discrete domain within the protein. The structure of this fragment has been determined using X-ray crystallographic techniques and shown to correlate to the outermost portion of VP5, referred to as the VP5 upper domain (VP5ud) (Baker *et al.*, 2003).

### **VP19C**

VP19C (also see Section 1.4.4.3) is encoded by the UL38 ORF (McGeoch *et al.*, 1988; Rixon *et al.*, 1990). It was originally designated VP19, but renamed VP19C in order to differentiate it from a 53 kDa glycosylated virion protein with which it co-migrated during SDS-PAGE (Heine *et al.*, 1974). It has a predicted MW of 50260 Da – a figure that correlates well with the observed MW of 50-54 kDa, indicating that VP19C does not undergo any significant post-translational modification (Gibson and Roizman, 1972; Cohen *et al.*, 1980; McGeoch *et al.*, 1988; Rixon *et al.*, 1990). VP19C forms part of the heterotrimeric complex known as the triplex, in which one copy of VP19C interacts with two copies of VP23. VP19C has also been shown to interact with VP5 (Rixon *et al.*, 1996; Saad *et al.*, 1999). VP19C is essential for capsid formation. Person and Desai (1998) constructed a null mutant of VP19C and demonstrated that capsids could not be formed unless wild-type VP19C was supplied *in trans* by a complementing cell line. HSV-1 capsid assembly studies using the baculovirus expression system have also shown that VP19C is essential for capsid assembly (Newcomb *et al.*, 1994; Thomsen *et al.*, 1994; Tatman *et al.*, 1994).

### **VP23**

VP23 (also see Section 1.4.4.3) is encoded by the UL18 ORF (McGeoch *et al.*, 1988; Rixon *et al.*, 1990). It has a predicted MW of 34268 Da; again, this is a figure that correlates well with the observed MW of 30-34k Da, indicating that VP23 does not undergo any significant post-translational modification (Gibson and Roizman, 1972; Cohen *et al.*, 1980; McGeoch *et al.*, 1988; Rixon *et al.*, 1990). Together with VP19C, VP23 forms the heterotrimeric triplex. VP23 is essential for capsid formation. Desai *et al.* (1993) constructed a null mutant of VP23 and demonstrated that capsids could not be formed unless wild-type VP23 was supplied *in trans* by a complementing cell line. HSV-1

capsid assembly studies using the baculovirus expression system have also shown that VP23 is essential for capsid assembly (Newcomb *et al.*, 1994; Thomsen *et al.*, 1994; Tatman *et al.*, 1994). Interestingly, although the formation of full-size (125 nm diameter) capsids requires the presence of VP23, VP5 and VP19C expressed in the absence of VP23 (and the scaffolding protein) interact to form smaller (88 nm diameter) particles with  $T=7$  icosahedral symmetry (Rixon *et al.*, 1996; Saad *et al.*, 1999). These results suggest that VP23 is potentially involved in controlling the dimensions of the assembling capsid.

## VP26

VP26 (also see Section 1.4.4.2) is encoded by the UL35 ORF (McGeoch *et al.*, 1988; Davison *et al.*, 1992). It has a predicted MW of 12095 Da. This figure correlates well with the observed MW of 12 kDa (Heilman *et al.*, 1979; Cohen *et al.*, 1980; McGeoch *et al.*, 1988; Rixon *et al.*, 1990). VP26 is located on the tips of the hexon subunits of the capsid, but not on the pentons. It has been shown to be present in A, B, and C-capsids, but not in the procapsid (Chi and Wilson, 2000; Newcomb *et al.*, 2000). Studies using the baculovirus expression system have shown that VP26 is not essential for capsid assembly (Newcomb *et al.*, 1994; Thomsen *et al.*, 1994; Tatman *et al.*, 1994). Furthermore, a VP26 null mutant (K $\Delta$ 26Z), which was made by replacing the UL35 ORF with the LacZ reporter gene, was able to replicate in cell culture to the same levels as WT virus (Desai *et al.*, 1998). Although VP26 is not essential for capsid assembly, it has been proposed to have a role in the production of infectious virus in nervous tissue. Desai *et al.* (1998) showed that when K $\Delta$ 26Z was grown *in vivo* in mouse trigeminal ganglia, the virus yield from infected cells was 30-100 fold less than with WT virus. They also showed that K $\Delta$ 26Z has reduced ability to reactivate from latent infection compared to WT virus. Douglas *et al.* (2004) demonstrated a role for VP26 in the retrograde transport of viral capsids towards the cell nucleus. They observed that 2-4 hours after infection, capsids containing VP26 move closer to the cell nucleus, while capsids lacking VP26 remained in a random distribution and proposed that although VP26 is not essential for virus replication *in vitro*, it is likely that that it is involved in the mediation of retrograde axonal transport *in vivo*.

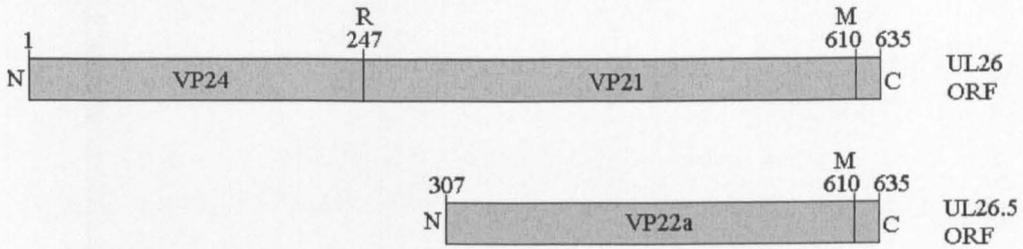
## VP21, VP24 and VP22a

VP21, VP24 and VP22a (also see Section 1.4.4.4) are scaffolding proteins. They are encoded by two overlapping in-frame genes: UL26 (which encodes VP21 and VP24) and UL26.5 (which encodes VP22a) (Liu and Roizman, 1991a and 1991b; Preston *et al.*, 1992)

(Figure 1.8). For proper formation of the HSV-1 capsid, the presence of either UL26 or UL26.5 is essential (Preston *et al.*, 1983; Gao *et al.*, 1994; Sheaffer *et al.*, 2000). However, if UL26 is absent, DNA packaging does not take place and infectious virus is not produced (Gao *et al.*, 1994). Production of infectious virus takes place in the absence of UL26.5, however this process is inefficient (Matusick-Kumar *et al.*, 1994). The role of the scaffolding proteins is to form the internal scaffold – a transient structure that directs the correct assembly of the procapsid shell. The scaffold disassociates and is removed from the capsid during maturation and DNA packaging (Dasgupta and Wilson, 1999; McClelland *et al.*, 2002). During this process the scaffolding proteins undergo proteolytic cleavage, as will be discussed further in Section 1.5.1.

The UL26 ORF initiates at nucleotide 50809 and terminates at nucleotide 52714 (McGeoch *et al.*, 1988; Davison *et al.*, 1992), encoding a 62466 Da protein. The N-terminus of the UL26 protein encodes a serine protease domain (Liu and Roizman, 1993) while the the C-terminus encodes an oligomerisation (Desai and Person, 1996; Pelletier *et al.*, 1997) and VP5 binding (Hong *et al.*, 1996; Oien *et al.*, 1997) domain. The UL26 protein is incorporated into the assembling capsid as a single polypeptide (Preston *et al.*, 1983; Kennard *et al.*, 1995; Thomsen *et al.*, 1995; Robertson *et al.*, 1996). During capsid maturation, the N-terminal protease domain self-cleaves the UL26 protein between amino acids A247 and S248 (known as the release site), resulting in the formation of the minor capsid proteins VP24 and preVP21, which constitute the N- and C- terminal domains of the UL26 protein respectively. Thus, VP24 is 247 amino acids in length while preVP21 is 387 amino acids in length (Liu and Roizman, 1993). During procapsid maturation, preVP21 is cleaved a second time by the VP24 protease, between amino acids A610 and S611 (known as the maturation site). This removes the C-terminal 25 amino acids of preVP21 to give VP21, which is 362 amino acids in length (Hong *et al.*, 1996; Oien *et al.*, 1997).

The UL26.5 ORF initiates at nucleotide 51727 (within the UL26 ORF) and like the UL26 ORF, terminates at nucleotide 52714 (McGeoch *et al.*, 1988), encoding the 40 kDa protein preVP22a. The primary sequence of preVP22a is identical to that of preVP21, with the exception that preVP21 has an additional 48 amino acids at the N-terminus. Therefore, preVP22a contains the C-terminal 328 amino acids of preVP21. preVP22a is incorporated into the assembling capsid and is cleaved by VP24 protease during procapsid maturation between the same two amino acids as preVP21 (amino acids A610 and S611 of VP21).



**Figure 1.8:** Structural organisation of the HSV-1 scaffolding genes.

The scaffolding proteins are encoded by the overlapping UL26 and UL26.5 ORFs. UL26 encodes the VP24 maturational protease (residues 1-247) and the VP21 minor scaffolding protein (residues 248-635). UL26.5 encodes the VP22a major scaffolding protein (residues 307-635). The R (release) and M (maturation) sites of proteolytic cleavage are indicated. Cleavage at the R site results in the release of proteins preVP21 and VP24. Cleavage at the M site results in the removal of 25 amino acids from both preVP21 and preVP22a to yield to mature forms of VP21 and VP22a respectively.

Cleavage removes the C-terminal 25 amino acids from preVP22a to give VP22a, which is 303 amino acids in length (Hong *et al.*, 1996; Oien *et al.*, 1997).

Differential regulation of UL26 and UL26.5 transcription results in preVP22a being expressed at higher levels than preVP21 and VP24. As a result, VP22a is present at much larger quantities in B-capsids than VP21 or VP24 (approximately 1000-1500 copies of VP22a per capsid, compared to approximately 100 copies of VP21 and VP24). Hence, VP22a is referred to as the major scaffolding protein (Newcomb *et al.*, 1993). VP21 and VP22a are released from the capsid during packaging, therefore VP21 and VP22a are found only in B-capsids. VP24 is retained within the capsid during procapsid maturation and is therefore a component of A, B, and C-capsids as well as mature virions (Gibson and Roizman, 1972; Spear and Roizman, 1972; Booy *et al.*, 1991; Stevenson *et al.*, 1997).

### **Minor Capsid Proteins**

A number of proteins have been identified as minor constituents of the HSV-1 capsid. These minor capsid proteins are encoded by the ORFs UL6, UL15, UL17, UL25, UL28, and UL33 (Homa and Brown, 1997). They are involved in the cleavage and packaging of the viral genome into the newly assembled capsid. The UL6 proteins have been shown to form the portal through which the viral genome enters and exits the capsid. The portal is located at a unique capsid vertex where it probably replaces a penton (Newcomb *et al.*, 2001a). It has also been demonstrated that the UL15 and UL28 proteins interact with the UL6 protein (White *et al.*, 2003), which leads to the conclusion that the minor capsid proteins assemble into a DNA cleavage and packaging complex located at a unique capsid vertex. Due to this arrangement, the minor capsid proteins are not arranged on the capsid with icosahedral symmetry. As a consequence, the location of these proteins and the structure of the DNA cleavage and packaging complex cannot be determined by cryo-EM and image reconstruction techniques, since these techniques require the subject material to have icosahedral symmetry (for example, Zhou *et al.*, 2000). UL17 is present in both the tegument and the capsid, and as the tegument is not attached to DNA-containing capsids in the nucleus (Booy *et al.*, 1991), it is likely that the UL17 protein participating in DNA packaging is bound to the capsid (Thurlow *et al.*, 2005). Both UL17 and UL25 appear to be present at multiple sites on the capsid, although their precise locations have not been determined (Ogasawara *et al.*, 2001; Thurlow *et al.*, 2006). It has recently been suggested that UL25 is located near the vertices on the external surface of the capsid (Newcomb *et al.*, 2006). UL33 is present in A, B, and C capsids, independent of the presence of UL6,

UL15 or UL28, and it has been suggested that it acts as part of the molecular machinery that translocates cleaved genomic DNA into the capsid (Beard *et al.*, 2004). The roles of minor capsid proteins in the packaging of the viral genome will be discussed further in Section 1.3.6.

### **1.2.3. The tegument**

#### **1.2.3.1. Composition of the tegument**

The tegument is the least-well understood component of the HSV-1 virion. It is an amorphous proteinaceous layer located between the capsid and the envelope (Figure 1.3). Its exact composition is uncertain, but it is thought to contain at least nineteen proteins (Table 1.3). The functions of many of these proteins are poorly understood. Some are important for the early stages of infection. The virion host shutoff protein (vhs, the gene product of UL41) shuts off host cell protein synthesis (reviewed by Smiley *et al.*, 2001). VP16 (UL48) acitvates HSV-1 immediate-early gene expression (Campbell *et al.*, 1984). The presence of these proteins in the tegument means that they can be released rapidly upon infection, without the delay that is inevitable when proteins have to be expressed *de novo* from the viral genome. This helps to ensure a rapid and efficient infection of the host cell. Other known functions performed by tegument proteins are summarised in Table 1.3.

The structural arrangement of tegument proteins is poorly understood. L-particles are non-infectious particles composed solely of envelope and tegument. Their existence shows that tegument proteins have an inherent ability to self-associate in the absence of capsids (McLauchlan and Rixon, 1992; Rixon *et al.*, 1992; Szilágyi and Berriman, 1994). Virus morphogenesis has been shown to proceed in the absence of several tegument proteins, such as US3, UL13, UL41, UL46, UL47, and UL49 (reviewed by Mettenleiter, 2002). Furthermore, over-expression of the tegument protein VP22 (the gene product of UL49) leads to increased amounts of the protein within the tegument, but does not appear to have any obvious impact on tegument structure (Leslie *et al.*, 1996). These observations suggest unlike the capsid, that the tegument does not have a precisely defined structure and exhibits flexibility in its arrangement. In addition, the level of protein synthesis can be a controlling factor in levels of incorporation into the tegument, as is the case for VP22 (Leslie *et al.*, 1996), while del Rio *et al.* (2005) used Green Fluorescent Protein (GFP) tagging to show that even without an increase in protein synthesis, the abundance of VP22 varies between individual virions. However, the organisation of proteins in the tegument is

Gene	Protein	MW (kDa)	Properties and Functions
UL4	-	60	Colocalises with UL3. Function not known.
UL11	gM	17-22	Myristylated, necessary for in envelopment and exocytosis
UL13	VP18.5	11	Protein kinase, phosphorylates ICP0, ICP22, VP22, vhs (UL41), gE (US8) etc.
UL14	-	24	Aids in cell to cell spread of virus
UL16	-	40	Associated with the capsid. Required for DNA packaging.
UL17	-	75	Required for DNA packaging.
UL21	-	58	Phosphoprotein
UL36	VP1-3	336	Implicated in capsid transport to nucleus, role in initiating tegument assembly around capsid
UL37	-	121	Cytoplasmic phosphoprotein, role in initiating tegument assembly around capsid
UL41	vhs	55	Virion host shutoff protein (vhs)
UL46	VP11/12	79	Phosphoprotein, modulates activity of UL48
UL47	VP13/14	74	Phosphoprotein, modulates activity of UL48
UL48	VP16	54	Transactivates IE gene expression
UL49	VP22	32	Phosphoprotein, secreted and exported to neighbouring cells
UL51	-	25	Function not known
UL56	-	-	Function not known
RL2	ICP0	110	IE gene transcription regulator
RS1	ICP4	133	IE gene transcription regulator
US3	-	53	Protein kinase, phosphorylates ICP22 and UL34, partially inhibits apoptosis
US10	-	34	Phosphoprotein, function unknown
US11	-	18	RNA binding protein, associates with 60s ribosomal subunit

**Table 1.3:** Summary of the HSV-1 tegument proteins and their known functions.

References: Leopardi *et al.*, 1997; Roizman and Knipe, 2001; Mettenleiter, 2002.

not simply random. Over-expression of the tegument protein UL37 did not result in any detectable increase in the amount of UL37 protein incorporated into virus particles (McLauchlan, 1997). More recently, work using cryo-electron tomography has revealed some interesting results regarding the structure and organisation of the tegument (Grünewald *et al.*, 2003). Most notably, the tegument was shown to have an asymmetric structure with respect to the capsid. On one side (the proximal pole), the tegument forms a very thin layer and the capsid is located in close proximity to the envelope, while on the other side of the virion (the distal pole) the capsid and envelope are separated by a thick layer of tegument. These results are interesting as they imply that the tegument and capsid layers do not move independently of each other, and the structure of the tegument contains some degree of organisation. The studies in question did not detect any structural features within the tegument other than filamentous structures believed to be actin filaments (Grünewald *et al.*, 2003). Therefore, while the tegument may not possess a single defined structure as is the case with the capsid, it does appear to have some level of structural organisation, although the exact nature of this organisation remains unclear.

Although the structure of the tegument as a whole is not well understood, cryo-EM examination of intact HSV-1 virions has identified regions of icosahedrally ordered tegument protein that form extensive interactions with the penton subunits of the capsid (Figure 1.9) (Zhou *et al.*, 1999). Calculations based on these reconstructions estimate that this icosahedrally ordered tegument protein has a molecular mass of 170-200 kDa and is approximately 200Å in length. While the identity of this protein has not been confirmed, the most likely candidate is VP1-3 – a tegument protein that is so closely associated with the capsid that it was originally identified as a capsid protein (Gibson and Roizman, 1972).

The function of VP1-3 is not well understood, but it is known to be essential. Studies of a temperature-sensitive mutant of UL36 (the gene that encodes VP1-3), known as *tsB7*, have revealed that at the non-permissive temperature, the release of viral DNA from the capsid is blocked (Knipe, *et al.*, 1981; Batterson *et al.*, 1983). More recent work has demonstrated the presence of a portal in the capsid, which is formed by the UL6 protein. The portal is thought to be the location at which HSV-1 genomic DNA enters and exits the capsid, and it has been demonstrated that the UL6 portal complex is located at a unique capsid vertex (Newcomb *et al.*, 2001a). While the nature of the interaction between UL6 and VP1-3 remains unclear, the implication that VP1-3 is involved in DNA exit from the capsid makes it appropriate for VP1-3 to be associated with the portal vertex, however association with the portal may not imply association with the other vertices.

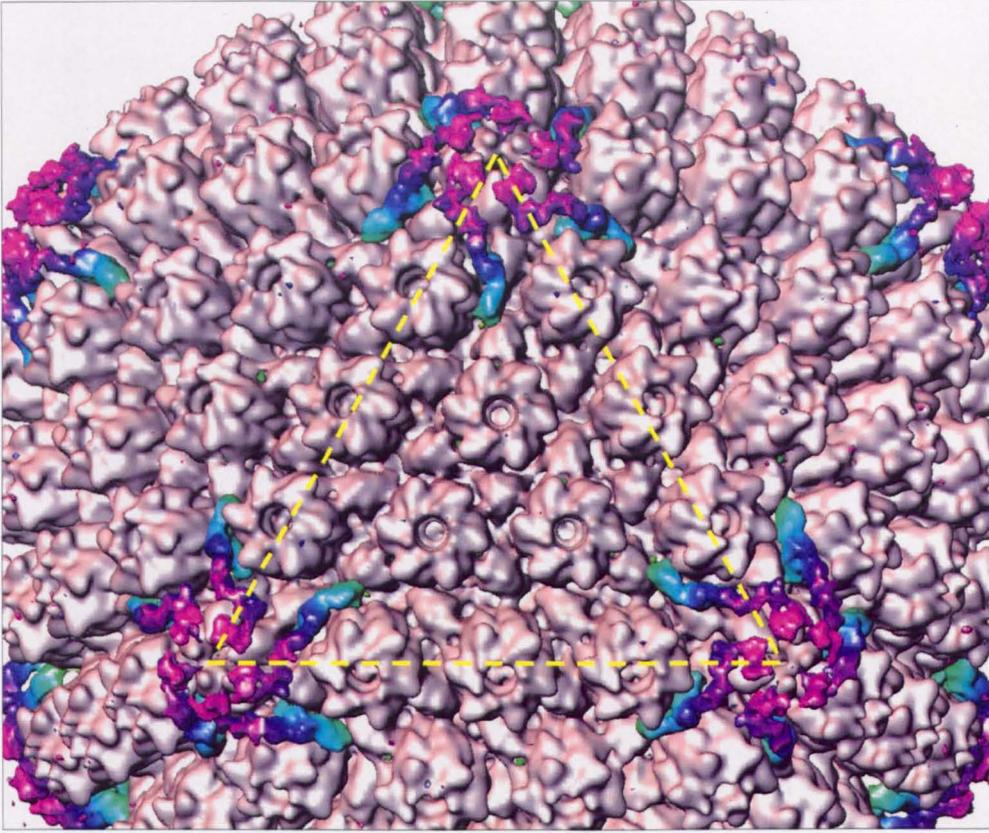


Figure 1.9: Visualisation of icosahedrally ordered tegument proteins associated with the HSV-1 capsid. The structure of the HSV-1 B-capsid is shown in grey. Associated tegument proteins are highlighted in colour. The dimensions of a single icosahedral capsid face are outlined by the broken yellow line. The resolution of this structure is 20 Å. This figure shows that the tegument densities are associated with only the capsid vertices (the pentons), and not with any other capsid subunit (the hexons or triplexes).

This figure was reproduced from Zhou *et al.*, (1999).

There are arguments against VP1-3 being the tegument protein that is associated with the pentons. It has been reported that the virion contains between 120-200 copies of VP1-3 (Heine *et al.*, 1974), which is more than the 55-60 copies that would be required if this protein bound to the capsid solely at the pentons in a 1:1 ratio with VP5 as suggested by cryo-EM. Furthermore, VP1-3 is known to be a component of L-particles, which lack capsids (Szilágyi and Cunningham, 1991). Although these observations are not sufficient evidence to dismiss VP1-3 as a candidate for the penton-associated protein, they do suggest that it does not interact exclusively with the capsid and that it is also a structural component of the tegument.

Icosahedrally ordered tegument proteins have been identified in HCMV and simian cytomegalovirus (SCMV) – both *Betaherpesvirinae* (Trus *et al.*, 1999; Chen *et al.*, 1999). The interactions in HCMV and SCMV differ significantly from those observed in HSV-1. While tegument/capsid interactions are restricted to pentons in HSV-1, in HCMV and SCMV extensive contacts are formed all over the surface of the capsid. In SCMV, two distinct species of tegument protein are clearly seen to be binding to the capsid; one to the major capsid protein and the other to the triplex. The tegument mass attached to the major capsid protein has been estimated at approximately 50-60 kDa, while the mass attached to the triplex has been estimated at approximately 100-120 kDa – both values considerably lower than the 170-200 kDa mass of the tegument proteins attached to the HSV-1 capsid. What is striking from these comparisons is that the organisation of the tegument in HSV-1 is completely different from that observed in both HCMV and SCMV – the proteins have different shapes, sizes and binding patterns with respect to the capsid. Trus *et al.* (1999) and Chen *et al.* (1999) proposed a number of proteins as candidates for the tegument proteins that bind to capsids in HCMV and SCMV. These proteins have been identified on the basis of the size of the protein and the abundance of the protein within the virion/tegument. The most likely candidate for the tegument protein that binds to the major capsid protein is the upper matrix protein (UM) (69 kDa, the gene product of UL82), which functions as a transactivator of immediate-early gene expression in HCMV (Baldick *et al.*, 1997). The candidate proposed for the tegument protein that binds to the triplex is the basic phosphoprotein (BPP) (119 kDa), which has been implicated in the transport of the capsid in the cytoplasm (Meyer *et al.*, 1997). Neither BPP nor UM have been shown to have any counterparts in HSV-1 (Chee *et al.*, 1990). The conclusion drawn from these results is that the structure and function of the HCMV/SCMV capsid-associated tegument proteins are entirely different to those of HSV-1.

It has also been suggested that SCMV capsids form interactions with the high-molecular-weight protein (HMWP), a 205 kDa protein, which is the homologue of HSV-1 VP1-3. SDS-Page analysis revealed small amounts of HMWP associated with cytoplasmic B-capsids isolated from SCMV infected cells (Chen *et al.*, 2001; Trus *et al.*, 1999). VP1-3 is poorly conserved; nevertheless, homologues have been identified in all mammalian and avian herpesviruses. These observations indicate an essential and possibly identical role for the protein across *Alpha*, *Beta* and *Gammaherpesvirinae*. Since HSV-1 VP1-3 has been identified as the most likely candidate to be the tegument protein associated with the HSV-1 capsid, it has been proposed that HMWP binds in a similar manner in SCMV capsids (Chen *et al.*, 2001). Thus, although the bulk of the icosahedrally ordered tegument proteins in HCMV and SCMV are likely to be structurally and functionally unrelated to those of HSV-1, it is plausible that VP1-3 homologues bind to capsids and function in a similar manner in all three viruses.

#### **1.2.4. The envelope**

The envelope forms the outermost boundary of the virion (Figure 1.3), and like the tegument, it is a poorly understood layer, proving intractable to structural analysis. It is derived from the host cell late in infection. There are at least eleven virally encoded glycoproteins in HSV-1 (Table 1.4). Although no glycoprotein has been shown to be essential for virion assembly, four are essential in the initial infection of the host cell (Steven and Spear, 1997).

The lack of symmetry in the envelope prevents the use of symmetry-based cryo-EM and image reconstruction techniques. However, cryo-EM of virions and L-particles has been used to give some limited structural information concerning the envelope. The membrane has been shown to be spherical and to contain spike-like projections of approximately 15 nm in length protruding from the surface. These presumably correspond to the glycoproteins (Szilágyi and Berriman, 1994). Negative staining has been employed to directly visualise the glycoproteins and structures ranging from 80Å to 140Å in length are clearly visible in the membrane. However, these preparations provide no information regarding the organisation of the glycoproteins within the undamaged membrane (Stannard *et al.*, 1987).

Gene	Protein	MW (kDa)	Properties and Functions
UL1	gL	40	Forms complex with gH, required for viral penetration and cell fusion
UL10	gM	60	Phosphorylated, role in acquisition of the viral envelope in the cytoplasm (secondary envelopment)
UL22	gH	115	Forms complex with gL, required for viral penetration and cell fusion
UL27	gB	120	Forms homooligomers, required for viral penetration and cell fusion
UL44	gC	120	Responsible for initial attachment to heparin sulphate
UL53	gK	40	Involved in cell fusion. Potential role in secondary envelopment in the cytoplasm
US4	gG	60	Required for entry into polarised cells, involved in viral egress and cell to cell spread
US5	gJ	10	Non-essential for viral replication in cultured cells
US6	gD	60	Required for receptor-binding, viral penetration and cell fusion
US7	gL	70	Forms complex with gE, role in acquisition of the viral envelope in the cytoplasm (secondary envelopment)
US8	gE	80	Forms complex with gI, role in acquisition of the viral envelope in the cytoplasm (secondary envelopment), involved in cell-to-cell spread

**Table 1.4:** Summary of the HSV-1 glycoproteins and their known functions.

References: Haarr and Skulstad, 1994; Steven and Spear, 1997; Subak-Sharpe and Dargan, 1998; Roizman and Knipe, 2001; Mettenleiter, 2002; Farnsworth, *et al*, 2003

Cryo-electron tomography has been used to provide some interesting results regarding the organisation of the envelope glycoproteins (Grünewald *et al.*, 2003). The envelope was shown to contain 650-700 glycoprotein spikes. The glycoprotein spikes ranged in length from 100Å to 250Å and also varied in morphology. Interestingly, the glycoproteins were shown to not be randomly distributed, with an accumulation of spikes at the distal pole, where the tegument is thickest. This has been proposed to be a result of lipid raft involvement in envelope assembly (Grünewald *et al.*, 2003).

### **1.3. Viral replication**

The stages of the HSV-1 life cycle, from initial infection to the release of infectious progeny virus, are discussed below and summarised schematically in Figure 1.10.

#### **1.3.1. Virus attachment and entry**

The entry of HSV-1 into cells occurs in at least three stages: (i) the attachment of the virion to the surface of the cell, (ii) the interaction of gD with one of several cellular receptors, and (iii) fusion of the viral envelope and plasma membrane to release the capsid/tegument structure into the cytoplasm. Most of the information currently available on this process is based on studies of non-polarised cells, in which the membrane proteins interacting with envelope proteins are most likely to be randomly distributed. Attachment and entry to polarised cells, such as most epithelial cells in the human body, may differ in detail.

The initial attachment of the virus particle to the cell is mediated by the interaction of glycoproteins gC and, to a lesser extent gB, with the glycosaminoglycans (GAGs) on the cell surface (WuDunn and Spear, 1989; Shieh *et al.*, 1992). Interestingly, although gC confers the greatest efficiency for attachment, it is not essential for either entry or viral replication (Heine *et al.*, 1974). In cases where gC is absent, binding to GAG is mediated by glycoprotein gB. The preferential GAG for this interaction is heparin sulphate. Nevertheless, cells devoid of heparin sulphate but expressing other GAGs such as chondroitin sulphate can also be infected (Banfield *et al.*, 1995).

The second stage of viral attachment involves the interaction of gD with one of several cell-surface receptors. There are three classes of cell-surface molecule that are known to serve as receptors for HSV-1 attachment: (i) herpesvirus entry mediator (HVEM or HveA), which is a member of the tumour necrosis factor receptor family (Montgomery *et al.*,



HSV-1 virion

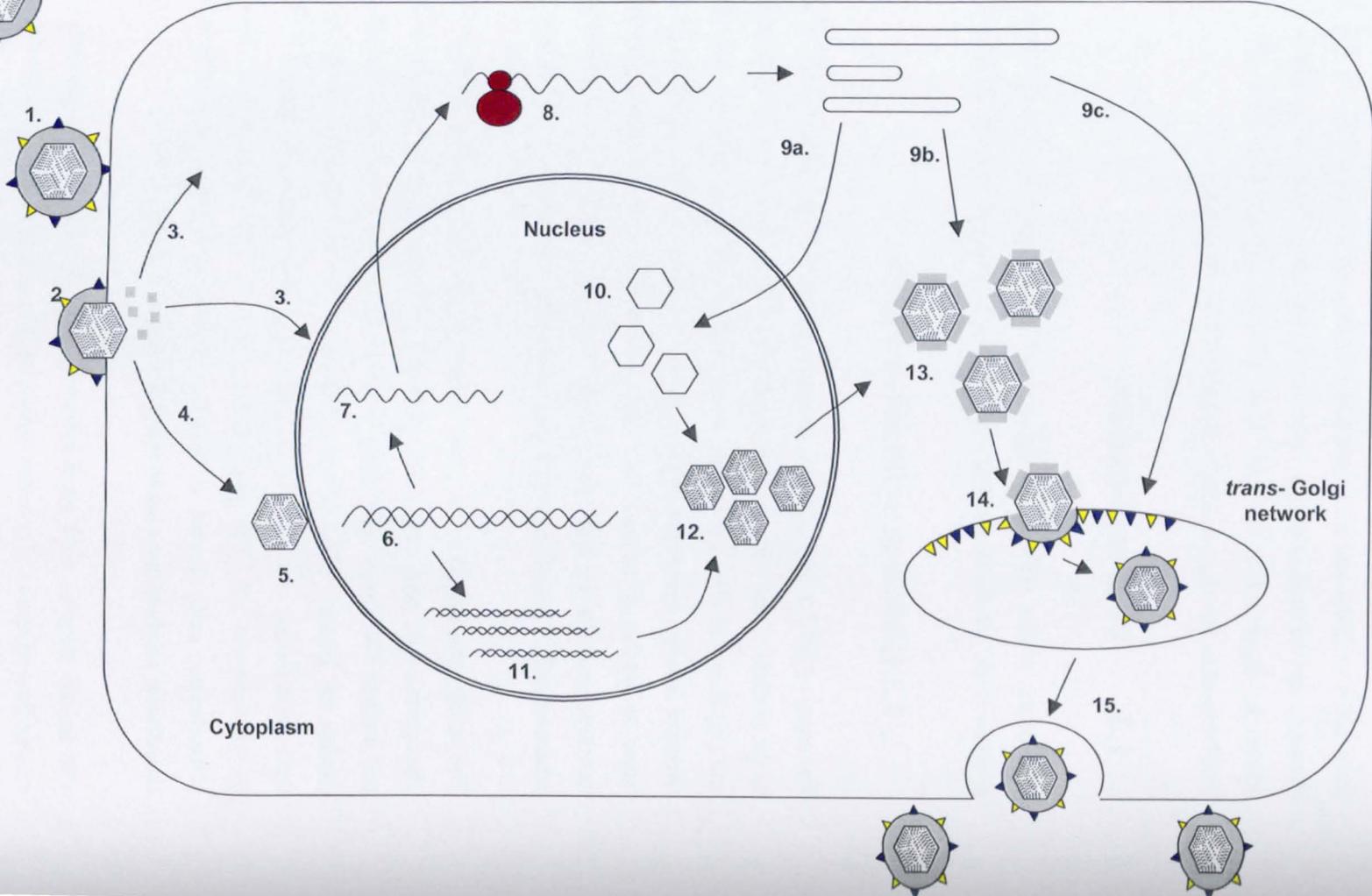


Figure 1.10: Overview of the HSV-1 lytic cycle

(1) Virions attach to the cell surface via specific interactions between viral glycoproteins (gB and gC) and the extracellular matrix. (2) Interaction of gD with cellular receptors initiates fusion of the cellular and viral membranes and the releases the capsid and tegument proteins into the cytoplasm. (3) Tegument proteins are either retained in the cytoplasm, where they influence host cell responses, or are transported to the nucleus where they are involved in the initiation of viral gene expression. (4) The nucleocapsid is transported to the nucleus via microtubules, and (5, 6) docks at the nuclear pore, where the viral genome is released into the nucleus. (7) HSV-1 gene expression is activated by the VP16 tegument protein. Viral genes are expressed in a tightly-regulated cascade. (8) Following transcription, viral mRNAs are exported to the cytoplasm for translation on cellular ribosomes. Virally-encoded proteins are targeted to the appropriate sites within the infected cell; (9a) some IE proteins, and those involved in DNA replication or capsid assembly are required within the nucleus. (9b) Other proteins, such as tegument proteins, remain in the cytoplasm. (9c) Viral glycoproteins enter the Golgi where they undergo post-translational modification. (10) Capsid proteins assemble in the nucleus, forming procapsids. (11) DNA replication initially takes place by the theta mode of replication, but soon changes to a rolling-circle mode of replication. The vast majority of progeny genomes form concatameric molecules. (12) These are cleaved to form monomeric unit-length genomes and are packaged into the procapsid. During packaging the procapsid undergoes extensive structural remodelling to form a mature polyhedral capsid. (13) DNA-containing capsids exit the nucleus by budding into the perinuclear space and then into the cytoplasm. Tegument assembly takes place in the cytoplasm, and is initiated by the interaction of tegument protein VP1-3 with the penton subunits of the capsid. (14, 15) Tegumented capsids acquire an envelope by budding into the trans-golgi network. This process involves the interaction of the tegument proteins with the cytoplasmic tails of viral glycoproteins gD, gE/gI, and gM. (16) Progeny virions are released from the infected cell at the membrane by exocytosis.

1996), (ii) nectin-1 (HveC) or nectin-2 (HveB), two members of the Immunoglobulin superfamily (Geraghty *et al.*, 1998; Warner *et al.*, 1998), and (iii) specific sites in heparin sulphate generated by certain isoforms of 3-O-sulphotransferases (Shukla *et al.*, 1999). Any one of these cell surface molecules can bind to gD to mediate viral entry, with HSV-1 and HSV-2 having different receptor preferences. For example, HVEM and nectin-1 are excellent entry receptors for both HSV-1 and HSV-2, nectin-2 is more effective for HSV-2 than for HSV-1 and 3-O-sulphated heparin sulphate is probably more effective for HSV-1 than for HSV-2. (Roizman and Knipe, 2001; reviewed by Spear and Longnecker, 2003). Viruses lacking gD are unable to interact with cell-surface receptors, therefore gD is essential for the entry of virus into cells (Johnson and Ligas, 1988).

The final stage of HSV-1 entry is the fusion of the viral envelope with the cellular membrane. This process requires the involvement of glycoproteins gD, gB, and the heterodimer gH-gL in a manner yet to be clarified (reviewed by Campadelli-Fiume *et al.*, 2000). Bender *et al.* (2003) showed that cholesterol-rich lipid rafts may be involved in the fusion event through an interaction with glycoprotein gB. It has also been demonstrated that as an alternative to fusing directly with the cellular membrane, a second pathway for HSV-1 entry exists, in which after binding to the cell surface, the virus particle is taken into the cell via endocytosis and membrane fusion occurs in a pH-dependent manner (Nicola *et al.*, 2003). The significance of these findings is uncertain, but it has been suggested that HSV-1 may use different methods of cell entry depending on the cell type that it is infecting.

### **1.3.2. Transport to the cell nucleus**

Following fusion of the viral envelope with the cellular membrane, the HSV-1 capsid and tegument proteins are released into the cytoplasm of the host cell. Since viral DNA replication occurs in the nucleus of the infected cell, the capsid has to be transported through the cytoplasm to the nucleus. High concentrations of protein along with the presence of the cytoskeleton and organelles can cause molecular crowding in the cytoplasm, which effectively restricts free diffusion of molecules larger than 50 kDa (Luby-Phelps, 2000). This indicates that virions are transported by active processes. It has been shown that the HSV-1 capsid is transported through the cytoplasm by an active process involving the microtubule network of the cell. In the presence of microtubule depolymerising agents such as colchicine, vinblastine or nocodazole, there was a significant reduction in the concentration of capsids observed at the nuclear rim, with the

majority remaining scattered throughout the cytosol (Sodeik *et al.*, 1997). A more detailed study of this process has shown that HSV-1 is transported by interaction with dynein (a minus end-directed microtubule motor protein) and dynactin (a co-factor of dynein) (Döhner *et al.*, 2002). The HSV-1 proteins that are responsible for mediating this reaction are currently unknown but are likely to involve outer capsid or inner tegument proteins. An interaction between the dynein intermediate chain (DIC) and UL34 has been reported (Ye *et al.*, 2000), however the UL34 protein is absent from mature virions in HSV-1 (Reynolds *et al.*, 2002) and PRV (Klupp *et al.*, 2000).

Recent work by Douglas *et al.* (2004) demonstrated an interaction between VP26 and dynein light chains RP3 and Tctex1. This interaction was first identified using the yeast two-hybrid method. Subsequently, Glutathione S-transferase pull-down assays confirmed the binding of VP26 to RP3, Tctex1 and intact cytoplasmic dynein complexes. Furthermore, they observed that 2-4 hours after infection, micro-injected capsids containing VP26 had moved closer to the cell nucleus while capsids lacking VP26 remained in a random distribution, prompting authors to conclude that VP26 mediates the binding of incoming HSV-1 capsids to cytoplasmic dynein during cellular infection, through interactions with the dynein light chain.

Other potential candidates for the microtubule binding protein are US11, UL25 and VP22 (Elliott and O'Hare, 1998; Kaelin *et al.*, 2000; Diefenbach *et al.*, 2002). Of these, the minor capsid protein UL25 is possibly the best candidate since it is known that UL25 is closely associated with the capsid and also that it remains with the capsid until it reaches the nucleus (Kaelin *et al.*, 2000). The large tegument protein VP1-3 is another potential candidate. Although it has not been demonstrated that VP1-3 interacts with microtubules, it has been shown that VP1-3 is very tightly associated with the pentons of the capsid (Zhou *et al.*, 1999). Sodeik *et al.* (1997) showed that the interaction with microtubules appeared to involve the pentons, correlating with the observed localisation of VP1-3.

Following transport through the cytoplasm, the capsid docks at the nuclear pore complexes (NPCs), possibly through an interaction between importin  $\beta$  of the NPC and an HSV-1 tegument protein such as VP1-3 (Ojala *et al.*, 2000). Following binding to the NPC, the viral DNA is released from the capsid and injected into the nucleus where it localises adjacent to nuclear ND10 domains and circularises (Maul *et al.*, 1996). Recent evidence indicates that although viral genomes can associate with pre-existing ND10 domains,

structures resembling ND10 domains form *de novo* and associate with viral genome complexes during the initial stages of infection (Everett and Murray, 2005).

### **1.3.3. Disruption of host cell protein synthesis**

The HSV-1 virion contains a number of regulatory proteins that prime the newly-infected cell to support efficient virus replication. These proteins are located in the tegument, and are delivered to the cytoplasm after fusion with the host cell plasma membrane. Among these proteins is the virion host shutoff protein (vhs), which triggers the rapid shutoff of host cell protein synthesis by accelerating the rate of degradation of both cellular and viral mRNAs in a non-specific manner (Kwong and Frenkel, 1987; Oroskar and Read, 1987; Elgadi *et al.*, 1999). However, although viral mRNA is degraded, its rate of synthesis is greater than the rate of vhs-mediated degradation (Kwong *et al.*, 1988; Elgadi *et al.*, 1999). It has been shown that vhs interacts with the cellular translation initiation factor eIF4H, which presumably serves to target vhs to mRNA molecules (Feng *et al.*, 2001), and also that vhs possesses nuclease activity, and therefore degrades mRNAs directly rather than targeting them for cellular-mediated degradation (Everly *et al.*, 2002).

The vhs protein is not essential for viral replication, since mutants lacking the UL41 ORF are viable. However, mutants lacking functional vhs show a marked reduction in virus growth in cell culture (Read and Frenkel, 1983; Read *et al.*, 1993), and wild-type virus rapidly outgrows vhs mutants in mixed infections (Kwong *et al.*, 1988).

Three lines of evidence demonstrate that the vhs protein is necessary and sufficient for virion-induced host shutoff. Read and Frenkel (1983) isolated viable HSV mutants deficient in virion-induced shutoff, and one of these mutants (*vhs1*) was subsequently mapped to the UL41 ORF (Kwong *et al.*, 1988; McGeoch *et al.*, 1988). The ORF encoding vhs was confirmed by targeted disruptions of the UL41 gene, which produced a vhs-deficient phenotype (Fenwick and Everett, 1990b). Secondly, viral recombinants in which the HSV-1 UL41 gene was replaced with the corresponding gene of HSV-2 display the more robust shutoff phenotype characteristic of HSV-2 (Fenwick and Everett, 1990a). Thirdly, vhs is sufficient to block reporter gene expression when it is expressed as the only HSV-1 protein in transiently transfected mammalian cells (Jones *et al.*, 1995; Pak *et al.*, 1995).

It should be noted that vhs is not the only mechanism employed by HSV-1 in the control of host cell protein synthesis. The IE protein ICP27 inhibits the splicing of cellular pre-mRNAs by disrupting spliceosomes and redistributing cellular small-nuclear ribonucleoproteins (snRNPs) (Phelan *et al.*, 1993; Hardy and Sandri-Goldin, 1994; Sandri-Goldin *et al.*, 1995; Bryant *et al.*, 2001). Since the majority of HSV-1 genes are intron-less and do not require splicing, this is an effective mechanism for reducing host cell protein synthesis. Four HSV-1 genes contain introns and therefore require splicing prior to translation. Three of these genes are IE genes (ICP0, ICP22, and ICP47), which are transcribed and processed before ICP27 can effectively inhibit the splicing of pre-mRNAs. Late in infection, ICP27 is localised to replication compartments. At this stage it may no longer repress RNA splicing, allowing UL15 RNA to be spliced (Roizman and Knipe, 2001). UL15 is the fourth HSV-1 gene that contains introns and is expressed late in infection.

Another mechanism by which HSV-1 might disrupt host cell protein synthesis is through the action of ICP22 and the UL13 protein kinase. These proteins have been shown to phosphorylate cellular RNA polymerase II (Rice *et al.*, 1995; Long *et al.*, 1999), although it is uncertain how this affects protein synthesis since this enzyme is also required for transcription of the viral genome.

#### **1.3.4. HSV-1 gene expression**

Transcription of viral DNA takes place in the nucleus. The host RNA polymerase II is responsible for the transcription of all viral DNA during infection (Alwine *et al.*, 1974; Costanzo *et al.*, 1977). The resulting mRNAs are then transported into the cytoplasm for translation on cellular ribosomes. During productive infection, 74 HSV-1 genes are expressed in a highly regulated cascade fashion (Hones and Roizman, 1974; Roizman and Knipe, 2001).

Very soon after infection (within two to four hours), the viral immediate-early ( $\alpha$  or IE) genes are expressed (Roizman and Knipe, 2001). Five such genes have been identified encoding ICP4 (RS1), ICP0 (RL2), ICP27 (UL54), ICP22 (US1), and ICP47 (US12). Expression of these genes is controlled by the action of VP16, a virion trans-activator protein. VP16, encoded by the UL48 gene, forms a complex with cellular proteins and mediates IE gene expression through a common regulatory element found in the promoters of each of the IE proteins (often designated as TAATGARAT – the best conserved portion

of the consensus sequence involved). Activation of the HSV-1 IE genes is initiated by the formation of a multi-protein complex containing the virion protein VP16, and two host cell proteins: oct-1 (octamer DNA-binding protein) and HCF (Host Cell Factor). The complex binds to TAATGARAT elements present in each IE gene promoter (Weir *et al.*, 2001). As IE gene expression is mediated by pre-existing viral and cellular proteins, it does not require *de novo* protein synthesis, thus allowing IE genes to be activated rapidly after infection. The function of the IE genes is primarily to serve as transactivators of E and L gene expression. However, as discussed in the previous section, ICP27 and ICP22 are also involved in the shutoff of host cell protein synthesis.

The second set of HSV-1 genes to be expressed are the early genes ( $\beta$  or E). These genes require the presence of functional IE protein ICP4 (Watson and Clements, 1980), but not the onset of viral DNA synthesis, and their expression peaks five to seven hours after infection (Hones and Roizman, 1974). Many early proteins promote viral DNA replication, and the expression of late genes is then stimulated. Two general groups of early proteins have been identified. The  $\beta_1$  genes are expressed shortly after the onset of synthesis of  $\alpha$  proteins. Examples include UL29, which encodes ICP8 – the single-stranded DNA binding protein, and UL39, which encodes ICP6 – the large subunit of ribonucleotide reductase (Huszar and Bacchetti, 1981). The  $\beta_2$  genes are expressed with more delay after  $\alpha$  protein expression, and are exemplified by UL23, which encodes thymidine kinase. Some  $\beta$  genes require ICP27 for their expression (Samaniego *et al.*, 1995; McGregor *et al.*, 1996; Uprichard and Knipe, 1996), and this may correlate with the later kinetics that are characteristic of  $\beta_2$  genes (Roizman and Knipe, 2001).

The final class of viral genes to be expressed are the late ( $\gamma$  or L) genes. The L genes primarily encode the structural proteins of the virion, such as the capsid proteins (UL19, UL18 etc), the scaffolding proteins (UL26 and UL26.5), the viral glycoproteins (gB, gD etc), and the proteins of the tegument (vhs, VP16 etc) (Roizman and Knipe, 2001; Weir, 2001). Expression of L genes begins concurrently with the initiation of DNA replication (Weir, 2001). They can be sub-divided into leaky-late ( $\gamma_1$ ) and true-late ( $\gamma_2$ ) genes. True-late genes, such as UL38 and UL25, are those that require DNA replication for any significant expression of the protein, while leaky-late genes, such as UL19 and UL18) can be expressed in the absence of DNA replication (Roizman and Knipe, 2001). L gene expression peaks eight to ten hours after infection, then continues for the remainder of the lytic replication cycle (Harris-Hamilton and Bachenheimer, 1985).

### **1.3.5. Viral DNA replication**

Viral DNA replication commences shortly after early gene expression, and the viral genome and several early gene products localise within the nucleus at specific sites known as prereplicative sites, located near ND10 structures (Ishov and Maul, 1996; Uprichard and Knipe, 1996, Sourvinos and Everett, 2002; Everett and Murray, 2005). Seven viral proteins are required for viral DNA replication. These are (i) the viral DNA polymerase (encoded by UL30) and (ii) its accessory protein or processivity factor (UL42), (iii) an origin-binding protein (UL9), (iv) the single-stranded DNA-binding protein (UL29), and (v)-(vii) a helicase-primase complex of three proteins (UL5, UL8 and UL52) (Boehmer and Lehman, 1997; Lehman and Boehmer, 1999; Roizman and Knipe, 2001). A number of host cell proteins that are involved in DNA replication, such as topoisomerase II and DNA ligase, are also likely to be involved in the replication of the viral genome (Roizman and Knipe, 2001).

A model for HSV-1 DNA replication has been proposed, in which initiation occurs at one of three cis-acting (ori) elements located within the HSV-1 genome. UL9, the origin-binding protein binds to one of these elements and begins to unwind the DNA. UL9 then recruits ICP8, the single-stranded DNA-binding protein, which binds to the unwound single-stranded DNA. UL9 and ICP8 then recruit the five remaining viral DNA replication proteins to the replication forks. The helicase-primase proteins and the viral polymerase complex assemble at each replication fork. It has been suggested that replication is initiated by a bi-directional, theta mode of replication. However, there is a rapid switch by an unknown mechanism to a rolling-circle mode of replication. UL9 is not necessary for rolling-circle replication, which is not origin-dependent. Due to this rapid change to rolling-circle replication, the majority of progeny viral genomes are in the form of long head-to-tail concatamers, which are cleaved into monomeric molecules during packaging (reviewed by Boehmer and Lehman, 1997; Roizman and Knipe, 2001).

### **1.3.6. DNA packaging**

Following DNA replication, individual viral genomes are cleaved from concatamers, and each is packaged into a preformed capsid (Homa and Brown, 1997). This step involves extensive interactions between the capsid, DNA and the packaging machinery. Genetic analyses have identified a total of seven HSV-1-encoded proteins, i.e. the products of genes UL6, UL15, UL17, UL25, UL28, UL32 and UL33, which are demonstrated to be

necessary for DNA packaging (Table 1.5). When cells are infected with HSV-1 mutants lacking the function of any of these seven genes, capsid formation and DNA replication occur normally, but no packaging takes place (Weller, 1995; Homa and Brown, 1997; Newcomb *et al.*, 2001a).

The portal forms the channel through which genomic DNA enters HSV-1 capsids (Newcomb *et al.*, 2001a). It is composed of a dodecamer of UL6 molecules, which form a ring-like structure (Figure 1.11) found at a single vertex in the icosahedron, replacing one of the pentons (Trus *et al.*, 2004). Recent studies have demonstrated that portal containing capsids only form if the portal is present at the outset of assembly, indicating that the portal is incorporated into the capsid at the initiation of assembly. If portals are not present, assembly is initiated in another, as yet undefined way (Newcomb *et al.*, 2005).

The UL15 and UL28 proteins have been shown to interact with each other (Yu and Weller, 1998; Krosky *et al.*, 1998; Koslowski *et al.*, 1999; Abbotts *et al.*, 2000). UL15 has been shown to share limited sequence homology with the ATPase subunit of the bacteriophage T4 terminase (Davison, 1992), and there is evidence that the resulting complex may be analogous to the terminase of double-stranded DNA bacteriophage. This enzyme usually consists of two subunits, binds to specific sequences in the bacteriophage genome and acts as a molecular motor, delivering the DNA through the portal complex. The terminase initiates the translocation of DNA into the capsid by cleaving replicated concatameric viral DNA at a specific site, and provides energy for the packaging process through the hydrolysis of ATP. It terminates the process by a second cleavage event (Fujisawa and Morita, 1997; Catalano *et al.*, 2000).

Further evidence that supports the role of UL15/UL28 as the HSV-1 terminase comes from a demonstration that both UL15 and UL28 interact with the UL6 portal protein (White *et al.*, 2003) – an interaction that would be expected if UL15 and UL28 were to function as the HSV-1 terminase. The UL33 protein has also been shown to form a complex with UL15 and UL28 in infected cells, and has been proposed as a third subunit of the HSV-1 terminase (Beard *et al.*, 2002).

While procapsids contain the UL15 and UL28 proteins, the levels of both decrease during capsid maturation. In contrast, the amount of UL25 protein present was seen to increase during maturation (McNab *et al.*, 1998; Sheaffer *et al.*, 2001). Levels of the UL6 protein are similar in procapsids, B-capsids and C-capsids (Patel and Maclean, 1995; Sheaffer *et*

Gene	MW (kDa)	Present in mature virion	Properties and Functions
UL6	74	Yes	Portal protein; forms the channel through which the DNA enters the capsid.
UL15	81	No	Putative terminase subunit, interacts with UL28, UL33 and the UL6 portal protein. Homologous to bacteriophage T4 terminase subunit gp15
UL17	77	Yes	No function assigned to date
UL25	62.5	Yes	Proposed to have a role in sealing the capsid following completion of DNA packaging.
UL28	85.5	No	Putative terminase subunit, interacts with UL15, UL33 and the UL6 portal protein
UL32	64	No	No function assigned to date
UL33	19	No	Putative terminase subunit, interacts with UL15 and UL28

Table 1.5: Summary of the HSV-1 DNA packaging proteins and their known functions.

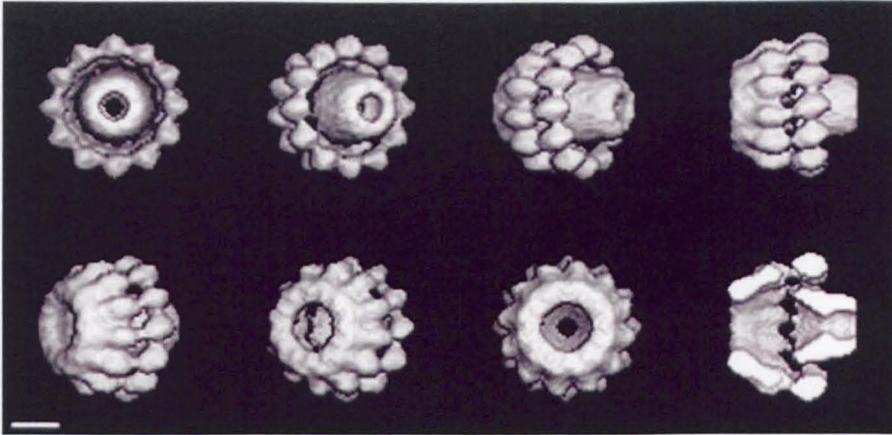


Figure 1.11: Various views of a 3D reconstruction of the HSV-1 portal at 16 Å resolution. The portal comprises twelve copies of UL6. It has an axial channel, peripheral flange-like domains, and its dimensions are compatible with those of the vertex site. Scale bar = 50 Å.

This figure was reproduced from Trus *et al.*, 2004.

*al.*, 2001). UL17 is found in procapsids, B-capsids and C-capsids, with increased levels found in C-capsids, suggesting that it may have a second function after the DNA has been packaged (Goshima *et al.*, 2000; Thurlow *et al.*, 2005). Only UL6, UL17 and UL25 are present in significant amounts in the C-capsid and mature virion. The remaining DNA packaging proteins are only transiently associated with the capsid and appear to disassociate on completion of their function. In null mutants of UL6 and UL17, DNA packaging does not take place. In UL25 null mutants a low level of viral DNA replication is observed, but most of the packaged DNA is not full length. Because of this it has been suggested that the UL25 protein is required at a later stage in DNA packaging than either UL6 or UL17 (NcNab *et al.*, 1998; Hodge, 2001). The UL25 protein has been observed in greater amounts in C-capsids than in B-capsids or the procapsid (Sheaffer *et al.*, 2001; Thurlow *et al.*, 2005), implicating it in a post-packing/scaffold-loss role. It has been suggested as potentially having a role in anchoring the DNA within the capsid following packaging, and in sealing the portal channel (Ogasawara *et al.*, 2001) – a function analogous to the ‘head completion’ proteins of bacteriophages such as P22 (Strauss and King, 1984).

Another protein that plays an important role in DNA cleavage and packaging is UL12, which encodes an alkaline exonuclease. The growth of a UL12 null mutant virus is severely impaired, with yields of progeny virus that are 0.1-1% of wild-type virus yields. Viruses lacking UL12 have been shown to form capsids containing DNA, but these capsids are not transported to the cytoplasm (Shao *et al.*, 1993). The role that has been proposed for the UL12 alkaline nuclease is in processing branched forms of concatameric genomic DNA molecules that arise during HSV-1 replication, presumably by recombination between repeat sequences in the viral genome (Zhang *et al.*, 1994; Martinez *et al.*, 1996a; Goldstein and Weller, 1998). A second protein that is translated from the UL12 ORF is UL12.5. It shares the 3’ terminus with UL12 alkaline nuclease and uses an internal methionine codon of UL12 to initiate translation. UL12.5 is unable to compensate for the absence of full-length UL12, it is not essential for viral growth in culture, it does not associate with viral capsids, and it is unlikely to play an important role in lytic infection (Martinez *et al.*, 1996b; Martinez *et al.*, 2002).

The mechanism by which viral DNA enters the capsid is currently not well defined. Cleavage of concatameric DNA molecules into unit length monomeric genomes is unable to take place in the absence of capsid assembly, indicating that that processes of DNA cleavage and packaging are linked (Desai *et al.*, 1993). A model has been proposed in

which the packaging proteins assemble on the capsid at the unique portal vertex (Newcomb *et al.*, 2001a). The 'a' sequence found in the concatameric genomic DNA contains the signals that are required for cleavage and encapsidation, and the packaging proteins bind to these sequences (Mocarski and Roizman, 1982; Spaete and Frenkel, 1982; Stow *et al.*, 1983; Varmuza and Smiley, 1985). It is believed that the packaging signal at one end of a unit-length genome is taken into the capsid and packaging continues until the packaging signal at the other end of the genome is reached. A cleavage event then takes place, presumably mediated by the UL15/UL28/UL33 terminase, yielding a monomeric DNA molecule. Amplicon vectors are 4-5 kb plasmids, which contain only the HSV-1 origin of replication and minimal DNA cleavage/packaging signals, yet contain all the information necessary to direct packaging of the plasmid into HSV-1 capsids (Spaete and Frenkel, 1982; Hodge and Stow, 2001). These vectors have provided an insight into the mechanisms involved in DNA packaging, as they show that although there appears to be no lower limit in the amount of DNA that will be encapsidated, DNA molecules that are significantly longer than the HSV-1 genome cannot be packaged, presumably because they simply will not fit into the capsid. It appears that UL25 has a role in ensuring that the complete HSV-1 genome is packaged. In a UL25 null mutant virus (KUL25NS), viral DNA was synthesised, but only A and B capsids were present in infected cell nuclei and the replicated KUL25NS DNA was susceptible to DNase (McNab *et al.*, 1998). However, unlike previously examined HSV-1 packaging mutants, both unit-length KUL25NS genomes and terminal genomic DNA fragments were detected in infected cells (McNab *et al.*, 1998). These observations suggest that the absence of UL25 may have resulted in an abortive packaging process in which viral DNA was only transiently associated with the capsid, and that the role of UL25 in packaging was to retain DNA in capsids. Subsequent work by Stow (2001) highlighted the presence of a DNase resistant KUL25NS DNA at low levels (in contrast to the earlier findings), demonstrated the ability of KUL25NS to package genomes of less than approximately 100 kbp, and also showed that the L-terminus of the HSV-1 genome is packaged first, with packaging subsequently proceeding towards the S terminus.

A limitation of the methods of cryo-EM and image reconstruction used with viruses is the requirement that the subject material must possess icosahedral symmetry. While the bulk of the HSV-1 capsid is icosahedrally symmetric, there are features of it that are not, namely the scaffold, the DNA genome, the bulk of the tegument and envelope, UL6, UL15, UL28, UL33 portal/terminase complex. As a result, these structures are not visible in current cryo-EM reconstructions. However, the technique of electron tomography has

elucidated three-dimensional structural information for structures in the HSV-1 virion that do not possess icosahedral symmetry (Grünewald *et al.*, 2003). More recently, single particle cryo-EM was used to determine the structure of epsilon15, a dsDNA bacteriophage that infects *Salmonella anatum*. Using this technique, non-icosahedral components could be seen clustering at one of the twelve capsid vertices (Jiang *et al.*, 2006). The application of these techniques to studies of the portal complex could be used in the future to produce further information on the structures that comprise it.

### **1.3.7. Acquisition of tegument and egress from the host cell**

Once DNA packaging has been completed, the capsid leaves the nucleus and acquires the tegument and envelope prior to exiting the host cell by the exocytic pathway.

The first stage of this process is the exit of the capsid from the nucleus. This is achieved by the budding of the capsid at the inner leaflet of the nuclear membrane (Vlazny *et al.*, 1982). This is known as primary envelopment, and results in the capsid and envelope derived from the inner leaflet being released into the perinuclear space. UL31 and UL34 have both been shown to be involved in this process. These proteins, homologues of which are present in members of all three mammal and avian herpesvirus subfamilies, encode a nuclear phosphoprotein and a C-terminally anchored membrane protein respectively (Purves *et al.*, 1992; Klupp *et al.*, 2000). UL31 and UL34 have been shown to interact (Reynolds *et al.*, 2001; Fuchs *et al.*, 2002a) and form a complex in the inner and outer leaflets of the nuclear membrane. The insoluble lamina underlying the inner nuclear membrane presents a barrier to herpesvirus envelopment. However, it appears that herpesviruses are able to modify the nuclear lamina to allow access to primary envelopment sites at the inner nuclear membrane. In murine cytomegalovirus (MCMV), M50/p38 and M50/p35 (the MCMV homologues of UL31 and UL34 respectively) recruit protein kinase C to phosphorylate and possibly disassemble the nuclear lamina (Muranyi *et al.*, 2002). If UL31 or UL34 are deleted, there is a drastic impairment of primary envelopment, with capsids being effectively trapped in the nucleus (Chang *et al.*, 1997; Klupp *et al.*, 2000; Roller *et al.*, 2000; Reynolds *et al.*, 2001; Fuchs *et al.*, 2002a).

Historically, the fate of DNA-containing capsids following entry into the perinuclear space was a source of controversy. In the luminal model, it was proposed that perinuclear virions contain their full complement of tegument and envelope proteins, which are never lost, and

that the envelope glycoproteins are modified *in situ* in the envelope of the mature virion during transit through the cell. The virions then leave the cell via the cellular secretory pathway (Enquist *et al.*, 1999). Recent research using several approaches has indicated that this model is incorrect. Immunoelectron microscopy has been used to demonstrate differences in composition between perinuclear and mature virions. Although UL31 and UL34 are present in the virions within the perinuclear space, they are not components of cytoplasmic or extracellular virions (Klupp *et al.*, 2000). In other work, a mutant strain of HSV-1 that encoded a form of gD modified to contain an endoplasmic reticulum retention signal was constructed (Whiteley *et al.*, 1999; Skepper *et al.*, 2001). Immuno-EM was used to examine cells that had been infected with either wild-type HSV-1 or the mutant strain of HSV-1 that encoded the ER-retained form of gD. The results showed that in cells infected with wild-type virus, gD was found in both perinuclear and mature extracellular virions. In contrast, when cells were infected with the mutant virus, although the inner and outer nuclear leaflets of the nuclear membrane and perinuclear virions were shown to contain gD, no gD was observed in extracellular virions. These results suggest that the envelope of the perinuclear virions has been lost during virion assembly (Skepper *et al.*, 2001). Further evidence against the luminal model comes from observations indicating that late in infection, tegument proteins accumulate in the cytoplasm of infected cells (Elliott and O'Hare, 1999; Sanchez *et al.*, 2000; Kopp *et al.*, 2002). This would suggest that virion assembly occurs in the cytoplasm rather than the nucleus. Moreover, studies have indicated that the lipid composition of virion membranes is more similar to that of cytoplasmic membranes than nuclear membranes (van Genderen, 1994). However, it should be noted that the host cell undergoes extensive biological changes during infection, and these changes may account for the differences in lipid composition observed.

In a departure from the established model, Leuzinger *et al.*, (2005) recently suggested a third mechanism to account for the fate of DNA-containing capsids. Their observations implied the existence of at least two different pathways of herpesvirus envelopment. One pathway is in accordance with the previously discounted luminal model. In the second pathway capsids escape the nucleus through dilated nuclear pores and are wrapped by Golgi membranes. This hypothesis has been controversial, and the observations reported have since been challenged (Mettenleiter and Minson, 2006).

A number of proteins have been identified as having a potential role in the exit of capsids from the perinuclear space. These are gK (Foster and Kousoulas, 1999), UL20 (Baines *et al.*, 1991), UL48 (Mossman *et al.*, 2000), and US3 (Wagenaar *et al.*, 1995; Klupp *et al.*,

2001). Mutants in these proteins are unable to exit the perinuclear space efficiently, resulting in an accumulation of perinuclear virions. Of these proteins, only UL48 is essential for virus replication. However, this is likely to reflect the importance of the protein in the initiation of viral gene expression, rather than its involvement in egress from the perinuclear space. These observations suggest that egress is not dependent on one protein, although the mechanisms by which these proteins act in egress are not well understood.

Tegumentation of the capsid takes place predominantly in the cytoplasm of the host cell. Sanchez *et al.* (2000) identified large accumulations of tegument proteins that may be sites of virion assembly in the cytoplasm. The tegument contains approximately twenty proteins (Homa and Brown, 1997; Mettenleiter, 2002; Mettenleiter, 2004). In addition to the tegument proteins that are associated with the capsid (Chen *et al.*, 1999; Trus *et al.*, 1999; Zhou *et al.*, 1999), several tegument proteins are known to form interactions with other tegument proteins (for example Read *et al.*, 1993; Elliott *et al.*, 1995; Vittone *et al.*, 2005). Tegument assembly probably begins with an interaction between the UL36 protein (VP1-3) and the capsid. VP1-3 appears to interact with VP5 at the pentons, and to form the innermost layer of tegument. The next stage of tegument formation is the addition of UL37, which has been identified as an interaction partner for UL36 in HCMV and PRV (Klupp *et al.*, 2001). Support for this model comes from the observation that in HSV-1 null mutants of either UL36 or UL37, virion morphogenesis is completely blocked, capsids accumulate in the cytoplasm and no infectious progeny virions arise (Desai, 2000; Desai *et al.*, 2001). Similar observations have been made for null mutants of the PRV homologues of UL36 (Fuchs *et al.*, 2004) and UL37 (Klupp *et al.*, 2001). The more distal parts of the tegument are less well defined. A number of gene products represent tegument constituents, but as yet, many have not been assigned a function in virus maturation.

The final stage in virion assembly is the acquisition of the envelope and the exit of progeny virus from host cells. Tegumented capsids become enveloped by budding into membrane-bound vesicles of the trans-golgi network (TGN) – this is where viral glycoproteins are acquired (Mettenleiter *et al.*, 2004). Viral glycoproteins are translated on ribosomes on the endoplasmic reticulum. From these ribosomes, they enter the cellular secretory pathway and undergo post-translational modification. As a result, the glycoproteins are already present in TGN vesicles prior to secondary envelopment. It has been hypothesised that to mediate envelopment, viral glycoproteins have to interact with tegument components via their cytosolic tails (Brack *et al.*, 2000; Fuchs *et al.*, 2002b; Farnsworth *et al.*, 2003).

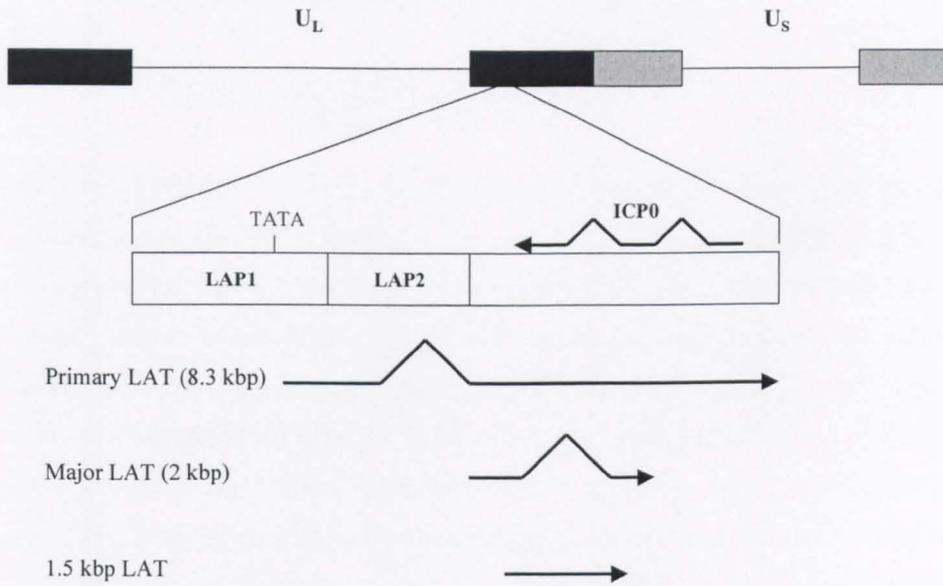
Glycoproteins gD, gE, gH, gI, and gM have all been proposed as candidates for this process (Farnsworth *et al.*, 2003; Gross *et al.*, 2003). While gE, gI and gM are not individually essential, if all three are deleted there is a significant reduction in the yield of progeny virus and an accumulation of unenveloped but tegumented capsids in the cytoplasm – a result which indicates a degree of redundancy in tegument/envelope interactions (Brack *et al.*, 1999). Later work has confirmed that the cytoplasmic tails of the glycoproteins are responsible for mediating this phenotype (Brack *et al.*, 2000). The yeast two-hybrid system was used to detect interactions between the tegument protein UL49 and the tails of glycoproteins gE and gM (Fuchs *et al.*, 2002b). (UL49 is not required for virus formation, which indicates that there must be additional redundancies on the part of tegument proteins.) In more recent work, Chi *et al.* (2005) used a glutathione S-transferase fusion protein binding assay was used to show that UL49 bound specifically to the gD tail. They also showed that HSV-1 capsids bound to the gD tail, suggesting that the UL49 protein may act as a linker, mediating the interaction of the capsid with gD. A similar interaction has been proposed for the tegument protein UL11 (Kopp *et al.*, 2003). UL11 has been proposed as being involved in directing herpesvirus tegument proteins to the envelopment site. Interestingly, in HSV-1 and PRV, UL11 does not appear to be essential for the production of progeny virus, since deletion mutants only moderately impair viral replication (MacLean *et al.*, 1992; Kopp *et al.*, 2003). Nevertheless, a similar deletion mutant in HCMV blocked replication (Silva *et al.*, 2003). A further glycoprotein/tegument interaction has been observed to take place between the cytoplasmic tail of gH and VP16 (Gross *et al.*, 2003). When VP16 is deleted, cytoplasmic capsids fail to envelope (Mossman *et al.*, 2000), suggesting that VP16 might be involved in sequestering membrane proteins at the site of envelope formation. However, its interactions with glycoproteins may exhibit a degree of redundancy as it is also able to interact with gB and gD (Zhu and Courtney, 1994). Overall, the results suggest a model for secondary envelopment that involves the recruitment of tegumented capsids to membranes of TGN vesicles by interactions between glycoproteins gE, gI, gD, gH and gM and tegument proteins UL11, UL48, and UL49.

### **1.3.8. Latency**

One of the defining characteristics of herpesviruses is their ability to establish potentially life-long latent infection in the host organism. Latent infection in herpesviruses can be viewed as having three distinct phases: establishment, maintenance, and reactivation.

During initial infection, HSV-1 virions infect nerve endings of the sensory nervous system that are located in the mucosal epithelium. They are then transported to the neuronal cell body, where the DNA enters the nucleus and is able to persist in an episomal state for the lifetime of the host. Periodic reactivation of HSV-1 can occur, leading to lytic infections in the same dermatomal distribution as the initial infection (reviewed by Jones, 2003). During latency the only abundant viral RNAs are the latency associated transcripts (LATs) (Stevens *et al.*, 1987), although very low levels of lytic gene RNAs have been detected by reverse transcription PCR techniques (Kramer and Coen, 1995; Chen Hua *et al.*, 2002). The location of the LATs in the HSV-1 genome is shown in Figure 1.12. The LAT region encodes multiple transcripts, including the 8.3 kb primary transcript and two stable introns of 2.0 kb and 1.5 kb (Farrell *et al.*, 1991), which are predominantly localised to the neuronal nucleus and are transcribed antisense and partially complementary to the coding sequence for ICP0.

Since the LATs are the only gene products transcribed during latency, it would seem reasonable to assume that they must have some role in the establishment, maintenance or reactivation from latency. However, their true role has been the subject of conflicting reports and currently remains unidentified. The observations that LATs are transcribed antisense and complementary to ICP0, and that ICP0 is a powerful transactivator of IE genes has led to the suggestion that the LATs themselves function to repress IE gene expression by preventing or reducing the expression of ICP0. This was supported by evidence that the 2 kb LAT could inhibit the transactivation of the HSV-1 tk gene promoter by ICP0 (Farrell *et al.*, 1991). In addition, mutant viruses unable to make the 2 kb LAT have an increased accumulation of ICP0 RNA (Arthur *et al.*, 1998). Furthermore, LATs can suppress IE gene expression in a neuronal cell line (Mador *et al.*, 1998). This theory would therefore predict that a virus lacking both copies of LAT would be unable to repress ICP0 expression and therefore reactivate from latency with enhanced efficiency. Interestingly, the opposite appears to be true, with observations demonstrating that LAT negative mutants are impaired for reactivation *in vivo* (Leib *et al.*, 1989; Hill *et al.*, 1990; Perng *et al.*, 1994; Bloom *et al.*, 1996). However, whether this impaired reactivation is simply a downstream effect of a reduced efficiency in latency establishment is debatable (Thomson and Sawtell, 1997).



**Figure 1.12:** The genetic organisation of the LAT region.

The LAT region is located on the long repeat regions of the genome, and is therefore present in two copies. The primary LAT is an unstable 8.3 kbp transcript. The major species are the 2 kbp and 1.5 kbp LATs, which may be smaller spliced introns from the primary LAT. LATs are transcribed antisense and at least partially complementary to the gene encoding ICP0. The two latency-associated promoters, LAP1 and LAP2, are indicated.

## 1.4. Capsid structure

### 1.4.1. Principles of capsid structure

#### 1.4.1.1. Quasi-equivalence

Crick and Watson (1956) first realised that viral genomes did not possess enough genetic information to encode a single protein coat molecule that could surround the viral genome. They suggested that virion coats would have to be made up of symmetric arrays of one or a small number of proteins, and that these symmetric arrays would have specific bonding properties, allowing them to make specific intersubunit contacts that would be repeated exactly throughout the particle. In helical viruses such as tobacco mosaic virus (TMV), the helical coat protein fulfils this prediction – each subunit in the TMV capsid particle interacts with its neighbouring protein in essentially the same way. The exceptions are those subunits at the ends of the virus particle, for which there are unsatisfied bonding capabilities (Crick and Watson, 1956; Namba *et al.*, 1989; reviewed by Casjens, 1997).

The situation is more complex in so-called spherical viruses with closed protein shells. Although in theory capsid shells could be tetrahedral, cubic, octahedral, dodecahedral, or icosahedral (the five kinds of cubic symmetry), only examples of icosahedral symmetry have been found. Crick and Watson (1956) proposed that in order to construct a virus particle using repeating units of a single polypeptide, in which each subunit interacts identically with its neighbours, the virus particle would have to be made up of exactly 60 subunits. This hypothesis holds true for some very small icosahedral virus particles such as the parvoviruses. Parvoviruses typically consist of a linear 5 kb single-stranded DNA genome surrounded by an icosahedral capsid shell made up of 60 copies of a single polypeptide unit. However, the hypothesis does not account for the many icosahedral viruses that are made up of considerably more than 60 identical subunits, such as HSV-1 (Casjens, 1997).

The contradiction between the large size of many icosahedral viruses and the requirements for identical interaction led Caspar and Klug (1962) to promote a theory that appears to reconcile these differences. Caspar and Klug suggested that all icosahedral viruses have rings of five subunits that interact around each of the 12 fivefold rotationally symmetric icosahedral vertices ( $5 \times 12 = 60$ ). For viruses in which more than 60 subunits are used, the extra subunits are systematically inserted between the pentamers as rings of six subunits

(hexamers). A combination of pentamers and hexamers is required to form a closed icosahedral shell. A repeating pattern of hexamers can form a flat plane, but this is not the case with pentamers, since removing one subunit from a hexamer causes it to adopt a convex shape. Hence, in capsids, pentamers form the corners while hexamers form the faces. The addition of extra hexamer subunits would allow the size of the icosahedron to be increased without disrupting the basic geometry of the capsid shell. Thus, the simplest virus capsids consist of an icosahedron composed solely of 12 pentamers at the capsid vertices (60 subunits). The smallest capsid shells composed of both hexamers and pentamers consist of 12 pentamers at the capsid vertices and 20 hexamers on the faces of the capsid (180 subunits). An icosahedral capsid constructed using 60 subunits is referred to as a  $T=1$  icosahedron. Likewise, a capsid constructed using 180 subunits is a  $T=3$  icosahedron. The multiples of 60 subunits could be any member of the infinite series  $T=1, 3, 4, 7, \dots$ , ( $T = h^2 + hk + k^2$ , where  $h$  is the number of moves in one direction and  $k$  is the number of moves in a second direction from one pentamer to the next on the surface lattice) corresponding to 60, 180, 240, 420, ..., subunits respectively (Caspar and Klug, 1962; Johnson and Speir, 1997). A consequence of this theory is that subunit proteins interact with their neighbours in slightly different ways, depending on their location within the capsid. Therefore, icosahedral shells that contain hexamers have 'quasi-equivalent' geometry: the morphology of the proteins and protein interactions formed in their construction is almost the same, but not identical, to what it would be in a capsid composed of just 60 subunits (reviewed by Casjens, 1997).

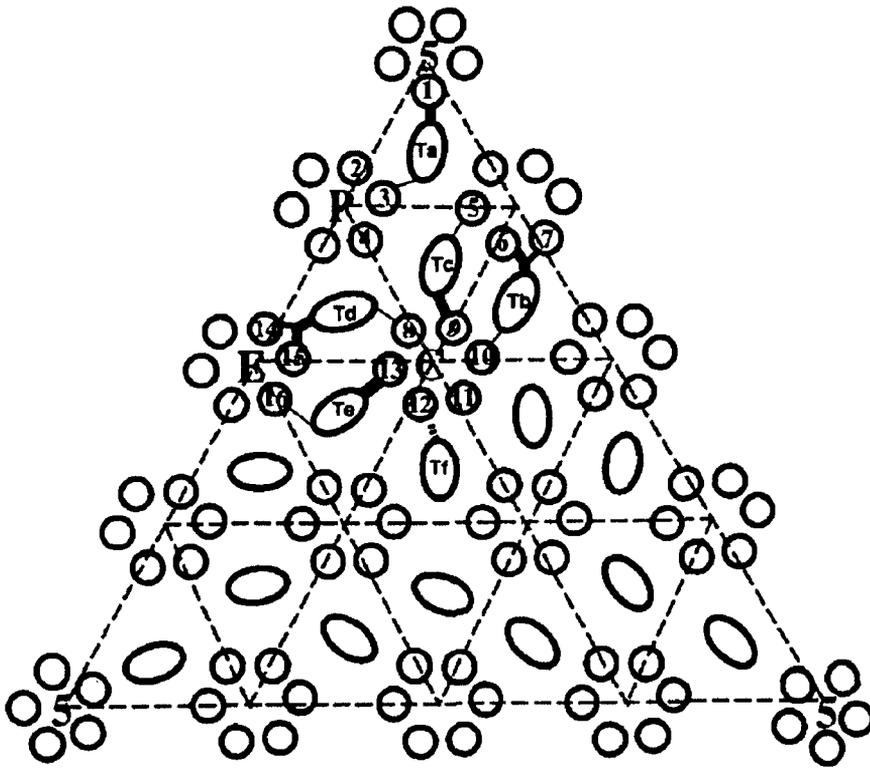
The models proposed by Crick and Watson (1956) and Caspar and Klug (1962) are purely mathematical attempts to describe means by which proteins can interact to form large icosahedral viruses. However, experimental analysis of virus particles has revealed that some viruses do not conform to these models, and that in some cases they diverge greatly. The polyomavirus capsid is composed of 360 subunits of the major capsid protein. According to the models described above, this would suggest that the capsid is made up of 62 capsomers (12 pentamers and 50 hexamers) arranged on a  $T=6$  lattice – an impossible configuration. In reality, the polyomavirus capsid is made up of 72 capsomers, which are all pentamers of the major capsid protein. The 72 capsomers conform to  $T=7$  icosahedral symmetry, with each hexavalent capsid being constructed from five copies of the major capsid protein (reviewed by Dokland, 2000). Another example of divergence from the models mentioned above can be seen in adenovirus capsids. The adenovirus capsid has 252 capsomers that are arranged with  $T=25$  symmetry (reviewed by Johnson and Speir, 1997). Quasi-equivalence would imply the presence of 1500 structural subunits (12 pentamers and

240 hexamers). However, in adenovirus capsids the hexons and pentons are constructed from two different types of protein. The twelve vertices are pentamers of the penton base protein, while the 240 'hexamers' are trimers of the major capsid protein. Thus, adenoviruses deviate from quasi-equivalence in two major ways.

#### 1.4.1.2. Quasi-equivalence in the HSV-1 capsid

Unusually for such a large virus, the HSV-1 capsid conforms closely to Casper and Klug's quasi-equivalence model. Until recently, it was thought that the capsid was constructed using 960 copies of VP5 arranged as 12 pentamers (pentons) and 60 hexamers (hexons), with the pentons and hexons being made up of five or six copies of VP5 respectively. It is now generally accepted that one vertex is occupied by the portal complex, although it is not yet certain whether this replaces or is present in addition to a penton. Apart from this, the structural arrangement of the HSV-1 capsid corresponds closely to  $T=16$  icosahedral symmetry, although the differences in structure between penton and hexon subunits and slight differences among hexon subunits (He *et al.*, 2001) are not consistent with precise quasi-equivalence.

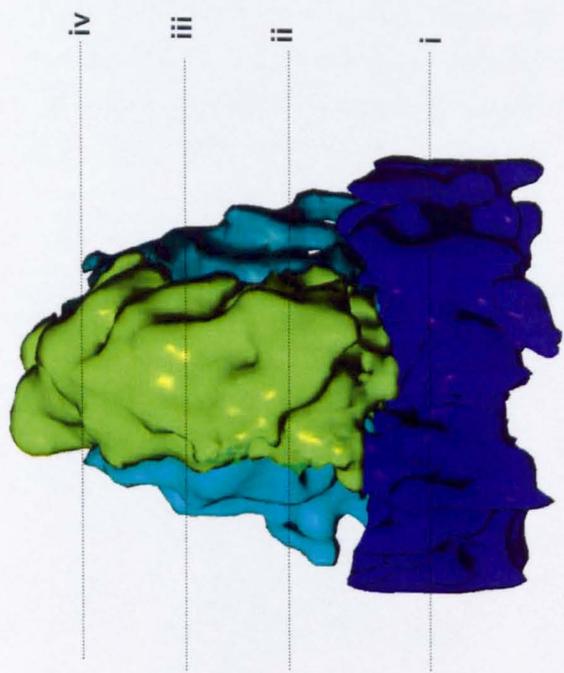
In contrast to the capsomers, the triplexes diverge significantly from quasi-equivalence. Triplexes occur in positions of three-fold symmetry between adjacent capsomers and form interactions with the middle and floor domains of the surrounding capsomers. Cryo-EM has been used to demonstrate that the interactions between adjacent capsomers are not uniform, with  $T_a$ - $T_f$  (Figure 1.13) having slightly different degrees of interaction with adjacent capsomers. The most striking differences are observed when one compares  $T_a$  to triplexes  $T_b$ - $T_c$  (Zhou *et al.*, 1994; Zhou *et al.*, 1998). This variation can also be detected biochemically, possibly reflecting differences in bonding interfaces. Peripentonal triplexes ( $T_a$  and  $T_c$ ) more readily disassociate from capsids treated with GuHCl than triplexes surrounding adjacent hexons (Newcomb *et al.*, 1993). The triplex is a heterotrimer composed of two copies of VP23 and one copy of VP19C. It is therefore, by definition, an asymmetric structure. As a result, it has been difficult to determine how the triplex occupies a three-fold symmetrical position within the icosahedral capsid. The floor of the capsid, which interacts with each triplex, exhibits precise three-fold radial symmetry. This paradox may be resolvable through the observation that the base of the triplex appears to form a pseudotrimer with VP19C and the two VP23s forming symmetrical connections with the three identical VP5 floor domains (Zhou *et al.*, 2000). Above the floor, the arrangement of the upper regions of the triplex proteins is markedly asymmetric (Figure 1.14).



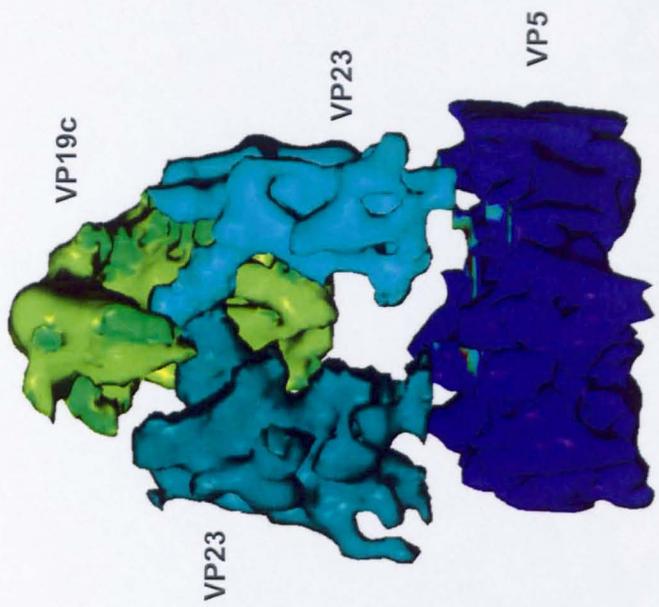
**Figure 1.13:** Schematic representation of the relative locations of the various structural components of the HSV-1 capsid. The quasi-equivalent subunits of pentons and hexons in one of the asymmetric units are labelled 1-16. The quasi-equivalent triplexes in the asymmetric unit are labelled Ta-Tf. Lines indicate the interactions between triplexes and their neighbouring capsomeric subunits.

This figure was reproduced from Zhou *et al.*, (1994).

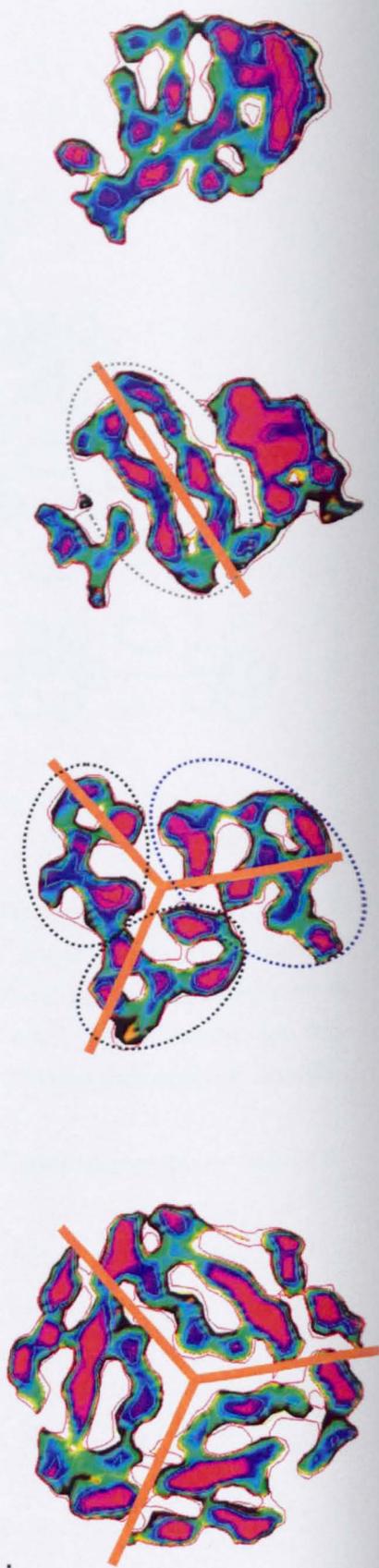
A.



B.



C.



**Figure 1.14:** Structure of the triplex at 8.5 Å resolution.

**A.** A single triplex subunit at 8.5 Å resolution. VP19C is shown in pale green. The two VP23s are shown in turquoise and bottle-green. The capsid floor (blue) is formed by domains of VP5. The structure shown was obtained by averaging the structural data from the four triplexes,  $T_b$ ,  $T_c$ ,  $T_d$ , and  $T_e$ , as labelled in Figure 1.15. **B.** The same triplex subunit rotated by 180° around the 3-fold axis of symmetry. The molecules have been separated to highlight the structural arrangement of the component proteins. **C.** Images i-iv represent horizontal cross-sections through the triplex structure at the positions indicated by broken lines in figure A. (i) Cross-section through the capsid floor. (ii) Cross-section through the base of the triplex where it makes contact with the capsid floor. The densities attributed to VP19C (blue ellipse) and VP23s (black ellipse) are indicated. The arrangement of the three protein densities approximates the three-fold radial symmetry of the capsid floor (image (i), orange lines). (iii) The two copies of VP23 form a flattened ring (grey ellipse) that now exhibits two-fold radial symmetry (orange line), with densities corresponding to VP19C on either side. (iv) The top of the triplex: a completely asymmetric structure composed almost entirely of VP19C.

This figure was reproduced from Zhou *et al*, (2000).

### 1.4.2. Overview of capsid structure

The HSV-1 capsid has been the subject of intense study, and its structure and composition are relatively well characterised compared to other virion compartments (Steven and Spear, 1997). It consists of a protein shell, which has  $T=16$  icosahedral symmetry. Three distinctive structural elements, designated pentons, hexons, and triplexes, make up the bulk of this shell. 150 hexons form the faces and edges of the icosahedron, while the vertices are formed by the pentons. Hexons and pentons are made up of five and six copies of the major capsid protein, VP5 (Section 1.2.2.2), respectively. Six copies of VP26 (Section 1.2.2.2) occupy the outer surface of each hexon. Three separate classes of hexons are formed: P (peripentonal), E (edge), and C (central), each of which has a slightly different morphology as a result of differences in quasi-equivalent environments in the icosahedral capsid face (Steven *et al.*, 1986; Schrag *et al.*, 1989; Steven and Spear, 1997; He *et al.*, 2001). The locations of the three forms of hexons are shown in Figure 1.15, and discussed further in Section 1.4.3. The triplexes, characteristic features of herpesvirus capsids, are heterotrimers made up of VP19C and VP23 (Section 1.2.2.2) in  $\alpha\beta_2$  organisation, which in HSV-1 is formed by a single molecule of VP19C ( $\alpha$ -subunit) and two copies of VP23 ( $\beta$ -subunit). They occupy the 3-fold positions between the hexons and pentons. The icosahedral geometry of the capsid largely determines the numbers of each component, but the precise numbers of pentons and triplexes is uncertain. This is because one of the twelve vertices is believed to be the location of a portal complex through which the genome enters and exits the capsid (Newcomb *et al.*, 2001a). It seems probable that the portal replaces one of the pentons leaving eleven occupying the remaining icosahedral vertices. Similarly, for the triplex, there are 320 local 3-fold positions in the capsid, including five surrounding each vertex. However, the relationship between the triplexes and the portal is unknown, making it unclear whether triplexes are present at the five positions surrounding the portal.

Early work to identify the structure of the capsid involved the EM analysis of negatively-stained specimens (Wildy *et al.*, 1960; Furlong, 1978). These studies demonstrated that the capsid is icosahedral in shape and is made up of 162 capsomers, thus producing a structure that has icosahedral symmetry of triangulation class  $T=16$ . More recent studies have employed the techniques of cryo-EM and image reconstruction to generate three-dimensional structures of the HSV-1 capsid. This approach was first described by Schrag *et al.* (1989), who determined the structure of the capsid to a resolution of 40Å. Advances

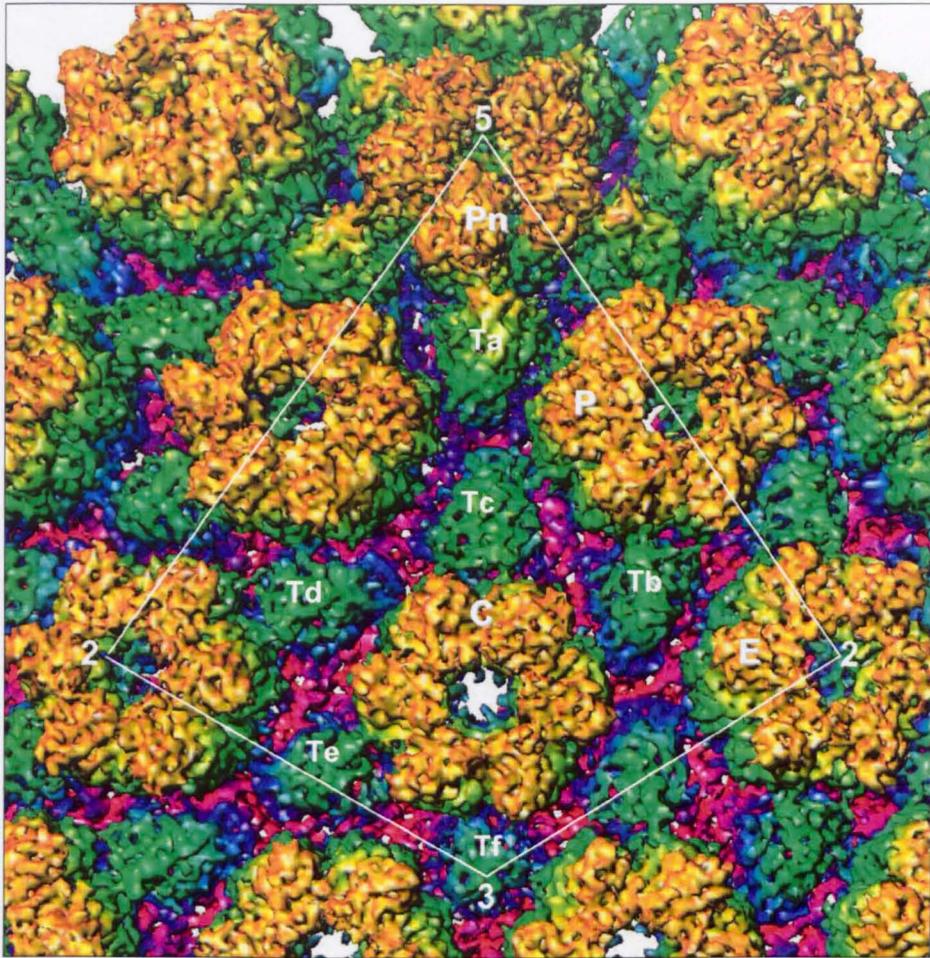


Figure 1.15: The asymmetric unit of the HSV-1 capsid.

An enlarged view of the HSV-1 capsid at 8.5 Å resolution. The structural components of the asymmetric unit are indicated: the penton (labelled Pn), the three types of hexon (P, E and C), and the six types of triplex (Ta to Tf). The image is radially colour-coded. Inner structures (the capsid floor) are blue and outer structures (capsomeres and triplexes) are green and yellow. The icosahedral five-fold, three-fold and two-fold axes of symmetry, which define the asymmetric unit, are indicated as 5, 3 and 2 respectively.

This figure was reproduced from Zhou *et al.*, (2000).

in the technology used to produce these images have resulted in a gradual refinement of the generated structures, and to date the HSV-1 capsid has been determined to a resolution of 8.5Å (Zhou *et al.*, 2000).

The capsid shell is approximately 150Å thick. Capsomers are held together by their innermost portions. These are packed closely together to form the floor of the capsid, which is approximately 35Å thick. Each capsomer has a chimney-like protrusion that rises approximately 100Å outward from the floor, and is traversed by an axial channel of non-uniform diameter. On the outer surface of the floor lie the triplexes. (Steven and Spear, 1997).

At the current resolution of 8.5Å (Figure 1.7) (Zhou *et al.*, 2000),  $\alpha$ -helices are the only distinguishable secondary structural feature (Baker *et al.*, 2003). However, our understanding of the structure of part of the HSV-1 capsid has been greatly increased following the publication of the crystal structure of a portion of the HSV-1 major capsid protein (Figure 1.16) (Bowman *et al.*, 2003). HSV-1 capsid proteins have generally been intractable to analysis by x-ray crystallographic techniques, but partial trypsin digestion isolated a 65 kDa fragment of VP5 that was soluble and crystallised. The structure of this fragment was solved to a resolution of 2.9Å. The fragment represented a central 604 amino acids of the protein sequence (residues 451-1054), and using a combination of x-ray crystallographic, EM and biochemical information it was identified as the upper domain of VP5 (VP5ud). Fitting the VP5ud into the 8.5Å cryo-EM reconstruction of the capsid (Zhou *et al.*, 2000) revealed details of the interactions between regions of VP5 within hexons and pentons. These results will be discussed fully in Section 1.4.4.1.

### **1.4.3. The asymmetric unit**

An asymmetric unit is the smallest portion of a structure to which symmetry can be applied to generate one full unit. In the context of the HSV-1 capsid, this is all the information required to build an icosahedral capsid shell. Each of the twenty triangular faces of the icosahedron consists of three asymmetric units, each of which is made up of one fifth of a penton, one P-hexon, one C-hexon, one half of an E-hexon, one each of triplexes  $T_a - T_e$ , and one third of triplex  $T_f$  (Figure 1.15; Zhou *et al.*, 1994). Thus, in the structural analysis of HSV-1 capsids it is standard practice to simplify the analysis by assuming that the

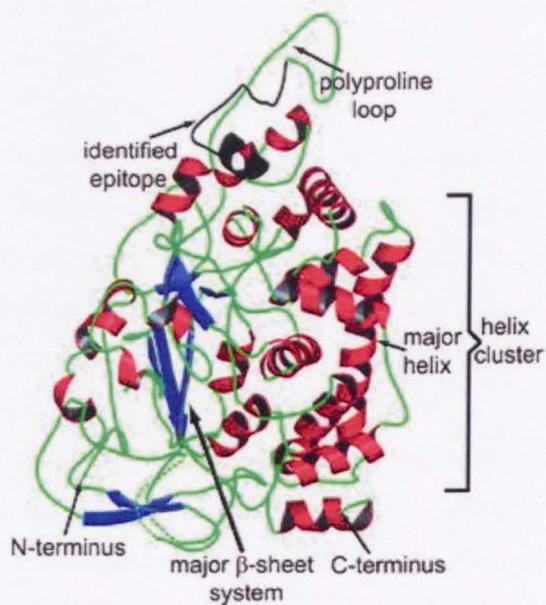


Figure 1.16: A ribbon representation of the crystal structure of the upper domain of VP5 (residues 451-1054).

The secondary structural elements in the structure are coloured:  $\alpha$ -helices are shown in red,  $\beta$ -strands are shown in blue, and loops are shown in green. Key elements are labelled in the structure.

Reproduced from Bowman *et al.*, 2003.

capsid is formed from 60 identical asymmetric units. The classifications of different types of hexon and triplex are described in more detail in the following sections.

#### **1.4.4. Structural composition of the capsid**

The seven proteins that make up the HSV-1 capsid have been known for some time: VP5, VP19C, VP21, VP22a, VP23, VP24, and VP26 (Section 1.2.2.2). Historically, two main approaches have been used to determine the structural organisation of these proteins, namely antibody labelling (Vernon *et al.*, 1981; Rixon *et al.*, 1988; Trus *et al.*, 1992) and the systematic depletion of component proteins from purified capsids (Newcomb and Brown, 1989; Newcomb and Brown, 1991; Newcomb *et al.*, 1993). Newcomb and Brown (1989) used argon ion (Ar) plasma etching to show that when purified B-capsids are exposed to low doses of ion plasma, VP5, VP19C, and VP23 more readily disassociate from the capsid than VP21, VP24 and, VP22a. They concluded that VP5, VP19C and VP23 are more likely to be present on the outer surface of the capsid, and that VP21, VP24 and VP22a are more likely to be internal components, requiring longer periods of exposure to disassociate. More recently, techniques such as cryo-EM and image reconstruction have made more detailed analysis of the organisation of the component proteins possible, culminating in the resolution of the 8.5Å structure (Zhou *et al.*, 2000) – currently the best available. A complete structural description requires atomic resolution, however at present this is only possible through X-ray crystallography.

##### **1.4.4.1. Hexons and pentons are respectively hexamers and pentamers of VP5**

It was Gibson and Roizman (1972) who originally proposed that VP5 is a major component of the HSV-1 capsid shell. This proposal was based on SDS-PAGE studies of capsids, which showed that VP5 comprised approximately 70% of capsid mass. Subsequently, Trus *et al.*, (1992) used cryo-EM to show that antibodies raised against VP5 bound only to the hexons, while an antibody raised against purified B-capsids bound only to the pentons. They concluded that the differences in antibody binding were likely to have been caused by differences in conformation for VP5 in hexons and pentons. Pentons in HSV-1 B-capsids treated with either guanidine hydrochloride (GuHCl) or 6M urea were shown to dissociate from the capsid. SDS-PAGE analysis of the material that had been removed confirmed that the pentons were composed of VP5, and that hexons and pentons were hexamers and pentamers of VP5 (Newcomb and Brown, 1993).

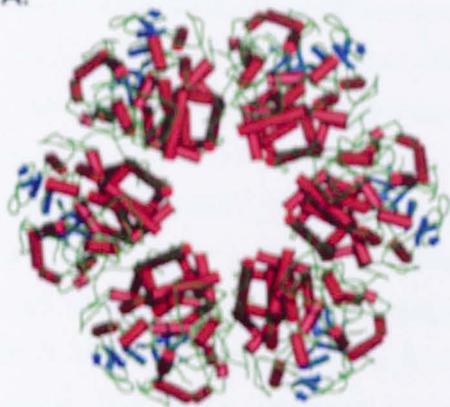
More recently, detailed information on the structure of hexons and pentons has been obtained from cryo-EM and X-ray crystallography experiments (such as Zhou *et al.*, 2000; Baker *et al.*, 2003; Bowman *et al.*, 2003). Hexons and pentons are both cylindrical and of similar size and dimensions. Hexons are 170Å in diameter and 140Å high, while pentons are 145Å in diameter and have a height of 140Å (Steven *et al.*, 1986; Schrag *et al.*, 1989; Zhou *et al.*, 1994). Analysis of the crystal structure of the VP5ud shows that the protein folds in a broadly similar manner in hexons and pentons (Bowman *et al.*, 2003). However, it is interesting to note that previous studies identified antibodies that bind to distinct conformational epitopes at the tops of either hexons or pentons (Trus *et al.*, 1992). This observation implies that each type of capsomer also contains distinctive features, which could explain the specific interactions formed by hexons and pentons (Wingfield *et al.*, 1997).

The five copies of VP5 found in the penton are less closely associated than those in the hexon, and are 'twisted' relative to each other (Figure 1.17), resulting in the penton channel being larger than that found in the hexon. These differences in arrangement have been proposed (Bowman *et al.*, 2003; Chen *et al.*, 2001) to account for VP26 only binding to the hexon (Trus *et al.*, 1995; Zhou *et al.*, 1995) and the tegument only binding to the penton (Zhou *et al.*, 1999) (Figure 1.17). Differential binding in hexons and pentons is not restricted to the tops of VP5. Scaffolding protein has been shown to interact with the floor domains of hexons, but not pentons (Zhou *et al.*, 1998b).

#### **1.4.4.2. VP26 is located at the tips of the hexons but not the pentons**

Newcomb and Brown (1991) originally proposed that VP26 was either associated with capsid vertices or formed the triplexes. They demonstrated that VP26 could be readily detached from purified capsids in the presence of either urea or GuHCl. Subsequent cryo-EM analysis of VP26 depleted and VP26 re-associated capsids revealed that VP26 is specifically located at the tips of hexons (Booy *et al.*, 1994). Studies using recombinant capsids have shown that six copies of VP26 bind to each hexon, forming a star-shaped ring at the hexon tip (Zhou *et al.*, 1995; Wingfield *et al.*, 1997). VP26 is absent from the equivalent position in pentons (Zhou *et al.*, 1995). It has been proposed that this is due to different interactions between neighbouring VP5s in hexons and pentons, as discussed in Section 1.4.4.1. This theory has been supported by the observation that VP26 does not bind

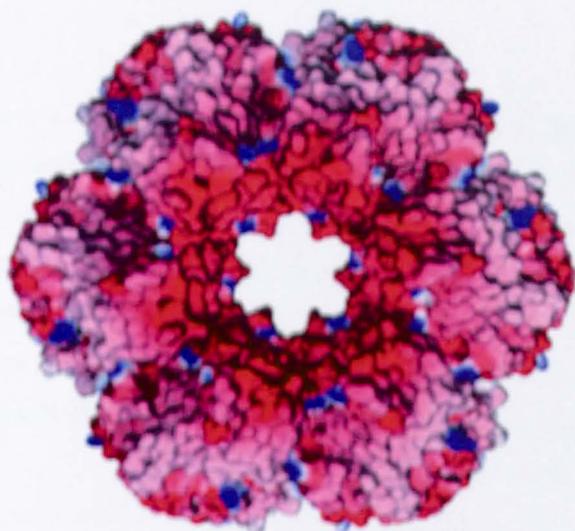
A.



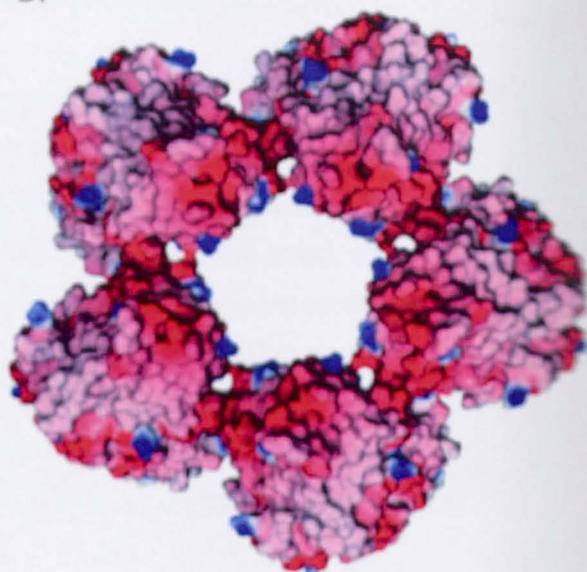
B.



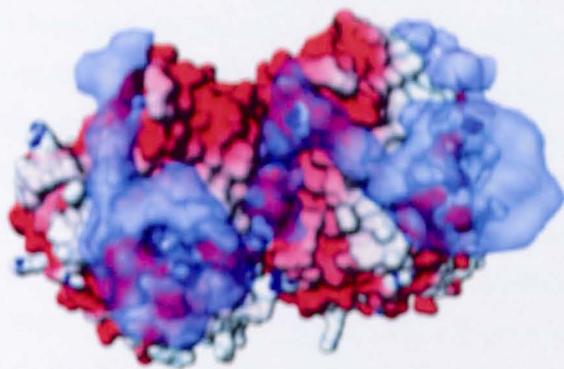
C.



D.



E.



F.

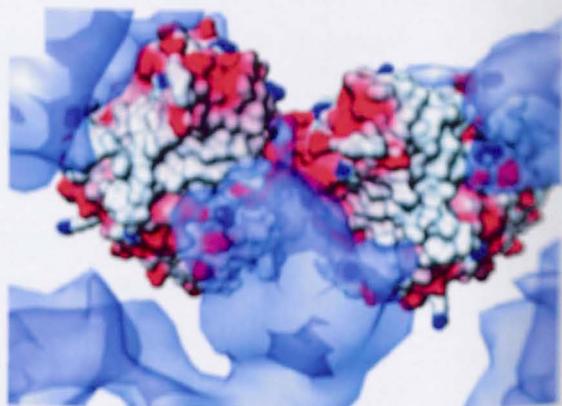


Figure 1.17: Structural comparisons of the hexon and penton (viewed from above).

**A.** Ribbon representation of the VP5ud 2.9 Å crystal structures as seen from above in the hexon (C-hexon). Alpha helices are shown as red cylinders, beta sheets as blue arrows, and looped regions are shown as green lines. **B.** Ribbon representation of the VP5ud structures in the penton as seen from above (same colour scheme as A). VP5 subunits in the penton are twisted in comparison to their organisation in the hexon. **C.** Space-filled representation of the VP5ud structures in the hexon (C-hexon) as seen from above. **D.** Space-filled representation of the VP5ud structures in the penton as seen from above. The penton channel has a wider diameter than that of the hexon. The structures in images C and D are coloured coded to reflect the electrostatic charges of the amino acids in each structure. Acidic residues are coloured red, basic residues are coloured blue and uncharged amino acids are coloured pink/grey. As a consequence of the differences in structural arrangements, the hexon channel is much more acidic than the penton channel. **E.** Space-filled representation of two hexon VP5ud subunits bound to the VP26 small capsid protein (blue-grey semi-transparent structure). **F.** Space-filled representation of two penton VP5ud subunits bound to the VP1-3 tegument protein (blue-grey semi-transparent structure).

Figure reproduced from Bowman *et al*, 2003

to the procapsid, where the conformation of VP5s in hexons greatly differs from that in mature capsids (Chi and Wilson, 2000; Newcomb *et al.*, 2000).

#### 1.4.4.3. Triplexes are heterotrimers of VP19C and VP23

Triplexes (Figure 1.14) were originally identified as a component of the HSV-1 capsid by negative stain EM analysis. They were shown to connect adjacent capsomers (Vernon *et al.*, 1974). Early cryo-EM work showed triplexes as Y-shaped masses located on the 3-fold axes of symmetry and connecting adjacent capsomers (Schrag *et al.*, 1989; Baker *et al.*, 1990; Booy *et al.*, 1991). 2M GuHCl treatment of capsids removed the triplexes that surround the penton ( $T_a$  and  $T_c$ ) (Newcomb *et al.*, 1993). Quantitative analysis of the proteins associated with the removal of these triplexes suggested that the triplexes were heterotrimers made up of one copy of VP19C and two copies of VP23. More recently, high-resolution cryo-EM has been used to show that contrary to earlier hypotheses, most triplexes are asymmetrical rather than Y-shaped (Zhou *et al.*, 1994; Zhou *et al.*, 1998b). For descriptive purposes, the triplex is separated into an upper, diamond-shaped portion comprising the 'head' and 'tail' regions, and a lower portion consisting of two 'legs'. The head and tail form connections with adjoining VP5 molecules, and the legs interact with the capsid floor. Studies of recombinant particles that lack VP23 have allowed the positions of each component to be determined (Saad *et al.*, 1999). These studies proposed that the two copies of VP23 form the legs, while VP19C forms the head and the tail of the triplex (Zhou *et al.*, 2000), although more recent work has cast doubt on this interpretation (Section 5.4).

A total of six different types of triplex ( $T_a$ - $T_f$ ) contribute to a single asymmetric unit, with a total of 16 triplexes making up each icosahedral face (Figure 1.13) (Zhou *et al.*, 1994). Each type of triplex has a slightly different morphology by virtue of differences in the quasi-equivalent environments in the icosahedral capsid face.  $T_a$ - $T_e$  are asymmetric, but  $T_f$ , which is located at the global three-fold axis, appears to be symmetrical in computer reconstructions, although this is believed to be an artefact of the reconstruction process. The locations and interactions between the triplexes and their adjacent capsomers varies according to the position of the triplex within the capsid.

#### 1.4.4.4. Structure of the internal scaffold

It was originally proposed that the major scaffolding protein, VP22a, was only present in B-capsids (Gibson and Roizman, 1972). This observation was later supported by Rixon *et*

*al.* (1988) who used immunogold labelling to demonstrate that VP22a is a component of empty capsids but not present in full capsids or mature virions, and suggested that VP22a is transiently associated with capsids at an early stage of assembly and is lost during DNA packaging. Ar<sup>+</sup> etching of purified B-capsids revealed that VP22a, VP21, and VP24 were all relatively well protected from erosion by low-intensity plasma, and it was concluded that they must form the internal components of the capsid (Newcomb and Brown, 1989). Subsequently, cryo-EM was used to compare A and B capsids. Additional protein masses were observed in B-capsids and it was suggested that these correlated to the presence of internal core proteins (Baker *et al.*, 1990). It is now known that the major internal scaffold component is VP22a, which is present in over 1000 copies in each capsid (VP21 and VP24 are both present in approximately 100 copies per capsid) (Newcomb *et al.*, 1993). The scaffold exists in two forms within the interior of capsids. In large cored capsids, the C-terminal 25 amino acids of preVP22a and/or preVP21 have not been cleaved by the VP24 protease, while in small cored capsids this cleavage event has taken place. Large cored capsids can be made by omitting UL26, which encodes VP21 and VP24 (Tatman *et al.*, 1994; Thomsen *et al.*, 1994). Within the UL26 gene, temperature sensitive mutants that are defective in the processing of UL26 into VP21 and VP24, such as *ts1201*, also produced large cored capsids at the non-permissive temperature. If the temperature was altered to become permissive, UL26 is processed correctly and small cored capsids are produced (Preston *et al.*, 1983). Gao *et al.* (1994) demonstrated that virus lacking UL26 was unable to grow on Vero cells, and that although the mutant virus produced capsids, the assembly of VP5 into hexons is altered, suggesting that the protease is required for the formation of functional capsid structures.

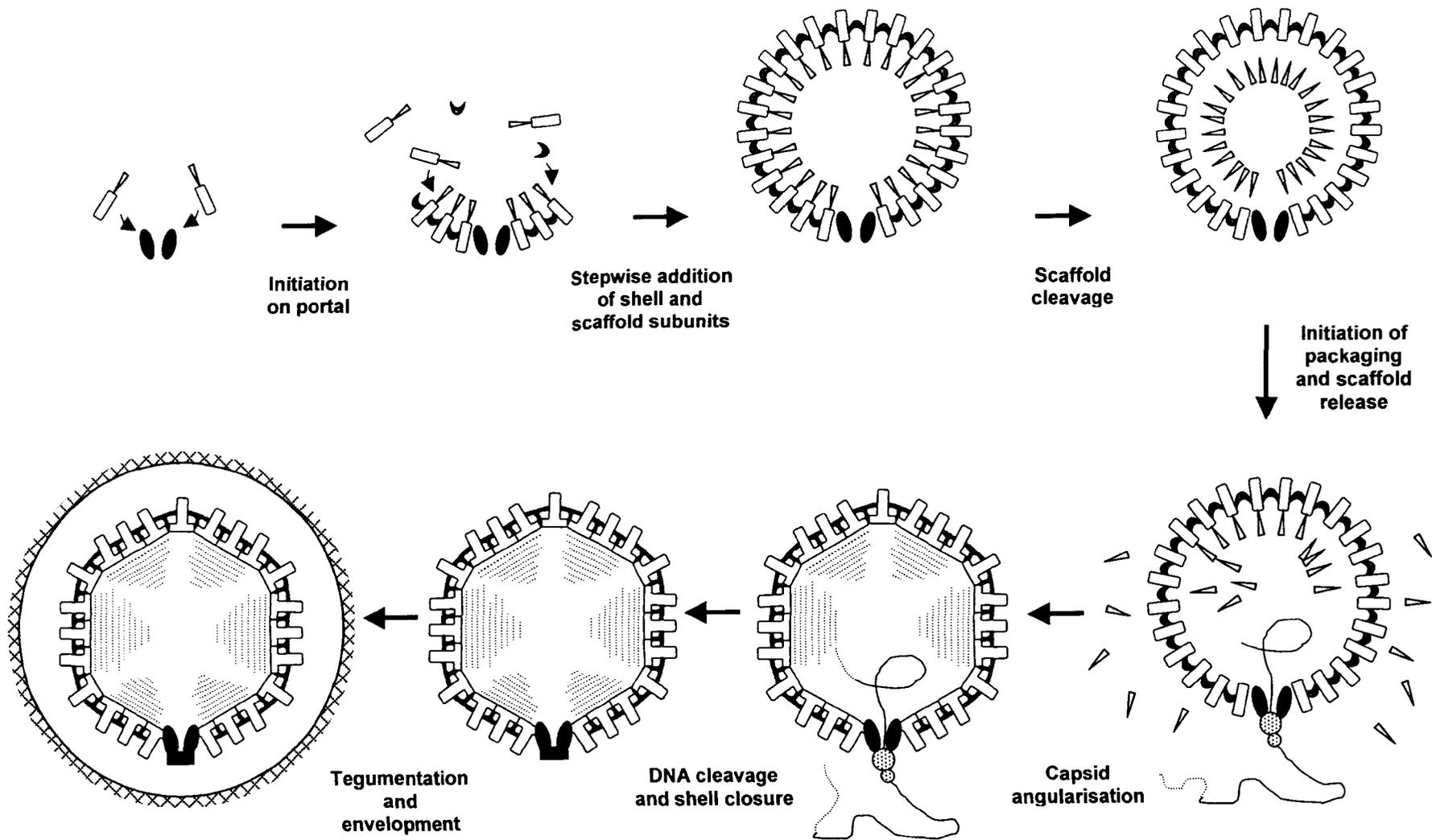
The interaction between preVP22a and VP5 has been shown to occur through the C-terminal 25 amino acids of VP22a, indicating that these residues are essential for capsid formation (Kennard *et al.*, 1995; Matusick-Kumar *et al.*, 1995; Hong *et al.*, 1996). These residues are thought to interact with the N-terminal of VP5 (Desai and Person, 1999; Warner *et al.*, 2000; Warner *et al.*, 2001). Cryo-EM analysis has revealed that most of the scaffold mass possesses no icosahedral symmetry, although some regions do appear to be icosahedrally arranged, forming rod-like densities that protrude from the capsid shell into the interior of capsids and are approximately 40Å in length. These densities appear to be located beneath triplexes T<sub>b</sub>, T<sub>c</sub>, and T<sub>e</sub>. Furthermore, it has been shown that interactions between the capsid and scaffold only take place with VP5s in hexons. These masses are thought to represent the positions at which the C-terminal of the scaffold interacts with the capsid shell (Zhou *et al.*, 1995; Zhou *et al.*, 1998b).

Through studies using the baculovirus system, it has been shown that preVP21 can substitute for preVP22a in the formation of capsids, although capsids are assembled at a greatly decreased efficiency and larger numbers of aberrant shells are formed, as compared to UL26.5 mediated assembly (Newcomb *et al.*, 1993; Thomsen *et al.*, 1994; Tatman *et al.*, 1994). Subsequently, it was demonstrated that in capsid assembly in the absence of UL26.5, there was no compensatory increase in the incorporation of VP21 and VP24 into the capsid (Sheaffer *et al.*, 2000). This led researchers to conclude that assembly could occur (albeit less efficiently) on a minimal scaffold of preVP21.

The mature forms of VP21 and VP22a are released from the capsid. It was proposed that in the procapsid, the onset of DNA packaging results in pH changes that result in the reversal of the self-interaction of the scaffold proteins, and that monomeric VP21 and VP22a are then able to leave the capsid through its numerous channels and pores, possibly by displacement when DNA enters (McClelland *et al.*, 2002). It is now accepted that the B-capsid is an abortive product resulting from failure in DNA packaging. This means that in B-capsids there would be no pH change and hence no disassociation (McClelland *et al.*, 2002). While VP21 and VP22a are removed from the capsid, VP24 is retained (Gibson and Roizman, 1972; Spear and Roizman, 1972; Booy *et al.*, 1991; Stevenson *et al.*, 1997). It has been proposed that this retention is mediated by retention signals carried by the protein, and that VP24 may have additional functions within the capsid, although these functions are as yet unknown (Rixon 1993; Sheaffer *et al.*, 2000).

## 1.5. Capsid assembly

Although the structure of the HSV-1 capsid has been under investigation for several decades, studies on the mechanisms by which it undergoes assembly began comparatively recently. Our understanding of the events and processes leading to capsid assembly has been aided by the introduction of insect cell-based protein expression systems. These have made it possible to develop *in vitro* capsid assembly systems and have revealed details of the mechanisms involved in the assembly of the capsid. A proposed model for HSV-1 capsid assembly is shown in Figure 1.18.



**Figure 1.18:** The main events in the HSV-1 virus assembly pathway showing the relationship between capsid assembly, maturation and DNA packaging. Copies of VP5 are depicted as open boxes, triplexes as crescents, and scaffolding proteins as triangles. Capsid assembly appears to be initiated at the UL6 portal complex. The release of the scaffold and the closure of the portal complex are poorly understood, and therefore they may differ in detail from the mechanisms shown.

Figure reproduced from Rixon and Chiu, (2003).

### **1.5.1. Nuclear assembly of HSV-1 capsids**

A characteristic feature of all herpesviruses is that DNA replication, capsid assembly and DNA packaging all occur inside the nucleus. Hence, HSV-1 capsid proteins translated in the cytoplasm must be transported to the nucleus. Nicholson *et al.* (1994) and Rixon *et al.* (1996) showed how the major structural components of the capsid go to the nucleus. VP19C and preVP22a are able to locate to the nucleus independently, which indicates that these proteins each contain a nuclear localisation signal (NLS), while VP5, VP23, and VP26 are not capable of independent nuclear localisation. PreVP22a associates with VP5, while VP19C associates with VP5 and VP23. In each case, the association results in the localisation of the proteins to the nucleus. VP26 does not localise to the nucleus when co-expressed with either VP19C or preVP22a alone, but nuclear localisation is observed when VP26 is co-expressed with VP5 and either VP19C or preVP22a. This result led the authors to conclude that VP26 binds directly to VP5, and the subsequent interaction of VP5 with VP19C or preVP22a results in VP26 nuclear localisation.

Some of the interactions involved in the nuclear localisation of capsid proteins have been confirmed by the yeast two-hybrid system. Using this system, Desai and Person (1996) detected interactions between VP19C and VP23, and between VP5 and VP22a. However, they did not detect the known interaction between VP5 and VP19C.

Detailed analysis of capsid protein nuclear localisation using confocal microscopy has revealed that capsid assembly takes place at defined foci in the nucleus, in close proximity to DNA replication compartments (Ward *et al.*, 1996; de Bruyn Kops *et al.*, 1998).

#### **1.5.1.1. HSV-1 capsid assembly analysis within insect cells**

The development of the baculovirus expression system has been significant in studies of protein interactions within the HSV-1 capsid (Bishop, 1992). Baculoviruses naturally infect arthropods, and they can be recombinantly manipulated to express single or multiple genes of interest within cultured insect cells. One of the most commonly used insect cell lines is that of *Spodoptera frugiperda* (SF). High levels of protein expression are achieved when SF cells are infected with recombinant baculoviruses that express gene(s) of interest under the strong late polyhedron promoter.

Tatman *et al.* (1994) and Thomsen *et al.* (1994) used the baculovirus expression system to identify the proteins involved in the assembly of capsids. Capsid particles that resemble wild-type B-capsids were produced by infecting insect cell lines with a panel of baculoviruses expressing six capsid protein genes: UL19 (VP5), UL18 (VP23), UL38 (VP19C), UL35 (VP26), UL26 (VP21/VP24), and UL26.5 (preVP22a). The omission of UL35 has little effect on capsid production, indicating that it is not required for capsid formation. However, capsid formation is prevented if UL19, UL18, or UL38 is omitted.

These studies also showed that neither of the scaffold protein genes, UL26 or UL26.5 is a *sine qua non* for capsid formation. Expression of either UL26 or UL26.5 is sufficient to serve as a scaffold for capsid formation. In the absence of UL26, large numbers of capsids form. The only apparent difference between these capsids and wild-type B-capsids is that they have larger scaffolds. By contrast, in the absence of UL26.5, the number of capsids produced decreases significantly. Moreover, large numbers of incomplete capsids are present; indicating that capsid formation in the absence of UL26.5 is inefficient. If both UL26 and UL26.5 are omitted, incomplete and distorted capsid shells form. If UL26.5 is expressed in the absence of other capsid proteins, large numbers of core-like structures form, indicating that preVP22a can self-assemble in the absence of other capsid proteins (Preston *et al.*, 1994).

The use of baculovirus expression systems has allowed the minimal requirements for HSV-1 capsid assembly (namely UL18, UL19, UL38, and either UL26 or UL26.5) to be identified. It should be emphasised that the HSV-1 capsid does contain additional proteins, such as the DNA packaging proteins UL6 and UL25. Nevertheless, the results described above demonstrate that although these proteins are essential for replication in mammalian cells, they are not required for the actual assembly of the capsids.

#### **1.5.1.2. In vitro HSV-1 capsid assembly and procapsid formation**

Recombinant expression of capsid proteins using the baculovirus expression system has led to the development of an *in vitro*-based capsid assembly model. *In vitro* studies have a significant advantage over cell-based analyses, in that they make it possible to study the assembly of the HSV-1 capsid under much more clearly defined experimental conditions. It was Newcomb *et al.* (1994) who showed that capsids were able to form from a mixture of lysed SF cell extracts that had been individually infected with viruses expressing VP5, VP19C, VP23, and preVP22a. The lysed cell extracts were incubated overnight at 28°C, and subsequent EM analysis revealed that the resulting capsids were similar to wild-type

capsids in terms of size and morphology ( $T=16$ ), and in the presentation of specific surface epitopes to monoclonal antibodies. The group went on to identify a series of intermediate capsid structures by the immunoprecipitation of complexes from *in vitro* assembly-competent reaction mixtures, with subsequent EM analysis (Newcomb *et al.*, 1996). Partial capsids, closed spherical particles (designated procapsids since they resemble the proheads formed in the capsid assembly pathways of dsDNA bacteriophages), and mature polyhedral capsids were identified after 1 minute, 90 minutes and 8 hours respectively of combined extract incubation. It was suggested that partial capsids are the precursors to procapsids – since they are rarely seen in reactions after 90 minutes of incubation.

When EM was carried out on procapsids (Newcomb *et al.*, 1996) they were seen to be highly uniform in morphology. They consist of a round closed outer shell with identifiable capsomer protrusions (composed of VP5) surrounding a scaffold core. The continued incubation of purified procapsids results in the production of mature angularised icosahedral particles. This led the authors to conclude that procapsids were the HSV-1 capsid assembly intermediate. Interestingly, partial capsids and procapsids disintegrated at 2°C, while angularised particles resembling those of mature B-capsids remained stable. The cold-sensitivity of partial capsids and procapsids is consistent with that seen in tailed bacteriophages that assemble through procapsid intermediates, such as T4 phage (Steven *et al.*, 1976).

Confirmation that the procapsid was an intermediate in the HSV-1 capsid assembly pathway came from work done by Rixon and McNab (1999). Cells were infected with a strain of HSV-1 that encodes a temperature-sensitive UL26 protease (*ts1201*) (Preston *et al.*, 1983). Cells infected with *ts1201* at the non-permissive temperature formed capsids that were cold-sensitive and which disassembled at 0°C in a similar manner to procapsids formed *in vitro* (Newcomb *et al.*, 1996). The lesion encoded by *ts1201* affects the ability of UL26 protease to cleave the scaffolding proteins at the maturation site. As a result, at the non-permissive temperature, the scaffolding proteins remain uncleaved and connected to VP5. This defect is reversible. When *ts1201* cells infected at the non-permissive temperature were incubated at the permissive temperature, the UL26 protease became activated and was able to cleave the scaffolding protein. Moreover, the capsids were able to undergo structural remodelling which allowed them to form mature, cold-stable polyhedral capsids (Preston *et al.*, 1983; Addison *et al.*, 1990; Rixon and McNab, 1999). As well as confirming the procapsid as an intermediate in the capsid assembly pathway, these observations implicated the protease as having a role in capsid maturation.

Current evidence suggests that capsid proteins are added to the growing procapsid shell as small preformed protein complexes rather than as individual proteins. This hypothesis was originally proposed by Trus *et al.* (1996). It was based on experiments showing that the nuclear localisation of HSV-1 capsid proteins requires that the capsids and scaffolding proteins interact in the cytoplasm prior to transport to the nucleus (Nicholson *et al.*, 1994; Rixon *et al.*, 1996). Subsequent experiments supported this model. Spencer *et al.* (1998) demonstrated that VP23 and VP19C initially interact to form heterotrimeric triplexes, before being incorporated into the assembling procapsid shell. Newcomb *et al.* (1999) demonstrated that VP5 and preVP22a interacted *in vitro*, forming complexes consisting of one or two copies of VP5 and up to six copies of preVP22a. On the basis of these results it was possible to add detail to the proposed model for procapsid assembly: VP5 and the scaffolding protein (preVP22a or UL26 protease) interact in the cytoplasm, forming small complexes. These complexes are transported to the nucleus, where they interact with preformed triplex subunits. Individual VP5/scaffolding protein complexes are cross-linked by the triplexes to form partial capsids. These are then extended by the addition of further complexes until a spherical procapsid is formed (Newcomb *et al.*, 1999). The importance of the interaction between the scaffolding protein and VP5 was further highlighted with the identification of hydrophobic residues in the N-terminus of VP5 that, when mutated, result in VP5 no longer being able to interact with the scaffold protein. These mutations also prevented the closure of capsid shells (Walters *et al.*, 2003). Recently, Newcomb *et al.*, (2005) demonstrated the importance of the portal in initiating the assembly of the procapsid (as discussed in Section 1.3.6), thereby adding further detail to our understanding of capsid assembly mechanisms.

Saad *et al.* (1999) infected insect cells with baculoviruses expressing only VP19C and VP5, and observed the formation of spherical particles. These were smaller than those observed in wild-type HSV-1, with a diameter of 880Å and  $T=7$  icosahedral symmetry. This would suggest that the formation of closed spherical structures with icosahedral symmetry does not require the presence of scaffolding proteins, and instead that it is an inherent property of capsid proteins. More recently, Newcomb *et al.* (2001b) examined cell-free assembly systems in which the concentration of preVP22a was varied while concentrations of VP5, VP19C and VP23 remained constant. They observed that below a threshold concentration of preVP22a the capsids produced were smaller. They suggested that they had  $T=9$  icosahedral symmetry, although this was not confirmed by reconstruction, indicating that the scaffolding protein concentration affects the structure of

nascent procapsids with a minimum amount required for the assembly of procapsids with the standard radius of curvature and scaffolding protein content. When scaffolding proteins are expressed in the absence of other capsid proteins, they self-assemble into ~600 Å diameter structures (Newcomb and Brown, 1991; Preston *et al.*, 1994; Tatman *et al.*, 1994; Kennard *et al.*, 1995). This suggests that scaffolding proteins have a role in preventing capsid proteins from forming smaller structures and in ensuring that they assemble into a shell of the correct size (Thuman-Commike *et al.*, 1999). It is interesting to note that if VP5, VP19C, and VP23 are co-expressed in the absence of scaffolding protein, no 880 Å particles are seen, although partial and distorted capsid shells are formed. This observation suggests that VP23 may modulate the interaction between VP19C and VP5. If this hypothesis is correct, one might conclude that VP23 could be considered to be an external scaffolding protein, acting in conjunction with the internal scaffold to form 1250 Å capsids. External scaffolding proteins have been described in other viruses, such as the ssDNA virus phiX174 (Dokland *et al.*, 1974). However as VP23 remains an integral part of the capsid in mature virions, it cannot be classed as a scaffolding protein.

#### **1.5.1.3. The structure of the procapsid**

The structure of the HSV-1 procapsid has been determined using cryo-EM and image reconstruction techniques (Trus *et al.*, 1996; Newcomb *et al.*, 1999; Newcomb *et al.*, 2000; Heymann *et al.*, 2003). These studies have demonstrated that although the procapsid is an icosahedral shell with  $T=16$  symmetry, there are a large number of visible differences between it and a mature HSV-1 capsid. The procapsid is more open and porous, with relatively little contact between neighbouring VP5 subunits in pentons and hexons other than at their bases. Furthermore, individual capsomers have no direct contact with the capsomers that surround them, and are held in place principally by the triplexes, indicating that the development of the capsid floor occurs at the stage of capsid maturation rather than initial assembly. The consequence of the loose association between the components of the procapsid is an increased fragility compared to the mature capsids – an observation that is reflected in the instability and resulting disassembly of procapsids at low temperatures (Newcomb *et al.*, 1996; Newcomb *et al.*, 2000). Several divergences from mature capsids have been noted in the hexons of procapsids. They do not exhibit six-fold symmetry, but form distorted oval (in the case of E- and P-hexons) and triangular shapes (C-hexons); moreover, the trans-hexon channel – a small circular opening in the mature capsid – is considerably wider in the procapsid. The triplexes appear more trimeric than those of mature capsids and VP26 does not bind to the tops of hexons in procapsids, presumably as a result of the difference in hexon structure (Chi and Wilson, 2000; Newcomb *et al.*, 2000).

#### 1.5.1.4. Maturation of the procapsid

The successful end product of procapsid maturation is the C-capsid: a mature capsid that has lost its internal scaffold, successfully packaged the DNA genome and undergone structural remodelling. The completion of these processes involves a number of different structural and molecular events. It has been proposed that these events are triggered by the activation of UL26 protease and the subsequent cleavage of scaffolding protein, and that they take place either simultaneously or in rapid succession during maturation (Gao *et al.*, 1994; Church and Wilson, 1997; Rixon and McNab, 1999; Newcomb *et al.*, 2000).

Sheaffer *et al.* (2000) demonstrated that in HSV-1, DNA packaging proteins interact with the procapsid – which is evidence that DNA packaging takes place, or is at least initiated in the procapsid state. It has also been shown that DNA packaging and the loss of the scaffold are linked (Dasgupta and Wilson, 1999), with DNA entry into the procapsid probably being the trigger for the disassociation of the internal scaffold following cleavage (McClelland *et al.*, 2002).

Although the successful result of procapsid maturation is the C-capsid, it is not an entirely efficient process. Cells also contain A-capsids (empty shells) and B-capsids (which retain a core composed of scaffold proteins). Both A- and B-capsids are abortive products, although angularisation takes place in A-, B-, and C-capsids. It has been suggested (McClelland *et al.* 2002) that the procapsid represents an unstable configuration for capsid shell proteins, and that it is only the interaction between VP5 and the scaffold that maintains the procapsid form. Thus, cleavage of the scaffolding protein renders the procapsid unstable, and as a result, the capsid proteins spontaneously rearrange themselves into a more stable configuration: the polyhedral capsid.

### 1.6. The capsid in virus evolution

All known herpesviruses share a characteristic four-component virion architecture, comprising a DNA core, capsid, tegument, and envelope. Our knowledge of herpesvirus capsid structures has been extended by the advent of cryo-EM. By this means, capsids have been defined to resolutions as high as 8.5Å in the case of HSV-1 (Zhou *et al.*, 2000).

### **1.6.1. Three-dimensional reconstructions of herpesvirus capsids**

Three-dimensional reconstructions for a number of alpha, beta, and gammaherpesviruses have now been generated: HSV-1 (Zhou *et al.*, 2000), HCMV (Butcher *et al.*, 1998; Chen *et al.*, 1999), SCMV (Trus *et al.*, 1999), HHV-8 (Wu *et al.*, 2000; Trus *et al.*, 2001), Equine herpesvirus 1 (Baker *et al.*, 1990), and Rhesus monkey rhadinovirus (Yu *et al.*, 2003). The appearances of hexons and pentons are similar in each case, with both structures being composed of the same protein and extending approximately 10 nm outwards from the capsid floor. A 3-D reconstruction has also been determined for CCV (Booy *et al.*, 1996), and although it represents a different class of herpesviruses (Section 1.1.1.5), the size and arrangement of the hexons and pentons in CCV is strongly suggestive of a similar pattern to that seen in alpha, beta, and gammaherpesviruses. More recently, a 3-D reconstruction of OsHV-1 has been carried out (Davison *et al.*, 2005b). Unlike other 3-D reconstructions of herpesvirus capsids, in OsHV-1 large voids are visible at the pentagonal vertices. However, this is thought to be due to the loss of pentons during preparation of the sample, so to date there is no data available concerning the penton structure of OsHV-1 (Davison *et al.*, 2005b).

All alpha, beta, and gammaherpesviruses examined by 3-D reconstruction contain triplexes, each consisting of two copies of one protein and one copy of a second protein. It is not known whether capsids from the fish and bivalve herpesviruses families are organised in this way, although CCV capsids contain suitably-sized candidate proteins with appropriate stoichiometries (Booy *et al.*, 1996; McGeoch *et al.*, in submission).

Although the information from CCV and OsHV-1 is less detailed, it is clear that the capsids of all herpesviruses studied display a common set of structural features. The 3-D structures that have been obtained provide evidence that all three families have a common evolutionary origin with respect to capsid architecture, and thus that they possess related sets of capsid proteins. The actual amino acid sequences of the respective proteins have, for the most part, diverged to the extent that there is no detectable homology between the three herpesvirus families. However, one protein provides an interesting exception to this observation, containing sequence features indicative of common descent, namely the ATPase subunit of terminase. This protein has a role in DNA packaging and thus interacts directly with the nascent capsid (McGeoch *et al.*, in submission). The presence of the conserved terminase has not been considered definitive evidence for a common origin of

the three disparate herpesviruses because the same relationship is seen with the ATPase subunit of bacteriophage T4 terminase, suggesting possible early horizontal exchange with a T4-like tailed bacteriophage (Davison *et al.*, 2002). However, recent evidence suggests that it might in fact be indicative of a wider evolutionary relationship that would encompass both herpesviruses and tailed bacteriophages.

### **1.6.2. Similarities between virus families**

Historically, although similarities had been observed in the coat proteins of small icosahedral viruses, the viral ‘universe’ was often viewed and studied as a collection of unrelated families. However, the discovery that P3, the major coat protein of the bacteriophage PRD1, has a very similar structural fold to that of hexon, the major coat protein of human adenovirus (Benson *et al.*, 1999; Benson *et al.*, 2002) altered the way in which the evolution of viruses was viewed.

PRD1 P3 and the adenovirus hexon are both trimeric molecules containing two jelly rolls (also known as  $\beta$ -barrels) per subunit, and both produce a pseudo-hexagonal base in the trimer. The adenovirus hexon is much larger than P3: there are 951 amino acids in adenovirus type 5, compared with 394 in P3. However, most of the additional residues in hexon form loops above the jelly rolls that lead to extensive, entangled towers (Figure 1.19). The similarity between PRD1 P3 and adenovirus hexon is the most striking of many structural similarities between PRD1 and adenovirus virions, which include virion architecture, vertex recognition spikes, and genomes with inverted terminal repeats and terminal proteins. These observations suggest that the two viruses are evolutionarily related (Benson *et al.*, 1999). The implication of this is that if an evolutionary relationship can be detected between viruses that infect two very different hosts who diverged billions of years ago, many other viruses that were previously thought to have arisen independently may in reality have originated from a common predecessor.

The relationship has now been extended, with more examples of viruses from families that were previously thought to be unrelated containing capsid proteins that exhibit similar folding patterns. The capsid of paramecium bursaria chlorella virus 1 (PBCV-1), an algal phycodnavirus, is composed of trimers of the Vp54 coat protein. This protein has been shown to contain two jelly rolls arranged similarly to adenovirus hexon and PRD1 P3 (Nandhagopal *et al.*, 2002). Recently, a similar arrangement was also described for the coat protein of *Sulfolobus* turreted icosahedral virus (STIV), a virus that infects the

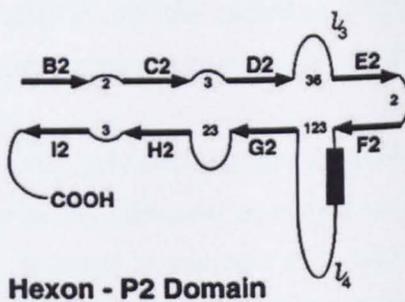
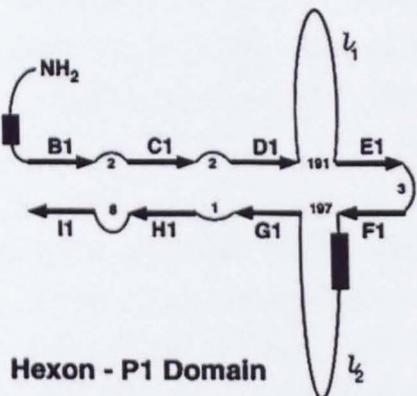
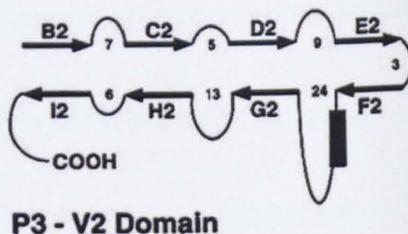
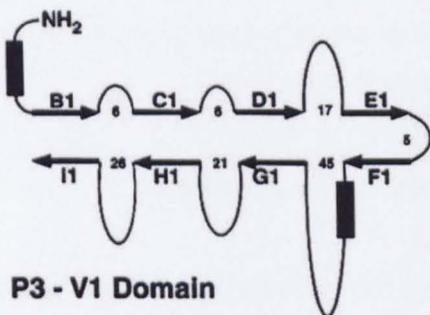
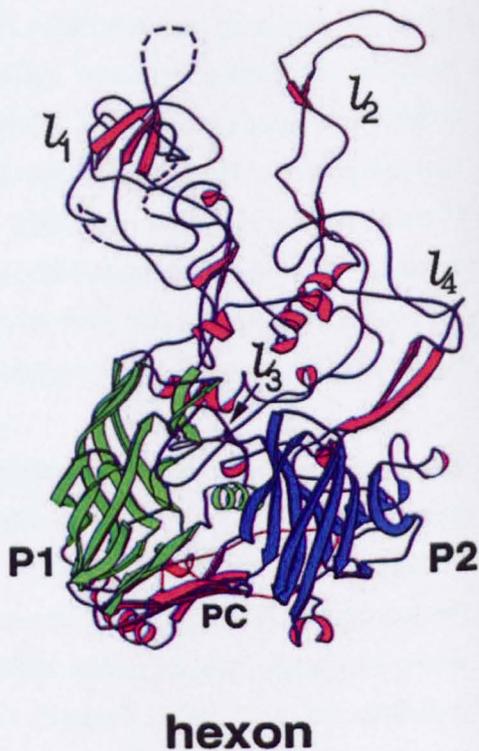
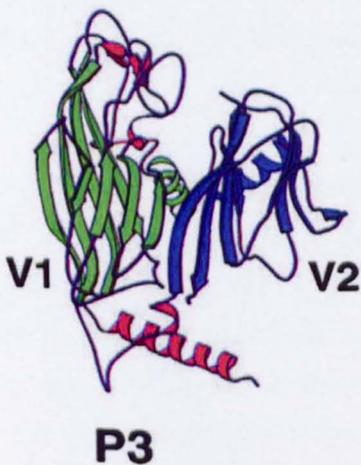
**A****B**

Figure 1.19: Comparison of the P3 and adenovirus hexon.

Schematic organisation of the two jelly rolls in P3 and adenovirus type 2 hexon. Adenovirus hexon is larger than P3, however most of the additional residues form loops, leading to extensive entangled towers in the trimer. **A.** The locations of  $\alpha$ -helices and loops are similar. The numbers of residues between the  $\beta$  strands are indicated. Arrows and black boxes indicate  $\beta$  strands and  $\alpha$ -helices respectively. **B.** Ribbon diagrams of P3 and hexon highlighting the jelly rolls (green and blue).

Figure reproduced from Benson *et al.* (1999).

hyperthermophilic archeon *Sulfolobus sulfataricus* (Rice *et al.*, 2004). The similar organisation of the capsid proteins from viruses that infect eukaryotes, bacteria and archaea suggests long-range evolutionary relationships between viruses from all three domains of life. It is possible that this similarity is a result of horizontal gene transfer between all three domains, however this would be unprecedented, and must have happened long enough ago for all sequence resemblance to have disappeared. If STIV, PBCV-1, PRD1 and adenovirus all evolved from a common ancestor, then that ancestor presumably must have existed prior to the divergence of eukaryotes, bacteria and archaea: more than three billion years ago. The striking similarity in the proteins of these viruses suggests that this is the case (Rice *et al.*, 2004).

### **1.6.3. Similarities between herpesviruses and dsDNA bacteriophages**

There is a growing body of evidence that demonstrates a possible evolutionary relationship between herpesviruses and tailed dsDNA bacteriophages, such as T4, P22 and  $\lambda$ . The hypothesis that these diverse viruses are related was originally based on parallels in capsid assembly pathways (Casjens and King, 1975; Steven and Spear, 1997). The likelihood of such a relationship existing was strengthened by analyses of HSV-1 portal complexes. However, until recently, no detailed evidence from protein structural comparisons between herpesviruses and tailed bacteriophages was available. However, the improvement in resolution of capsid reconstructions during recent years for both herpesviruses and tailed bacteriophages has progressed to the point that it has been possible to compare the capsids of HSV-1 and certain tailed bacteriophages. Such comparisons have revealed some previously unsuspected similarities.

#### **1.6.3.1. Parallels in capsid assembly**

Capsid assembly in tailed bacteriophage can be compared to that in herpesviruses. For example, a mature capsid of P22, a dsDNA bacteriophage that infects the bacterium *Salmonella tryphimurium* (Hendrix and Garcea, 1994; King and Chiu, 1997), consists of 420 copies of the coat protein gp5. It comprises 60 hexamers and 12 pentamers, and it is arranged in  $T=7$  icosahedral symmetry.

Assembly of the P22 capsid involves the interaction of gp5 with a scaffolding protein to form a spherical capsid intermediate. A maturation process then follows, which involves

extensive structural remodelling, the loss of the internal scaffold, and the packaging of viral DNA. The procapsid normally initiates around a portal complex made up of twelve copies of gp1. Recent studies have demonstrated that the HSV-1 procapsid also initiates around a portal complex (Section 1.3.6). The P22 procapsid grows by sequential addition of gp5 subunits around an internal core of 300 copies of the scaffolding protein. Following assembly, minor proteins such as gp16, gp20 and gp7 attach to the capsid. These proteins are required for the injection of DNA into a new host following attachment. However, as in HSV-1, these proteins, along with the portal complex, are nonessential for capsid assembly *in vitro*.

The P22 procapsid is spherical. In striking similarity to the HSV-1 procapsid, P22 capsomers appear skewed, but become more regularly hexagonal upon maturation, while holes that are present at the six-fold axes of the P22 procapsid are absent in the mature capsid. During maturation, the most extensive conformational changes take place at the trimer clusters. In another parallel with HSV-1, these three-fold symmetrical axes have been shown to be the positions at which the scaffold binds to the inner surface of the capsid shell. The observation that the site of scaffold binding is also that of considerable conformational rearrangement implicates the scaffold in both the assembly of P22 capsids, and also in the maturation of the capsid during DNA packaging.

### **1.6.3.2. Parallels in portal/terminase complexes**

In HSV-1 and some dsDNA bacteriophages, the replicated viral genome is a concatameric DNA molecule, which is cleaved into a unit length genome and packaged into a preformed capsid. This process involves two separate protein complexes: the portal (the channel through which viral DNA enters the capsid), and the terminase (an enzyme complex which recognises the viral genome and cleaves it into monomers).

In HSV-1, the portal is composed of twelve UL6 molecules, which form a ring-like structure found at a single vertex in the icosahedron as discussed in Section 1.3.6. Similar organisation is seen at the portals of tailed bacteriophages such as  $\lambda$ , which also contain a unique vertex composed of a dodecameric ring of twelve copies of the gpB protein, through which the DNA enters the capsid (Bazin et al., 1985; Valpuesta and Carrascosa, 1994).

In bacteriophage  $\lambda$ , the terminase consists of two subunits: gpA and gpNu1. These subunits interact through the N-terminal 48 amino acids of gpA, and the C-terminal 91 amino acids

of gpNu1. The N-terminus of gpNu1 binds specifically to the  $\lambda$  DNA packaging signal (*cos*). The gpA C-terminus is involved in procapsid binding, and in endonuclease and helicase activities (reviewed by Catalano, 2000).

It has been hypothesised that  $\lambda$  terminase mediates DNA cleavage by gpNu1 binding to the *cos* sequence in conjunction with a host cell protein known as integration host factor (IHF). GpNu1 then recruits gpA, which binds as a dimer to either side of the *cos* sequence, and its endonuclease activity results in symmetrical cleavage at the *cos* site. The two members of the gpA dimer then separate, causing the DNA strands to separate, effectively cleaving the DNA at the *cos* site. The result of these actions is a complex known as complex I, a stable intermediate in the DNA packaging process in which the mature left end of the viral genome is tightly bound by the terminase complex. This complex then interacts with the pre-assembled procapsid at the portal complex to form complex II. The viral genome is then driven into the procapsid in an ATP-dependent manner. As this takes place, gpNu1 'scans' the incoming viral genome until a second *cos* sequence is reached. This *cos* sequence is then cleaved by the gpA endonuclease, resulting in the reconstitution of complex I on the newly-formed end of the concatameric DNA, which can then interact with another procapsid (reviewed by Catalano, 2000).

In HSV-1, UL6 forms the portal channel, and the DNA packaging proteins UL15, UL28, and UL33 have been proposed as subunits of the terminase enzyme (Davison, 1992; Beard *et al.*, 2003; White *et al.*, 2003). It appears likely that all herpesviruses package their genomes in a similar fashion. However, while the  $\lambda$  terminase is known to contain two subunits, the composition of the terminase has not been determined in herpesviruses so it is not possible to determine how many components it contains.

### 1.6.3.3. Parallels in capsid structure

It was previously thought that  $T=16$  capsid architecture was unique to herpesviruses. However this arrangement has now also been observed in the bacteriophage SPO1. Furthermore, the SPO1 capsid has been suggested to contain heterotrimeric triplexes at positions of threefold symmetry (Duda *et al.*, 2006).

The secondary structure of HSV-1 VP5 has been compared to gp5, the capsid shell protein of the lambdoid bacteriophage HK97 (Baker *et al.*, 2005). Although gp5 is much smaller (282 amino acids compared with 1374 in VP5), it has very similar dimensions to the floor domain of VP5, and when comparison is confined to this domain (i.e. excluding the middle

and upper domains of VP5), the overall shapes of the two proteins and the dispositions of their secondary structural elements appear strikingly similar (Figure 1.20). Capsids from several more bacteriophages have now been analysed at sufficient resolution to reveal secondary structural features, and in each case, despite limited sequence similarity, a fold pattern similar to HK97 gp5 has been observed (Wikoff *et al.*, 2000; Jiang *et al.*, 2003; Fokine *et al.*, 2004; Morais *et al.*, 2005). No herpesvirus capsid structure other than that of HSV-1 is known in sufficient detail to carry out similar analyses, but the general uniformity of herpesvirus capsid structures makes it very likely that the floor domain fold in VP5 is maintained throughout herpesviruses.

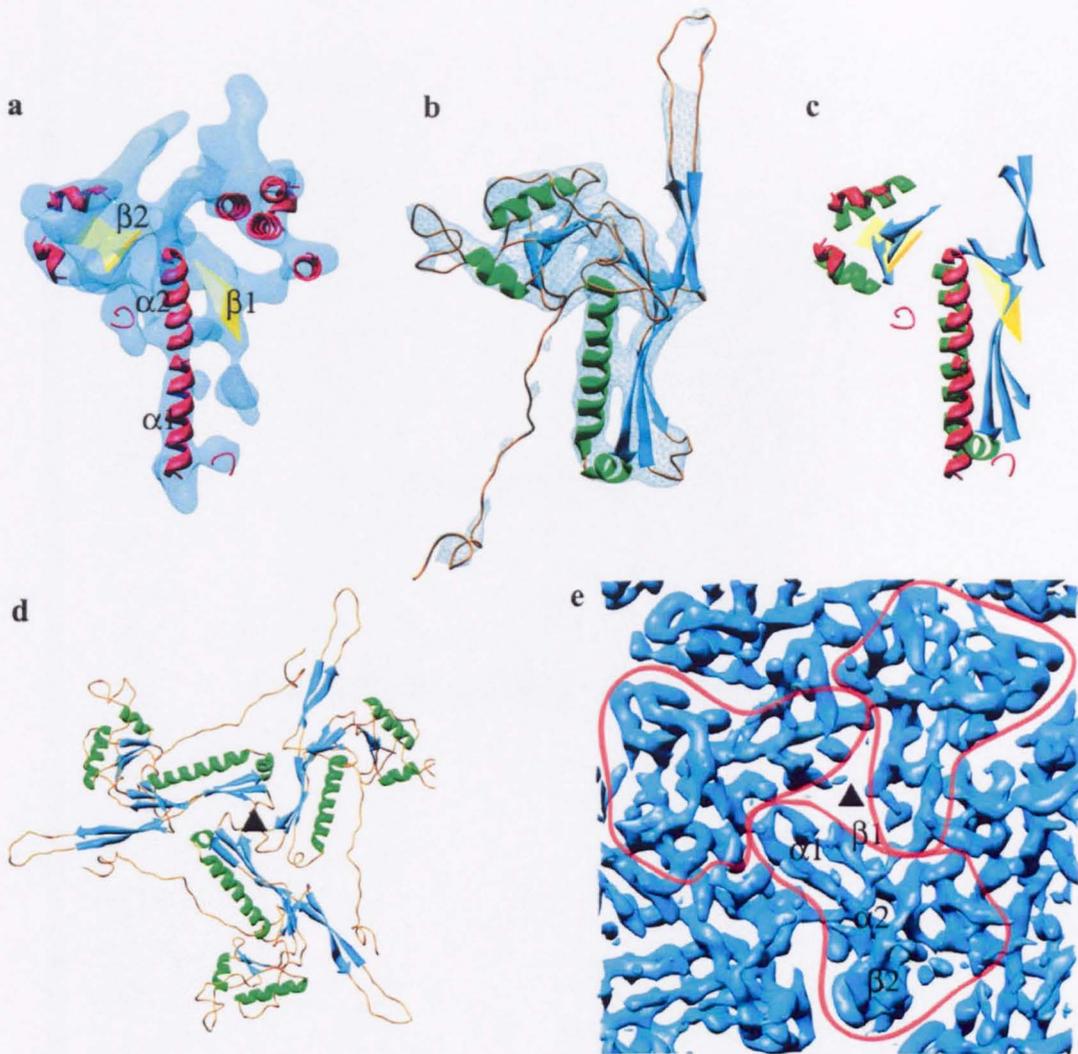
Taken together these findings indicate that the entire capsid packaging machinery is of ancient origin, and that it has been passed down to herpesviruses and tailed bacteriophages from a common progenitor. This helps to explain why tailed bacteriophages have provided such useful models in examining herpesvirus capsid assembly.

## **1.7. Experimental aims**

The aim of the work presented in this thesis was to add to the current understanding of the proteins that make up the HSV-1 capsid and the interactions that take place between them.

Insertional and deletional mutagenesis was carried out on the triplex protein VP19C in order to identify regions that were important for its function. The results of this work identified the N-terminal region of the protein as a particular region of interest, and led to attempts to map the nuclear localisation signal that this region contained, and attempts to identify the effect that removing amino acids from this region of the protein has on overall capsid morphology.

The remaining work aimed to further characterise the major capsid protein, VP5. The possibility of fluorescently labelling VP5 by amending the protein sequence to include a six amino acid motif was investigated. Amino acids in VP5 that were proposed to be responsible for interaction with VP26 were examined by producing a panel of VP5 mutants containing single amino acid substitutions. These particular topics are introduced in more detail at the beginning of the relevant results chapters (chapters 3, 4, 6, and 7).



**Figure 1.20:** The secondary structural elements of HSV-1 VP5 and HK97 gp5 and their molecular interactions in capsids. **(a)** The VP5 floor domain (turquoise), viewed from outside the capsid. The structure has been predicted to contain two long  $\alpha$ -helices (red;  $\alpha 1$  and  $\alpha 2$ ) adjacent to a large  $\beta$  sheet (yellow;  $\beta 1$ ) in the floor domain, as well as a second  $\beta$  sheet (yellow,  $\beta 2$ ), and several smaller helices flanking  $\alpha 1$  and  $\alpha 2$ . **(b)** HK97 gp5 shown in the same view reveals a similar structural motif. **(c)** Alignment of secondary structural elements demonstrates a clear match between the floor domain of VP5 and gp5. **(d)** Arrangement of gp5 subunits around a local three-fold axis (denoted with a black triangle) in the HK97 capsid, as viewed from inside the capsid. **(e)** Organisation of the HSV-1 capsid floor as shown in the same view as (d). Individual VP5 subunit floor domains are shown, with the long  $\alpha$ -helices ( $\alpha 1$  and  $\alpha 2$ ) and associated  $\beta$  sheets ( $\beta 1$  and  $\beta 2$ ) annotated in one subunit.

Figure reproduced from Baker *et al.*, 2005.

## **CHAPTER 2**

### **MATERIALS AND METHODS**

## 2. Materials and Methods

### 2.1. Materials

#### 2.1.1. Chemicals

The chemicals and reagents used for EM analysis were purchased from either Agar Aids or TAAB laboratories. All other chemicals and reagents were purchased from either Sigma-Aldrich or BDH Chemicals UK, unless otherwise stated below or in subsequent sections.

<b>Chemical</b>	<b>Supplier</b>
Acetic Acid	Rhone-Poulenc Ltd
Ammonium Persulphate	Biorad
Butanol	Prolabo
Chloroform	Prolabo
DMSO (dimethyl sulphoxide)	Koch-Light Laboratories Ltd
DNTPs	Pharmacia
ECL	Amersham Biosciences
Ethanol (ultra pure)	Fisher Scientific
Glacial Acetic Acid	Prolabo
Hydrochloric Acid	Rhone-Poulenc Ltd
Methanol	Prolabo
Nitro-Cellulose Membrane	Hybond
Rainbow Markers	Amersham Biosciences
S.O.C.	Invitrogen
Sodium Chloride	Prolabo
Tris-HCl	Roche

#### 2.1.2. Oligonucleotides

All oligonucleotides were supplied by either Sigma-Genosys or TAG Newcastle.

#### 2.1.3. Enzymes

All restriction enzymes and DNA modifying enzymes were supplied by Amersham Biosciences, Fermentas Life Sciences, New England Biolabs or Sigma-Aldrich.

#### 2.1.4. Cell lines

Baby hamster kidney-21 C13 (BHK-21) cells: A fibroblastic cell line originally derived from the kidney of a baby hamster (MacPherson and Stoker, 1962).

Rabbit skin cells expressing UL19 (UL19RSC): Rabbit skin cells are a fibroblastic cell line originally derived from the skin of a rabbit (Baines and Roizmann, 1991). In UL19RSCs, rabbit skin cells have been modified to constitutively express the HSV-1 major capsid protein VP5. Supplied by Dr. V. Preston.

Rabbit skin cells expressing UL38 (UL38RSC): As above, but modified to constitutively express the HSV-1 major capsid protein VP19C. (Thurlow *et al.*, 2005).

*Spodoptera frugiperda* 21 (SF21): Insect cells derived from ovarian tissue (Vaughn *et al.*, 1977).

Human bone osteosarcoma epithelial cells (U2OS): Cultivated from the bone tissue of a human female suffering from osteosarcoma (Ponten and Saksela, 1967). Supplied by Dr C. Boutell.

### **2.1.5. Tissue culture medium**

All tissue culture media and media supplements were obtained from Invitrogen. The growth media and the supplement requirements for each of the cell lines used were as follows:

BHK-21 cells: BHK-21 medium (Glasgow Modified Eagles Medium) supplemented with 10% tryptose phosphate broth and 10% newborn calf serum.

RSCUL19, RSCUL38 and U2OS cells: Dulbeccos Modified Eagles Medium supplemented with 10% foetal calf serum and 1% non-essential amino acids.

Sf21 cells: TC100 medium supplemented with 5% foetal calf serum.

Optimem: A commercially available growth medium that has a reduced serum content compared with other growth media; this medium was used during liposome transfections.

Overlay medium: BHK-21 medium (Glasgow Modified Eagles Medium) supplemented with 10% tryptose phosphate broth, 10% newborn calf serum, 2% L-glutamine, 0.26% sodium bicarbonate, 1% penicillin/streptomycin and 1.125% carboxymethylcellulose.

### **2.1.6. Viruses**

Wild-type HSV-1: strain 17 syn+, originally described by Brown *et al.*, 1973.

vΔ38YFP: Constructed by removing nucleotide sequences 84522 to 85982 in the HSV-1 strain 17 genome and replacing them with the yellow fluorescent protein ORF (Thurlow *et al.* 2005).

K5ΔZ: A mutated strain of HSV-1 in which the UL19 ORF has been disrupted by insertion of the LacZ gene (Desai *et al.*, 1993).

AcUL38: Baculovirus AcPAK6 vector expressing full-length VP19C. (Tatman *et al.*, 1994).

AcUL19: Baculovirus AcPAK6 vector expressing full-length VP5. (Tatman *et al.*, 1994).

AcUL26.5: Baculovirus AcPAK6 vector expressing full-length preVP22a. (Tatman *et al.*, 1994).

AcAB3.12: Baculovirus AcPAK6 vector expressing full-length VP23, VP26, and the UL26 protease: Made by cloning the ORFs into the triple expression transfer plasmid pAcAB3 (BD Biosciences, Pharmingen) and recombining them into AcPAK6 (Bishop, 1992).

Recombinant baculoviruses expressing modified VP19C proteins were produced using the BAC-TO-BAC Baculovirus Expression System (Invitrogen) following the manufacturer's protocol (Section 2.2.7).

### **2.1.7. Bacterial culture medium**

L-broth (LB)	10 g/l NaCl, 10 g/l Bacto™ peptone, 5 g/l Bacto™ yeast extract
L-broth agar (LB agar)	L-broth with 1.5% (w/v) Bacto™ agar
LB/LB agar + antibiotic	LB / LB agar supplemented with: Ampicillin at 100 µg/ml and/or kanamycin at 25 µg/ml (Unless alternative concentrations stated)
LB agar plates for BAC-TO-BAC Baculovirus Expression Systems (Gibco)	LB agar supplemented with: 50 µg/ml kanamycin 7 µg/ml gentamycin 10 µg/ml tetracycline 100 µg/ml Bluo-gal 40 µg/ml IPTG

### **2.1.8. Bacterial strains**

DH5 $\alpha$ : A strain of *E. coli* containing the *endA* and *recA* mutations (Invitrogen). The *endA* mutation eliminates endonuclease I activity, preventing degradation of plasmid DNA within the bacterial cell, thereby improving the quality of DNA preparations. The *recA* mutation reduces homologous recombination between plasmid DNAs, thereby increasing the stability of the plasmid within the bacterial cell.

GM48: A strain of *E. coli* that lacks both *dam* and *dcm* methylase activity, thereby producing plasmid DNA that is un-methylated at both *dam* and *dcm* recognition sites (Marinus, 1973). This strain is used to prepare plasmid DNA for situations where digestion with a methylation-sensitive restriction enzyme is required.

Note: Both of the above strains used were grown from our own glycerol stocks and were made competent for transformation using the method described in Section 2.2.3.3.

*E. coli* DH10B: Commercially available transformation competent bacterial cells (Invitrogen), used in situations where increased efficiency of transformation is required.

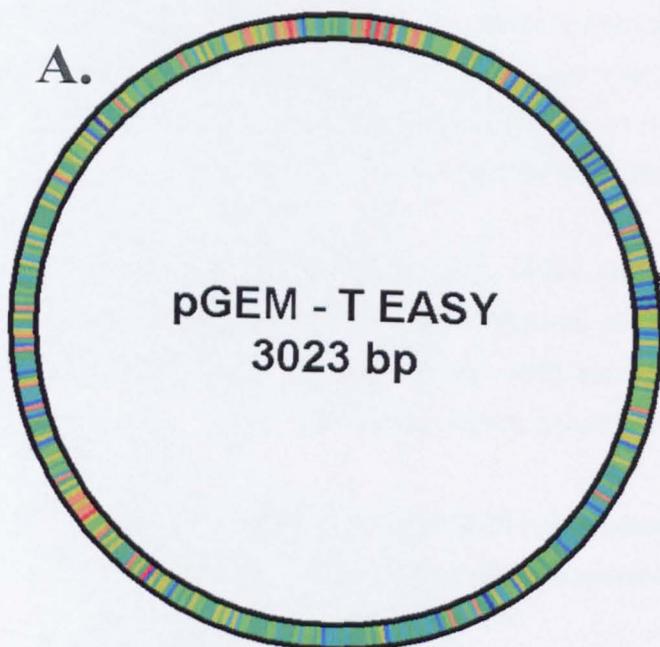
### **2.1.9. Plasmids**

pGEM-T Easy: Commercially available plasmid (Promega) used to directly clone PCR products into a plasmid vector. pGEM-T Easy has been linearised and has 3'-terminal thymidine nucleotides added. As *taq* polymerase adds a single adenine to the 3' end of PCR products, they can be directly cloned into pGEM-T Easy without further enzymatic manipulation (Figure 2.1).

pFASTBAC1: Commercially available plasmid (Invitrogen) containing a baculovirus-specific promoter (the polyhedrin promoter from *Autographa californica* nucleopolyhedrovirus) for expression of proteins in insect cells (Figure 2.1).

pFBpCI (Roger Everett, unpublished) was made by inserting a *BglII/EcoRI* fragment containing the human cytomegalovirus (HCMV) promoter/enhancer from pCI-neo (Promega) into *BamHI/EcoRI* digested pFASTBAC HTa (Invitrogen) (Figure 2.1).

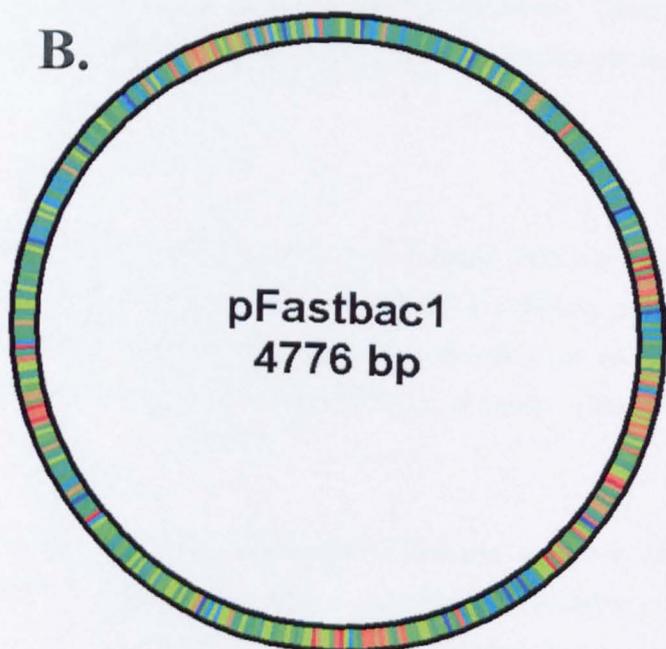
A.



pGEM - T EASY  
3023 bp

*Psp*OMI - 657 - G'GGCC\_C  
*Apa*I - 661 - G\_GGCC'C  
*Zra*I - 665 - GAC'GTC  
*Aaf*II - 667 - G\_ACGT'C  
*Sph*I - 673 - G\_CATG'C  
*Nco*I - 684 - C'CATG\_G  
*Sty*I - 684 - C'CwwG\_G  
*Sac*II - 696 - CC\_GC'GG  
**EcoRI - 699 - G'AATT\_C**  
*Spe*I - 719 - A'CTAG\_T  
**EcoRI - 725 - G'AATT\_C**  
*Bsp*MI - 732 - ACCTGCnnnn'n  
*Sbf*I - 743 - CC\_TGCA'GG  
*Pst*I - 743 - C\_TGCA'G  
*Sal*I - 745 - G'TCGA\_C  
*Acc*I - 746 - GT'mk\_AC  
*Hinc*II - 747 - G'Ty'rAC  
*Nde*I - 752 - CA'TA\_TG  
*Eco*ICRI - 762 - GAG'CTC  
*Sac*I - 764 - G\_AGCT'C  
*Mlu*I - 769 - A'CGCG\_T  
*Bst*XI - 773 - CCA nnnn'nTGG  
*Bfr*BI - 780 - ATG'CAT  
*Nsi*I - 782 - A\_TGCA'T

B.



pFastbac1  
4776 bp

*Bam*HI - 1107 - G'GATC\_C  
*Rsr*II - 1114 - CG'GwC\_CG  
*Bss*HIII - 1122 - G'CGCG\_C  
*Eco*RI - 1129 - G'AATT\_C  
*Stu*I - 1139 - AGG'CCT  
*Sal*I - 1145 - G'TCGA\_C  
*Eco*ICRI - 1153 - GAG'CTC  
*Sac*I - 1155 - G\_AGCT'C  
*Spe*I - 1157 - A'CTAG\_T  
*Not*I - 1165 - GC'GGCC\_GC  
*Bst*BI - 1174 - TT'CG\_AA  
*Xba*I - 1179 - T'CTAG\_A  
*Pst*I - 1191 - C\_TGCA'G  
*Xho*I - 1194 - C'TCGA\_G  
*Sph*I - 1204 - G\_CATG'C  
*Acc*65I - 1206 - G'GTAC\_C  
*Kpn*I - 1210 - G\_GTAC'C  
*Hind*III - 1212 - A'AGCT\_T

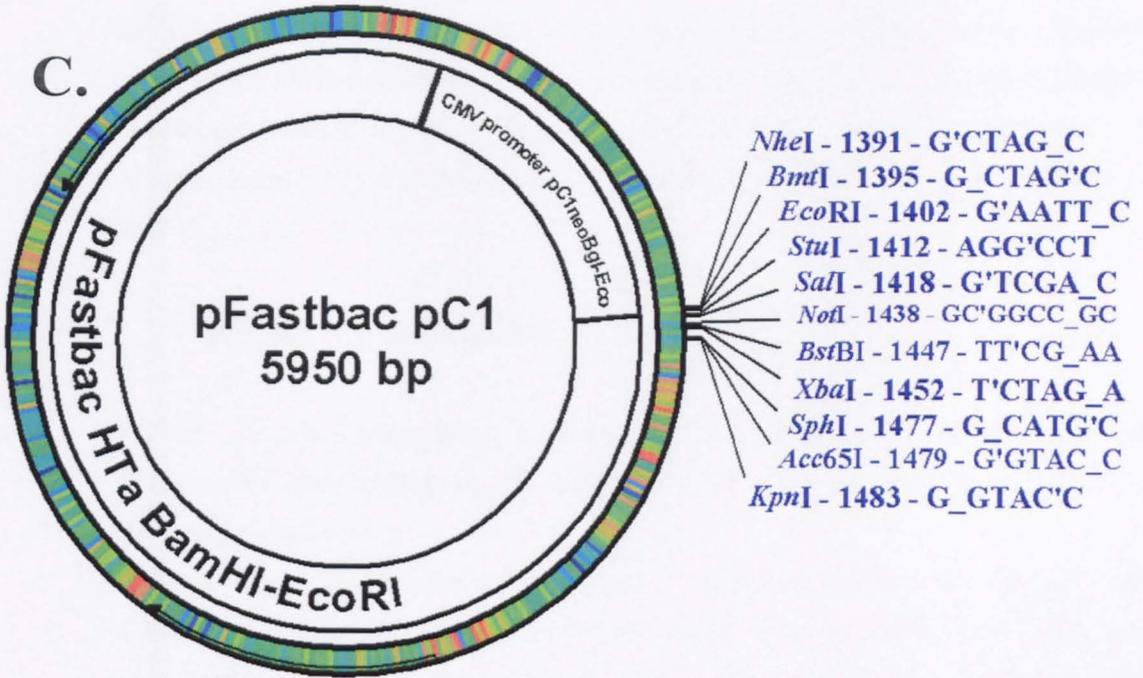


Figure 2.1: A. pGEM-T Easy. B. pFASTBAC1. C. pFASTBAC pC1.

Multiple cloning sites are shown for each plasmid.

PBS complete

PBSA + 6.8 mM CaCl<sub>2</sub>, 4.9 mM MgCl<sub>2</sub>.

PBSA/T

PBSA, 0.05% Tween-20.

Phenol/chloroform (1:1)

Mixed phenol/chloroform (1:1) saturated with 10 mM Tris (pH 8.0).

Alkaline Lysis Solution 1

25 mM Tris (pH 8.0), 10 mM EDTA, 50 mM glucose.

Alkaline Lysis Solution 2

0.2M NaOH, 1% SDS

Alkaline Lysis Solution 3

150 mM NaCl, 15mM trisodium citrate (pH 7.5)

10x Loading Buffer

40% (w/v) sucrose, 250 mM EDTA, 0.75% (w/v) bromophenol blue.

TBE

90 mM Tris (pH 8.0), 90 mM boric acid, 1 mM EDTA

TAE

40 mM Tris (pH 7.5), 20 mM glacial acetic acid, 1 mM EDTA pH 8.0

TBS

20 mM Tris-HCl (pH 7.5), 500 mM NaCl

Trypsin

0.25% (w/v) trypsin (Sigma-Aldrich) dissolved in TBS

Tryptose Phosphate Broth

10% (w/v) tryptose phosphate (Difco) dissolved in PBSA.

### Versene

600 mM EDTA dissolved in PBS A containing 0.0002% phenol red.

### Antibody Solution A

PBS complete, 5% newborn calf serum, 0.05% Tween-20

### Acid washing

Wash solution: 0.14 M NaCl

Acid solution: 0.14 M NaCl, 0.1 M glycine, pH 3.

### Formaldehyde Fix

PBSA, 2% sucrose, 2% formaldehyde

### 2x Cell Lysis Buffer (for packaging assay)

10 mM Tris-HCl (pH7.5), 1 mM EDTA, 0.6% SDS

### Resuspension Buffer (for packaging assay)

1 mM Tris-HCl (pH 7.5), 10 mM KCl, 1.5 mM MgCl<sub>2</sub>

### Resolving Gel Buffer

0.75 M Tris-HCl (pH 8.0), 1% SDS

### Stacking Gel Buffer

0.125 M Tris-HCl (pH 6.7), 0.1% SDS

### 3x SDS PAGE Sample Buffer

Stacking gel buffer: 25% SDS:β mercaptoethanol:glycerol (10:8:5:10)

### Tank Buffer

50 mM Tris, 50 mM glycine, 0.1% SDS

### Destain

Methanol:H<sub>2</sub>O:acetic acid (50:880:70)

### Coomassie Brilliant Blue Stain

0.2% Coomassie Brilliant Blue R250 in Destain

**Towbin Buffer**

40 mM Tris, 50 mM glycine, 20% methanol, 0.4% SDS

**Blocking Buffer**

PBSA/T, 5% Marvel milk powder

**Reticulocyte Standard Buffer (RSB)**

10 mM Tris-HCl (pH 7.5), 10 mM KCl, 1.5 mM MgCl<sub>2</sub>

### **2.1.12. Commercial kits**

QIAprep 8 Miniprep Kit	Qiagen Ltd
Sephaglas BandPrep Kit	Amersham Pharmacia Biotech Inc
pGEM-T Easy Vector System I	Promega
Expand High Fidelity PCR System	Roche
FLAsH <sup>TM</sup> -EDT <sub>2</sub> Labeling Kit	Panvera
Mutation Generation System	Finnzymes

### **2.1.13. Miscellaneous reagents**

Ampicillin	Beecham Research Laboratories Ltd
Gene Pulser ® cuvettes, 0.1cm electrode	Biorad
100 bp ladder, Lipofectamine reagent, Plus <sup>TM</sup> reagent.	Invitrogen
1 Kb ladder, lambda DNA- <i>Bst</i> E II digest	New England Biolabs

## **2.2. Methods**

### **2.2.1. Gel electrophoresis**

#### **2.2.1.1. Analytical agarose gel electrophoresis of DNA**

1% agarose/TBE gels were prepared by mixing the agarose in 1x TBE and heating the solution in a microwave oven until the agarose had completely dissolved. The solution was allowed to cool to approximately 50°C before ethidium bromide was added to a concentration of 0.5 µg/ml. When the gel had set it was submerged in 1x TBE buffer containing 0.5 µg/ml ethidium bromide. DNA samples were mixed at a ratio of 10:1 with 10x loading buffer, before being loaded on to the gel alongside appropriate size markers and electrophoresed at 100 V for 30-40 minutes. The DNA was examined and photographed under short-wave UV transillumination.

#### **2.2.1.2. Preparative agarose gel electrophoresis of DNA**

1% agarose gels were prepared using the procedure described above (analytical gel electrophoresis), except that TAE was used in place of TBE. DNA samples were mixed at a ratio of 10:1 with 10x loading buffer, before being loaded on to the gel alongside appropriate size markers and electrophoresed at 60 V for 60-90 minutes. If very small DNA fragments were to be purified, the length of time that the gel was electrophoresed was reduced. The gel was examined under long-wave UV transillumination and the desired DNA fragments were cut from the gel using a sterile scalpel. Gel slices were transferred to a 1.5 ml microfuge tube and stored at -20°C until required for purification (Section 2.2.2.4).

#### **2.2.1.3. Polyacrylamide gel electrophoresis of DNA**

5% acrylamide/TBE gels were prepared by mixing 3.5 ml 30% acrylamide-bis (Bio-Rad), 2 ml 10x TBE, 14.3 ml H<sub>2</sub>O, 20 µl TEMED (Sigma-Aldrich), and 200 µl 10% ammonium persulphate. The mixture was immediately poured between the glass plates of a Biorad Mini-Protean II electrophoresis apparatus and a comb was inserted into the top of the gel. Once set, the samples to be analysed were mixed with 10x loading buffer at a ratio of 10:1 and loaded on to the gel. All the samples were run alongside the 100 bp ladder. The gel was electrophoresed in 1x TBE at 100 V for 20-30 minutes until the bromophenol blue dye

front had reached the bottom of the gel. However, if very small DNA fragments were being analysed, electrophoresis was stopped before the dye front reached the bottom of the gel. The gel was stained by submerging it for 30 minutes in a solution of 1x TBE containing 0.5µg /ml ethidium bromide. The gel was then examined by short-wave UV transillumination and photographed.

## **2.2.2. DNA cloning and manipulation**

### **2.2.2.1. Polymerase chain reaction**

Each PCR reaction was set up in a thin-walled 0.2 ml PCR tube (Applied Biosystems) using the enzyme and buffer from the Expand™ High Fidelity PCR system. A typical PCR reaction mixture contained the following ingredients:

Template DNA (50-100ng)	1 µl
Forward primer (100pmol/µl)	0.5 µl
Reverse primer (100pmol/µl)	0.5 µl
5x PCR buffer	5 µl
2mM dNTP mix	2 µl
Taq expand (1.7 units)	0.5 µl
H <sub>2</sub> O	41.5 µl

The reaction's components were mixed by flicking the tube, taking care not to introduce air bubbles. PCR reactions were performed using an Applied Biosystems GeneAmp 9700 thermocycler running the following PCR profile:

Step 1: 1 cycle	96°C for 4 minutes
	55°C for 2 minutes
	72°C for 3 minutes

Step 2: 35 cycles	96°C for 1 minute
	55°C for 2 minutes
	72°C for 3 minutes

Step 3: Constant temperature 4°C

To determine whether the PCR reaction was successful, 5  $\mu\text{l}$  of the reaction was mixed with 0.5  $\mu\text{l}$  10x loading dye and analysed by agarose gel electrophoresis. A frequent drawback of this method is the presence of unwanted minor bands. Thus a comparison of the DNA profile of the PCR reaction with DNA size markers is essential to identify the desired product. If required, the remainder of the PCR reaction was mixed with 4.5  $\mu\text{l}$  10x loading dye and run on a preparative agarose gel. The desired band was then extracted and purified.

### **2.2.2.2. Restriction enzyme digestion**

#### **Analytical digests**

0.2-0.5  $\mu\text{g}$  of plasmid DNA was digested in a 20  $\mu\text{l}$  reaction containing the appropriate restriction buffer, 2  $\mu\text{g}$  BSA (if recommended) and 10 units of the desired restriction enzyme. The mixture was incubated at the recommended temperature for 60 minutes.

In the case of double digests, the buffer and temperature requirements for the two restriction enzymes were taken into consideration. If the enzymes functioned well under the same conditions, both could be added to the same reaction. However, if they had incompatible requirements the restriction digest had to be carried out in two sequential reactions. In this case, the DNA was digested with the first enzyme for 60 minutes in a 20 $\mu\text{l}$  reaction, under the conditions optimal for that enzyme. After 60 minutes, the reaction mixture was modified by adding the second enzyme, 3  $\mu\text{l}$  of the appropriate restriction buffer for that enzyme, and the volume of the reaction was increased to 30  $\mu\text{l}$ . The DNA was then digested for a further 60 minutes under the new conditions. All restriction digests were confirmed by analytical gel electrophoresis.

#### **Preparative digests**

The same basic principles as those used in analytical digests apply. The differences were that typically, 1-3  $\mu\text{g}$  of plasmid DNA and 20 units of restriction enzyme were used in the reaction. In addition, to confirm that the digest was successful, aliquots of the reaction were examined by analytical gel electrophoresis before the remainder of the reaction was purified using a preparative agarose gel.

### **2.2.2.3. De-phosphorylation of digested DNA**

To reduce self-ligation, the digested vector DNA was treated with calf intestinal phosphatase (CIP), since de-phosphorylation of the ends of the vector DNA renders them

unable to self-ligate. The CIP digestion was performed after restriction digestion by adding 10 units (1  $\mu$ l) of CIP directly to the digest reaction and incubating for a further 30-45 minutes at 37°C. CIP is functional in all restriction buffers.

#### **2.2.2.4. Purifying DNA fragments from a gel slice**

Purification of DNA fragments from agarose/TAE gel slices was performed using the Sephaglas<sup>TM</sup> BandPrep Kit in accordance with the standard protocols supplied. The yield and purity of the purified fragment could be determined by running a small fraction (typically 10%) on an analytical agarose gel.

#### **2.2.2.5. DNA ligation**

##### **Standard ligation**

Purified insert and vector DNA pre-digested with compatible restriction enzymes were mixed at an approximate 3:1 molar ratio in a 20  $\mu$ l reaction containing 1 unit of T4 DNA ligase and 1x T4 ligase buffer. The reaction components were mixed by flicking the tube and incubated overnight at 15°C. Following overnight incubation, the ligation reaction was stored at -20°C until required.

##### **Ligation of oligonucleotides**

Complementary nucleotides (0.5  $\mu$ g of each) in 20  $\mu$ l of H<sub>2</sub>O were heated for two minutes at 100°C, then transferred to a 37°C waterbath and allowed to cool for two minutes. 5  $\mu$ l of 5x T4 ligase buffer was added, after which the reaction was returned to the 37°C waterbath and incubated for a further 20 minutes. At this point, one unit of T4 ligase and 1  $\mu$ g of appropriately digested and de-phosphorylated vector DNA were added. The reaction mixture was transferred to a 15°C waterbath and incubated overnight.

##### **Ligation of PCR products**

1  $\mu$ l of gel-purified PCR product was ligated into the pGEM-T Easy vector in accordance with the standard ligation protocols supplied. The identity of the cloned PCR product was confirmed by DNA sequencing and restriction digestion.

#### **2.2.2.6. Small-scale plasmid DNA preparation ('miniprep')**

1.5 ml of a bacterial culture (Section 2.2.3.1) was transferred to a 1.5 ml microfuge tube and pelleted by centrifugation at 13 000 rpm for one minute in a bench-top microfuge. The supernatant was discarded and the pellet was processed using the QIAprep 8 Miniprep Kit

in accordance with the standard protocols supplied. The only deviation from this protocol was an increase in the elution volume from 100  $\mu\text{l}$  to 150  $\mu\text{l}$ , which allowed improved recovery of the purified plasmid. The vacuum manifold used was a Qiagen QIAvac 6S, used with a KNF vacuum pump (model number N022AN.18). Following purification, the miniprep DNA was analysed by restriction digestion and agarose gel electrophoresis.

In situations where the DNA under preparation was to be used for further cloning steps other than analysis or sequencing (such as in the production of intermediate plasmids during the process of creating random insertional mutants), the preparation was carried out using an alkaline lysis miniprep. Cells were pelleted as described above. The pellet was then resuspended in 100  $\mu\text{l}$  of ice-cold alkaline lysis solution 1 (Materials) and incubated on ice for five minutes. 200  $\mu\text{l}$  of freshly prepared alkaline lysis solution 2 (Materials) was added and mixing was carried out by inverting the microfuge tube six or seven times. 150  $\mu\text{l}$  of ice-cold alkaline lysis solution 3 (Materials) was then added. Mixing was again carried out by inversion and the mixture was incubated on ice for five minutes. The resulting white precipitate, consisting of cell debris, was pelleted by centrifugation at 13 000 rpm for five minutes at RT in a microfuge. The supernatant was transferred to another 1.5 ml reaction tube and the DNA was recovered by phenol/chloroform extraction and ethanol precipitation (Section 2.2.2.7). The DNA pellets were typically resuspended in 50  $\mu\text{l}$  of  $\text{dH}_2\text{O}$  containing 10  $\mu\text{g}/\text{ml}$  of RNaseA.

#### **2.2.2.7. Phenol/chloroform extraction and ethanol precipitation of DNA**

If the volume of the original DNA solution was less than 200  $\mu\text{l}$ , the volume was brought up to 200  $\mu\text{l}$  by the addition of distilled  $\text{H}_2\text{O}$ . An equal volume of phenol/chloroform (1:1) was added to the DNA solution and briefly vortexed before being centrifuged at 13 000 rpm for five minutes at RT. The upper (aqueous) phase was then removed. If the aqueous phase appeared cloudy then the phenol/chloroform extraction was repeated until the aqueous phase appeared clear. The DNA was precipitated by the addition of a 1/10 volume of 3 M sodium acetate and two volumes of 100% ethanol. The mixture was inverted several times and incubated for at least 30 minutes at  $-20^\circ\text{C}$ . Precipitated DNA was pelleted by centrifugation at 13 000 rpm for five minutes at RT in a microfuge. The DNA pellets were washed in 200  $\mu\text{l}$  of 70% ethanol to remove excess salts and the DNA was pelleted by centrifugation at 13,000 rpm for two minutes at RT in a microfuge. The DNA pellets were then either dried for ten minutes in an Eppendorf Concentrator 5301 or

allowed to air dry before being resuspended in the appropriate volume of d H<sub>2</sub>O containing 10 µg/ml RnaseA and stored at -20°C until required.

#### **2.2.2.8. Oligonucleotides**

All the oligonucleotides ordered from commercial sources were supplied as a lyophilised powder. This was resuspended in a volume of H<sub>2</sub>O that would provide a final oligonucleotide concentration of 100 pmoles/µl, vortexed vigorously, then incubated briefly at 37°C to ensure that all the material had dissolved.

#### **2.2.2.9. DNA sequencing**

All sequencing was carried out by either the Molecular Biology Support Unit at Glasgow University, or by Claire Addison at the MRC Virology Unit, Glasgow. The quality of the sequence data was confirmed by analysis of the chromatogram using a sequence viewing program called Chromas. This was obtained via the Internet from the URL: <http://www.technelysium.com.au>.

### **2.2.3. Bacterial methods**

#### **2.2.3.1. Bacterial culture**

##### (i) Large scale cultures

A 10ml starter culture was set up by inoculating 10ml of LB broth + antibiotic with a single isolated colony from a LB agar plate or from a glycerol stock (Section 2.2.3.5) and incubating overnight at 37°C in an orbital shaker at approximately 200 rpm. The following day, 100 ml of LB broth + antibiotic was inoculated with the 10 ml starter culture and grown overnight as before. Plasmid DNA was prepared as described above.

##### (ii) Miniprep cultures

2.5 ml of LB + antibiotic was inoculated with a single isolated colony from a LB agar plate and incubated overnight at 37°C in an orbital shaker at 200 rpm. On the following day, the culture was harvested to make DNA as described above (Section 2.2.2.6).

### 2.2.3.2. Preparation of electrocompetent *E. Coli* cells

*E. coli* cells were recovered from a  $-70^{\circ}\text{C}$  glycerol stock, placed in a 10ml starter culture of unsupplemented LB (containing no antibiotics) and grown overnight. On the following day, the starter culture was used to inoculate 100 ml of unsupplemented LB and again grown overnight. 50 ml of this culture was used to inoculate 1 litre of unsupplemented LB and grown until the culture reached an OD<sub>600</sub> of 0.5-0.6.

The culture was transferred to six pre-chilled 250 ml Falcon tubes (each containing approximately 160 ml of culture) and incubated on ice for at least 30 minutes. The cells were pelleted by centrifugation at 3500 rpm for 15 minutes at  $0^{\circ}\text{C}$  using a Sorvall SLA-1500 rotor. The supernatant was discarded and each of the six pellets was resuspended in 160 ml of ice-cold, sterile  $\text{H}_2\text{O}$ .

The cells were pelleted again as before and the supernatants discarded. The six pellets were resuspended in a total of 500 ml of ice-cold, sterile  $\text{H}_2\text{O}$  and combined to give  $\sim 160$  ml in each of three 250 ml Falcon tubes. The cells were pelleted again as before and the supernatants discarded. The three pellets were pooled and resuspended in a total volume of 40 ml ice-cold, sterile 10% glycerol, transferred to a pre-chilled 50 ml SS34 tube and centrifuged at 5800 rpm for 15 minutes at  $0^{\circ}\text{C}$  using the Sorvall SS34 rotor. The supernatant was discarded and the pellet resuspended in 2 ml ice-cold, sterile 10% (v/v) glycerol. The competent cells were aliquoted into pre-chilled sterile 1.5 ml microfuge tubes (40  $\mu\text{l}$  per aliquot) and stored at  $-70^{\circ}\text{C}$  until required.

### 2.2.3.3. Transformation of *E. coli* by electroporation

The transformation efficiency of electro-competent *E. coli* cells is reduced with repeated freeze-thaw cycles. For this reason, only as many vials as required for each experiment were removed from the  $-70^{\circ}\text{C}$  freezer and kept on dry ice. The cells were thawed on ice immediately prior to use.

5-10 ng of plasmid DNA, or 1  $\mu\text{l}$  of a ligation reaction was transferred to a Gene Pulser cuvette (Biorad). 25  $\mu\text{l}$  of electro-competent cells (Section 2.2.3.2) were then added to the cuvette and mixed with the plasmid DNA or ligation reaction by pipetting up and down. The cuvette was placed in a Hybaid Cell Shock, and the apparatus set to 1800 V and activated. The electroporated cells were resuspended in 1 ml of S.O.C. medium, transferred to a 15 ml Falcon tube and incubated for 1 hour at  $37^{\circ}\text{C}$  in an orbital shaker. 50-200  $\mu\text{l}$

aliquots of the transformation reaction were plated out on to LB agar plates containing the appropriate antibiotic and incubated overnight at 37°C.

#### **2.2.3.4. Making glycerol stocks**

1 ml of bacterial culture was transferred to a 1.5 ml microfuge tube. The cells were pelleted by centrifugation at 13 000 rpm for 1 minute and the supernatant discarded. The cell pellet was then resuspended in 250 µl 2% Bacto™ peptone (Difco) and 250 µl 80% glycerol, vortexed thoroughly, labelled, and stored at -70°C.

#### **2.2.3.5. Recovering cells from glycerol stocks**

The culture to be recovered was removed from the -70°C freezer and kept on dry ice. Using a sterile toothpick, a small amount of the frozen glycerol stock was used to inoculate a 10 ml starter culture as described above (Section 2.2.3.1).

### **2.2.4. Cell culture methods**

#### **2.2.4.1. Mammalian cell culture**

##### **(i) Passage**

BHK-21 cells were passaged routinely every 3-4 days in 850 cm<sup>2</sup> roller bottles. The cells were seeded at  $2 \times 10^7$  cell/bottle in 100 ml of BHK-21 medium (Section 2.1.5), gassed with CO<sub>2</sub> to 5%, and rotated slowly at 37°C. To harvest the confluent monolayers, the growth medium was removed, the cells were rinsed twice, briefly, in 20 ml of trypsin/versene (1:4) and incubated at 37°C until the cells had detached. The cells were resuspended in 20 ml of BHK-21 medium. The cell density was measured by placing a diluted aliquot of the cell suspension on to a haemocytometer (Neubauer) and counting the cells under an inverted light microscope (Olympus). The typical yield from an 850 cm<sup>2</sup> roller bottle was 20 ml of cell suspension at approximately  $1 \times 10^7$  cells/ml. Cells in suspension remained viable for up to one week at 4°C.

Rabbit skin cells and U2OS cells were passaged in 150 cm<sup>2</sup> flasks. The procedure was essentially the same as for BHK-21 cells except that they were grown in Dulbeccos medium (Section 2.1.5).

(ii) Freezing and storage

Cells grown in a 150 cm<sup>2</sup> flask were harvested by trypsinisation as described above (Section 2.2.4.1i). After removal from the flask, the cells were pelleted by centrifugation at 1000 rpm for 10 minutes using a Sorvall RT7 centrifuge. The growth medium was discarded and the cells were resuspended in 3 ml of 'freezing medium', i.e. 40% appropriate growth medium (minus calf serum), 40% calf serum, and 20% glycerol. The medium was aliquoted into 3 screw cap cryo-vials, placed in a polystyrene container, and stored at -70°C overnight before transfer to liquid nitrogen for long-term storage.

(iii) Recovery of frozen stocks

One vial of cells was thawed quickly in a water bath at 37°C and recovered into a 150 cm<sup>2</sup> flask containing the appropriate growth medium and incubated at 37°C. The medium was changed the following day and, when confluent, the cells were passaged in the normal way (Section 2.2.4.1i).

#### **2.2.4.2. Sf21 cell culture**

Sf21 cells were passaged every 3-4 days in 150 cm<sup>2</sup> tissue culture flasks containing 50 ml of TC100/5 medium. The cells were removed from the flask by the addition of 10 ml of TC100/5 medium and gentle tapping of the flask. Cell clumps were broken up by pipetting the medium up and down. Confluent flasks were typically passaged using a 1:4 split. Where appropriate, cell densities were calculated using a Neubauer counting chamber. Cells were counted using an inverted light microscope (Olympus).

(i) Freezing and storage of Sf21 cells

Sf21 cells were grown in 150 cm<sup>2</sup> flasks at 28°C until confluent. They were then harvested in 10 ml of TC100/5 medium and the cell suspension transferred to a 15 ml Falcon tube. Cells were pelleted by centrifugation at 1500 rpm for five minutes in a Sorvall RT7 centrifuge. The supernatant was removed and the cells were resuspended in 2 ml TC100/10. An equal volume of ice-cold freezing mix (80% TC100/10, 20% DMSO) was added. The cell suspension was aliquoted into cryogenic vials (1 ml/vial) before being frozen overnight at -70°C. Frozen cells were then transferred to liquid nitrogen for long-term storage.

(ii) Recovery of Sf21 cells

Aliquots of Sf21 cells removed from liquid nitrogen were thawed quickly in a water bath at 37°C before being resuspended in a small volume of TC100/5 (approximately 5 ml). The

cell suspension was then transferred to a 15 ml Falcon tube and cells were pelleted by centrifugation at 1000 rpm for five minutes in a Sorvall RT7 centrifuge. The supernatant was discarded and the cell pellet was resuspended in TC100/5, which was used to seed to a medium flask (200 ml volume). The flask was incubated at 28°C until the cells had firmly attached to the solid substrate. The medium was then replaced with fresh TC100/5. Once the monolayer had reached confluency, the cells were passaged into a large flask and cultured in the normal way.

#### **2.2.4.3. Liposome transfection of mammalian cells**

The procedure was carried out as described in the protocols supplied with the lipofectamine reagent (Invitrogen). Cells were set up in 24 well plates (Nunc) at a density of  $1 \times 10^5$  cells/well and incubated overnight at 37°C. On the following day, the DNA to be transfected was diluted in Optimem. For transfection with a single plasmid, 0.5 µg of DNA was required and for double and triple transfections, 0.25 µg of each plasmid was used. In each case, the DNA was diluted to a final volume of 21 µl with Optimem. 4 µl of Plus reagent was added to the diluted DNA/Optimem, then the solution was mixed and incubated at room temperature for 15 minutes. During this incubation, lipofectamine reagent was diluted 1 in 25 with Optimem. 25 µl of diluted lipofectamine was added to each DNA sample and mixed. The sample was then incubated for a further 15 minutes at room temperature. During the incubation, the cells to be transfected were washed twice with unsupplemented growth medium and then overlaid with 200 µl of the same medium. The transfection mix (50 µl) was added to the cells and the cells were incubated at 37°C. At three hours post-transfection, the medium was removed and 1 ml of supplemented growth medium was added to each well. The cells were then incubated at 37°C. The cells were now ready for processing as required, either for virus yield or immunofluorescence.

In transfections carried out in 40 mm plates (Nunc), cells were transfected in a similar fashion to that described above, with the modification that they were set up at a density of  $10^6$  cells/plate, and the volumes and concentrations of the reagents used during transfections were increased fourfold.

If the cells were to be processed for immunofluorescence (Section 2.2.4.5) or Green Fluorescent Protein fluorescence (Section 2.2.4.6), they were grown on 13 mm coverslips placed at the bottom of the wells of the 24 well plate. This allows the cell sheet to be easily removed and manipulated. However, if the cells were to be harvested for titration

following superinfection (Sections 2.2.4.4 and 2.2.4.7), no coverslips were required and the cells were grown directly in the base of the well.

#### **2.2.4.4. Superinfection of mammalian cells with HSV-1 and acid washing**

##### **(i) Superinfection with HSV-1**

The cells to be infected were transfected using the procedure described above (Section 2.2.4.1) At five hours after transfection the growth medium was removed from the wells and the cells washed with supplemented growth medium.

For BHK-21 cells, virus to be infected was added to the cells at a titre of 2 pfu per cell in a total volume of 100  $\mu$ l. The cells were incubated in the presence of the virus for one hour at 37°C.

##### **(ii) Acid washing**

After one hour of absorption, the medium containing the virus was removed and the cells were washed twice in 1 ml of wash solution (Section 2.1.11). The wash solution was removed and replaced with 0.5 ml of acid solution (Section 2.1.11). After 30 seconds the acid solution was removed and 1ml of PBS complete (Section 2.1.11) added to the cells to neutralise the acid. It was important not to incubate the cells for longer than 60 seconds as they are adversely affected by low pH. Therefore, if a large number of samples were to be processed they were typically dealt with in batches of four or five at a time. When all cells had been treated, the PBS complete was replaced by 1ml of supplemented growth medium and the cells were incubated for 18-24 hours at 37°C.

#### **2.2.4.5. Immunofluorescence**

Following transfection (Section 2.2.4.3) and overnight incubation at 37°C, the growth medium was removed, the cells were washed twice in 1 ml PBS complete. They were then fixed and stored in 1 ml methanol at -20°C for at least 30 minutes. When required, the methanol was removed from the well and the cells were washed twice with 1 ml of PBS complete. The cells were rehydrated by adding 1 ml of antibody solution A (Section 2.1.11) to the well and incubating at room temperature for 10-15 minutes. During this incubation the primary antibody (Section 2.1.10) was diluted in antibody solution A and stored on ice until required.

Following rehydration, the coverslips were removed from the wells and drained. They were then placed upside-down in a 20  $\mu$ l drop of diluted antibody, which had previously been placed on a surface such as the lid of a 24 well plate (NUNC). To prevent the coverslips from drying out, they were incubated in a humidified chamber. After 45 minutes, the coverslips were drained, inverted, and washed in antibody solution A for five minutes. To ensure that all traces of primary antibody were removed, the wash step was repeated three times. The secondary antibody (Section 2.1.10) was diluted in antibody solution A and 20  $\mu$ l was added to the upside-down coverslip. The coverslips were incubated for a further 45 minutes, again using a humidified chamber, then drained and washed four times in antibody solution A (five minutes per wash) and once in H<sub>2</sub>O for ten seconds. They were then drained thoroughly and mounted onto microscope slides using Mowiol mounting medium (Harco) containing 2.5% DABCO (Sigma). The coverslips were then examined on a Zeiss LSM 510 confocal microscope.

#### **2.2.4.6. Green fluorescent protein fluorescence**

Following transfection (Section 2.2.4.3) and overnight incubation at 37°C, the growth medium was removed. The cells were then washed in 1ml PBS complete, fixed with formaldehyde fix (Section 2.1.11), and incubated at 4°C for at least 30 minutes. When required, the formaldehyde fix was removed and cells were washed twice in 1ml PBS complete and once in H<sub>2</sub>O. They were then drained thoroughly, mounted on microscope slides and examined as described for immunofluorescence (Section 2.2.4.5).

#### **2.2.4.7. Determination of virus titre**

Mammalian cells were set up in 40 mm plates at a density of  $1 \times 10^6$  cells in 2 ml of appropriate growth medium. They were incubated overnight at 37°C until they were approximately 80% confluent. The virus sample to be assayed was prepared by making a series of tenfold serial dilutions in PBS plus Newborn Calf Serum. For assaying high-titre virus stock, dilutions were made to the order of  $1 \times 10^{-9}$ , but for complementation studies the amount of virus in the samples was considerably lower. In the latter cases, dilutions to the order of  $1 \times 10^{-3}$  or  $1 \times 10^{-4}$  were appropriate.

The growth medium was removed from the cells and 100  $\mu$ l of diluted virus solution was added. The cells were incubated at 37°C with periodic shaking. After 1 hour the virus inoculum was removed and replaced with 3 ml of overlay medium containing 1.125% carboxymethylcellulose (Section 2.1.5). The cells were incubated for two days at 37°C

until virus plaques became visible. The overlay medium was then removed and replaced with 3 ml of Geimsa stain. After 2-3 hours the stain was removed. The cell sheets were then washed in copious amounts of water and allowed to dry. The plaques were counted using a dissection microscope and the virus titre was calculated.

### **2.2.5. Electron microscopy**

#### **2.2.5.1. Preparation of samples for electron microscopy**

To prepare cells for thin sectioning, the culture medium was removed and replaced with 0.5ml of complete PBS. The cells were harvested using a disposable cell scraper (Becton Dickinson), transferred to a BEEM capsule (TAAB), and pelleted by centrifugation at 1000 rpm for 5 minutes using a Beckman microfuge 12, which has a vertical rotor to ensure the formation of a compact and symmetrical cell pellet.

The supernatant was discarded and the cell pellet was fixed by overnight incubation at 4°C in 500µl of 2.5% glutaraldehyde in complete PBS. The fixed pellets were rinsed twice with complete PBS and incubated at room temperature for 1 hour in 1% osmium tetroxide (TAAB). The pellet was again rinsed twice in complete PBS, then dehydrated through a graded alcohol series comprising sequential incubations in 30, 50, 70, 90, and 100% ethanol. Following dehydration, the ethanol was carefully removed and replaced by newly prepared Epon 812 resin (TAAB). The capsules were left overnight with their lids open to allow residual ethanol to evaporate. The cell pellets were then drained and fresh resin was added. The samples were left overnight before being labelled and placed in an embedding oven at 65°C for two days in order to harden.

#### **2.2.5.2. Analysis of ultra-thin sections by transmission electron microscopy**

Embedded samples were sectioned using a Leica Ultracut E and mounted on to 400 mesh uncoated copper grids (Agar Aids). The cell sections were stained by immersing them in a saturated solution of uranyl acetate (BDH) in 90% ethanol for one hour. They were then washed by immersion (3 x 10 minutes) in water, before being stained with lead citrate for one minute, and given further washes (3 x 10 minutes) in water. After this they were dried and examined using a JEOL 100S electron microscope.

### **2.2.5.3. Analysis of purified proteins**

#### SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were resolved using a 10% polyacrylamide gel, prepared in 1x resolving gel buffer (see Materials), polymerised by the addition of 150  $\mu$ l of APS and 15  $\mu$ l of TEMED, and poured within a vertical mini gel plate system (Bio-Rad). Until the acrylamide became polymerised, it was overlaid with butanol, which produces a consistent upper boundary in the resolving gel. The butanol was removed following polymerisation, and the resolving gel was overlaid with a 5% polyacrylamide stacking gel. The stacking gel was prepared in 1x stacking gel buffer (see Materials) and polymerised by the addition of 75  $\mu$ l of APS and 5  $\mu$ l of TEMED.

Protein samples were prepared in SDS PAGE sample buffer (Materials) and heated at 100°C for approximately two minutes prior to loading. Gels were typically electrophoresed at 150 V until the dye had reached the bottom of the resolving gel. If protein standards were required, 3  $\mu$ l of Rainbow<sup>TM</sup> markers (Amersham) were analysed next to the protein samples for comparison.

Gels that were not going to be used for Western blotting were fixed and stained using Coomassie Brilliant Blue stain solution (Materials) for 10-20 minutes before being transferred into a destain solution (Materials) in order to remove excess background staining. Gels were photographed using a digital camera and images were processed using Photoshop 3.1 software (Adobe).

### **2.2.6. Semi-dry western blot analysis**

#### **2.2.6.1. Protein transfer**

The transfer of proteins to a nitro-cellulose membrane was carried out by semi-dry transfer. SDS-PAGE (Section 2.2.5.3) was carried out. The stacking gel was removed from the resolving gel which was soaked in Towbin buffer (Materials) for approximately ten minutes, as was a nitro-cellulose membrane (Hybond). The nitro-cellulose membrane was placed on a stack of three pieces of 3 mm Whatman paper, which had been cut to the appropriate size, and soaked briefly in Towbin buffer. The resolving gel was then placed on to the nitro-cellulose membrane, and an additional stack of three pieces of 3 mm Whatman paper (also soaked briefly in Towbin buffer) was placed on top of the resolving

gel. The transfer of proteins on to the nitro-cellulose membrane was carried out using a 2117 Multiphor II Electrophoresis Unit (Pharmacia), typically at 40 mA, for at least two hours.

### **2.2.6.2. Detection of proteins**

When the transfer was complete, non-specific binding sites on the nitro-cellulose membrane were blocked by immersing the membrane in 50 ml of 2.5% Marvel milk powder in PBS-T. Incubation was carried out at RT for a minimum of one hour on an orbital shaker. The blocked membrane was then washed (3 x 15 minutes) in approximately 50 ml of PBS-T. The membrane was subsequently probed at room temperature in 10 ml of PBS-T containing the appropriately diluted primary antibody, for a minimum of one hour on an orbital shaker. The membrane was then washed as described above, before the addition of the secondary antibody, protein-A peroxidase, which was diluted 1:1000 in PBS-T. The membrane was incubated at room temperature in approximately 10 ml of diluted secondary antibody for a minimum of one hour on an orbital shaker. The membrane was then washed again as described above. Immuno-detection was carried out using ECL (Amersham) in accordance with the manufacturer's instructions. Kodak XS-1 film was used for the visualisation of immuno-complexes.

### **2.2.7. Construction of recombinant baculoviruses**

L-broth agar plates containing kanamycin, gentamycin, tetracycline, Blue-gal and IPTG were prepared as described in Section 2.1.7. DH10Bac cells (Gibco) were transformed as previously described, with DNA containing pFASTBAC (Section 2.1.9) modified to contain the gene of interest. Plates were incubated overnight at 37°C. Colonies containing bacmids into which the modified pFASTBAC had inserted could be identified by their white colour (as opposed to the blue colour observed when the *lacZ* gene of the bacmid had not been interrupted by the insertion of pFASTBAC.) Four white colonies were picked and streaked out on L-broth agar plates identical to those described above, in order to confirm that the colonies selected were indeed white in colour. (Blue colour in colonies can take time to form, resulting in the possibility of false-positive samples). A single colony was picked from each of the plates used in this second round of growth, and inoculated into 2 ml of L-broth containing kanamycin, gentamycin and tetracycline. The colony was then grown overnight at 37°C on an orbital shaker. On the following day, an alkaline lysis miniprep (Section 2.2.2.6) was carried out on each of the samples.

### **2.2.7.1. Transfection of Sf21 cells with recombinant bacmid DNA**

Sf21 cells were set up in 40mm plates as previously described (Section 2.2.4.2). On the following day, transfections were set up in a similar fashion to that described for BHK-21 cells in Section 2.2.4.3, with the modification that the incubations were carried out at 28°C and the BHK-21 medium was replaced with TC100/5 medium. Virus was harvested from cell culture medium 72 hours after transfection. The supernatant was transferred to a sterile 1.5 ml reaction vial. Any cells that were carried over were pelleted by centrifugation at 13 000 rpm for two minutes in a microfuge. The supernatant was transferred to a fresh 1.5 ml cryo-vial, frozen on dry ice and stored at -70°C until required.

### **2.2.7.2. Production of low-titre baculovirus stocks**

Sf21 cells in TC100/5 media were seeded in small tissue culture flasks (200 ml) and grown at 28°C until the monolayer reached a confluency of 50%. The media was removed, then 500µl medium containing baculovirus, harvested as described in Section 2.2.7.1 was inoculated on to the cells and allowed to absorb for one hour at 28°C. 10ml of fresh TC100/5 medium was added to the cells, which were then incubated for 4-5 days at 28°C. The medium was then transferred to a fresh 15 ml Falcon tube, aliquotted, and stored at -70°C until required. Baculovirus DNA was made from 500 µl of this supernatant, and PCR was carried out to confirm the presence of the desired inserts. 5 ml of this supernatant was used to infect a 50% confluent monolayer of Sf21 cells in 150 cm<sup>2</sup> flasks (as described above). After absorption, 50 ml of TC100/5 medium was added to the cells. These were subsequently incubated for 4-5 days at 28°C, after which the supernatant was harvested and stored at -70°C. The flask harvest was then titrated as described in Section 2.2.7.4.

### **2.2.7.3. Production of high-titre baculovirus stocks**

High-titre stocks of recombinant baculoviruses were grown in roller bottles containing Sf21 cells at a density of 5x10<sup>5</sup> cells/ml (typically 300 ml per bottle). Cells were infected at a m.o.i. of 0.1 pfu/ml and cultured for five days at 28°C. Cultures were transferred to 250 ml Falcon tubes and the cells were pelleted by centrifugation at 3000 rpm for five minutes in a Sorvall RT7 centrifuge. The supernatant was transferred to a 500 ml Sorvall Swan-Neck centrifuge bottle and baculovirus was pelleted by centrifugation at 12 000 rpm for two hours at 4°C in a Sorvall SLA-3000 rotor. Virus pellets were resuspended in TC100/5 medium, transferred to sterile 1.5 ml cryo-vials and stored at -70°C until required.

#### **2.2.7.4. Baculovirus titration**

High-titre stocks of baculovirus were titrated using the method described by Brown and Faulkner (1997). Sf21 cells were seeded on 40mm plates at a density of  $1 \times 10^6$  cells per plate and grown overnight at 28°C to generate a 70-80% confluent monolayer. Serial tenfold dilutions of the virus were performed in TC100/5 medium. The medium was removed from the cells, then 100 µl of diluted virus, from  $10^{-5}$  to  $10^{-9}$ , was added to each 40mm plate and allowed to absorb for one hour at 28°C, with occasional rocking to prevent the cell monolayer from drying out. Following absorption, 1 ml of overlay was added. The overlay was a 1:1 mixture of TC100/5 and 3% LTG agarose (SeaPlaque), which had previously been melted and incubated at 45°C to prevent the overlay from setting. Once the overlay medium had set (approximately ten minutes), 1.5 ml of TC100/5 was added and the plates were incubated for a further 4-5 days at 28°C. The medium was removed and 1.5 ml of neutral red stain (0.4% neutral chloride mixed 1:50 with TC100/5) was added on top of the agar. The plates were then incubated for a further five hours at 28°C. The stain was removed and the plates were inverted on to tissue paper and allowed to air dry overnight at RT. Plaques were counted and the virus titre was calculated.

#### **2.2.8. FIAsh labelling**

BHK-21 cells in which FIAsh labelling was to be carried out were set up and transfected as previously described (Section 2.2.4.3) in 24 well plates. At three hours post-transfection, 1ml of growth medium (Section 2.1.5) was added to each well and the cells were incubated at 37°C overnight.

25 mM EDT was prepared by mixing 2.1 µl 1,2-ethanedithiol (EDT) with 1 ml of DMSO. This was freshly prepared each time the procedure was carried out. 1 µl of 25mM EDT was added to 2 µl of 1 µM FIAsh (Panvera), and incubated at room temperature for 5-10 minutes to ensure that all FIAsh was in FIAsh-EDT<sub>2</sub> form. The cells were washed three times with PBS complete supplemented with 1 g/litre D+ glucose.

1.88µl of FIAsh-EDT<sub>2</sub> (from the 5-10 minute incubation described in the paragraph above) was added to 2.5 ml of PBS/glucose. PBS/glucose was removed from the cells and replaced with 250 µl PBS/glucose/FIAsh/EDT. The cells were then incubated at 37°C for one hour. Washing solution was prepared by mixing 36 µl of British Anti-Lewisite in 1 ml

DMSO, and then transferring 50  $\mu$ l of this solution to 50 ml of PBS/glucose. Following incubation, the labelling solution was removed, and the cells were given two rinses and three ten minute washes in washing solution. The cells were then overlaid with 4% formaldehyde, incubated at 4°C for at least 30 minutes, and processed for viewing on a confocal microscope, as previously described.

If the cells were also to be probed with a VP5 antibody, this was carried out as previously described (Section 2.2.4.5).

### **2.2.9. Transient DNA packaging assay**

Packaging assays were carried out as described by Hodge and Stow (2001). BHK-21 cells were seeded on 40 mm plates at a density of  $1 \times 10^6$  cells per plate and grown overnight at 37°C, in order to generate a 70-80% confluent monolayer. The cells were transfected with DNA using the lipofectamine transfection method as described in Section 2.2.4.3. Three hours after transfection, the medium was removed and replaced with 800  $\mu$ l of supplemented BHK-21 growth medium. Five hours after transfection, cells were infected at a multiplicity of 5 pfu/cell with an appropriate virus.

Incubation was continued for 16 hours at 37°C. The medium was then removed, the cells were resuspended in 2.3 ml TBS and divided into two equal samples, which were used to prepare total cellular and DNase-resistant (encapsidated) DNA. The cells from both samples were pelleted and resuspended in 184  $\mu$ l of reticulocyte standard buffer (RSB) containing 0.5% NP-40. An equal volume of packaging assay cell lysis buffer (Materials) containing 1 mg/ml protease was added. In the case of the total cellular DNA, this was done immediately. In the case of the encapsidated DNA, the sample was first incubated in the presence of 200  $\mu$ g of DNase I/ml, with occasional mixing, for 20 minutes at 37°C. After the addition of protease, all the samples were incubated for one hour at 37°C, extracted sequentially with phenol and chloroform, and precipitated with ethanol. The nucleic acids were redissolved in H<sub>2</sub>O containing 10  $\mu$ g/ml of RNaseA.

Following digestion with *DpnI*, the digested DNA was fractionated by agarose gel electrophoresis and transferred to a Hybond-N membrane (Amersham). Replicated (*DpnI*-resistant) plasmid DNA was detected by hybridisation to a <sup>32</sup>P-labelled probe prepared from the vector pAT153. Phosphoimages of Southern blots were acquired using a Biorad Personal Molecular Imager FX.

## **CHAPTER 3**

# **USING INSERTIONAL MUTAGENESIS TO STUDY VP19C**

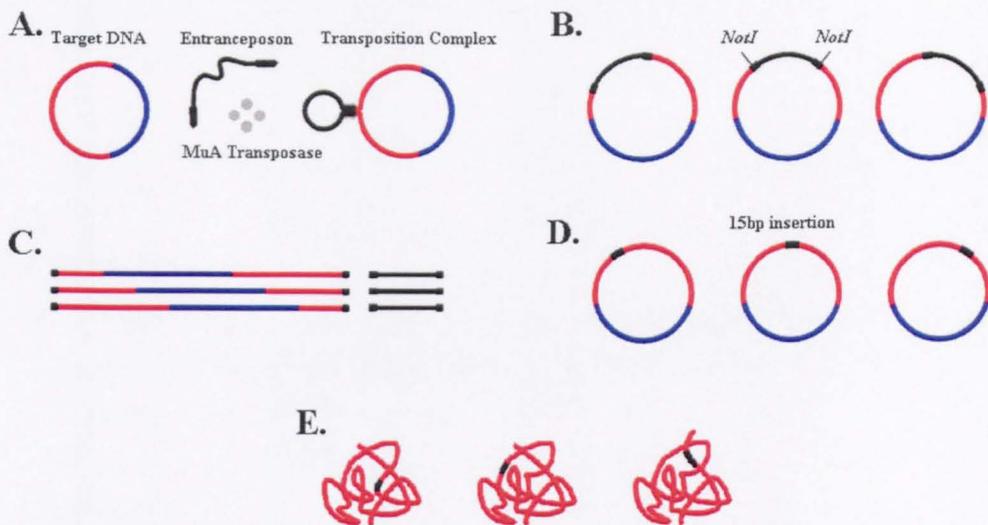
## 3. Using insertional mutagenesis to study VP19C

### 3.1. Introduction

Transposons are mobile genetic elements that were originally discovered in maize (McClintock, 1949). Since then they have become useful tools for genetic analysis of a variety of organisms (Berg and Berg, 1995; Kaiser *et al.*, 1995). Bacteriophage Mu, which uses DNA transposition to replicate its genome, is one of the best-characterised mobile genetic elements, and the first for which an *in vitro* transposition reaction was established (Mizuuchi, 1983). Subsequently, the bacteriophage Mu transposition reaction has been modified for use as a tool in genetic and molecular analysis (Happa *et al.*, 1999). In principle, any DNA sequence inserted between the ends of the Mu DNA can be used as a transposon.

The Mutation Generation System (Finnzymes) was based on the transposition machinery of Bacteriophage Mu (Figure 3.1). Using this system, it is possible to construct a library of 15 bp insertional mutations for the gene of interest, and this library can subsequently be used to carry out a functional analysis of the protein encoded by the gene. One transposition reaction can generate thousands of mutants with, in theory, random coverage throughout the gene. The insert sequences produced are short (five amino acids). Due to the sequence of the nucleotides inserted it is not possible for a stop codon to be introduced (as described in Figure 3.2), and as 15 nucleotides are introduced, the reading frame does not change and the amino acids downstream of the insertion site are not altered. Furthermore, the nucleotides inserted include a *NotI* restriction enzyme recognition sequence (GCGGCCGC). This can be used to estimate the position of the insertion within the gene of interest. Thereafter, the precise position can be identified by sequencing the gene.

In all the work described in this thesis, the transposon inserted during the Mutation Generation System transposition reaction (designated as the Entranceposon) is 1139 bp in length and contains a gene that codes for kanamycin resistance. This gene is flanked by 50 bp inverted terminal repeat sequences, which act as a recognition sequence for the transposase enzyme. The distal end of each terminal repeat sequence is straddled by a *NotI* restriction enzyme recognition sequence, therefore cleavage with *NotI* removes the intervening transposon sequences (Figure 3.1). Based on this arrangement for the Entranceposon, there are certain prerequisites for any target DNA that the transposition



**Figure 3.1:** The Mutation Generation System

The Mutation Generation System is designed for the rapid construction of libraries of insertional mutants. The system employs the transposition machinery of the bacteriophage Mu to generate a pool of 15 bp insertional mutants. **A.** The transposition reaction is catalysed by MuA transposase. **B.** The entranceposon is randomly inserted into the target DNA. **C.** The plasmid DNA is digested with *NotI*, and the central portion of the entranceposon is removed. **D.** *NotI* digested plasmids are then self-ligated, resulting in a 15 bp insertion in the target plasmid. The location of the insertion can be estimated by carrying out a series of restriction digests, and confirmed by DNA sequencing. **E.** If the 15 bp insertion is in an ORF, it is translated into five extra amino acids in the encoded protein.

Figure adapted from manufacturers' Internet site (<http://www.finnzymes.fi/>).

Frame 1:

5'	N12	345	TGC	GGC	CGC	A12	345	NNN	3'
	X	X	Cys	Gly	Arg	Ile	X	X	
						Met			
						Thr			
						Asn			
						Lys			
						Ser			
						Arg			

Frame 2:

5'	NN1	234	5TG	CGG	CCG	CA1	234	5NN	3'
	X	X	Leu	Arg	Pro	His	X	X	
			Met			Gln			
			Val						

Frame 3:

5'	NNN	123	45T	GCG	GCC	GCA	123	45N	3'
	X	X	X*	Ala	Ala	Ala	X	X	

X = Any amino acid

X\* = Any amino acid except Gln, Glu, Lys, Met, Trp

Figure 3.2: Translation of the 15 bp insertions

The amino acid patterns possible in mutant proteins following insertional mutagenesis.

reaction is to be performed on: namely that the target DNA should not itself contain any *NotI* restriction sites, and that it should not have a gene for kanamycin resistance.

The aim of the work described in this chapter was to use the mutation generation system to examine the triplex protein VP19C and identify regions of the protein that are important for its correct function.

### 3.2. Construction of insertional mutants

A previously constructed plasmid, pBJ382 (Tatman *et al.*, 1994), contains the UL38 open reading frame within a plasmid construct that also contains a gene encoding for ampicillin resistance. Furthermore, pBJ382 does not contain any *NotI* restriction sites, thus making it a suitable target for the Mutation Generation System. Following mutagenesis, mutated UL38 open reading frames can be transferred to a plasmid containing suitable promoter and enhancer sequences, and then analysed. A further advantage of using pBJ382 as the target DNA for mutagenesis is its small size – 4169 bp. Since the UL38 ORF is 1395 bp, approximately 33% of the mutant plasmids obtained by means of the transposition reaction should have an insertion within the UL38 ORF.

For the investigation described in this thesis, a transposition reaction was carried out in accordance with the manufacturers' instructions. A mass of approximately 160 ng of target DNA was used in the reaction. After one hour, and following inactivation of the transposase, 1 µl of the reaction mixture was diluted in 9 µl of H<sub>2</sub>O. The dilution was necessary due to the high salt concentrations in the transposition reaction mixture, which can result in arcing during electroporation. Thereafter, 5 µl of the resulting mixture was used to transform DH5α cells (Section 2.2.3.3). The resulting bacteria were spread over four LB/agar + amp + kan plates and incubated at 37°C overnight.

Each plate produced approximately 15-20 colonies. Each colony was picked and grown overnight at 37°C in 200 µl of LB + amp + kan. On the following day the resulting cultures were pooled, and a vacuum miniprep (Qiagen) used to obtain DNA.

In pBJ382, the UL38 ORF is flanked by *XbaI* and *HincII* restriction sites. Digestion of pBJ382 with these enzymes would normally produce fragments of 1491 bp (UL38 ORF) and 2678 bp (plasmid vector). However, as shown in Figure 3.3, the pooled pBJ382 DNAs– each with a full-length transposon inserted into it (pBJ382<sup>m</sup>) – will produce bands

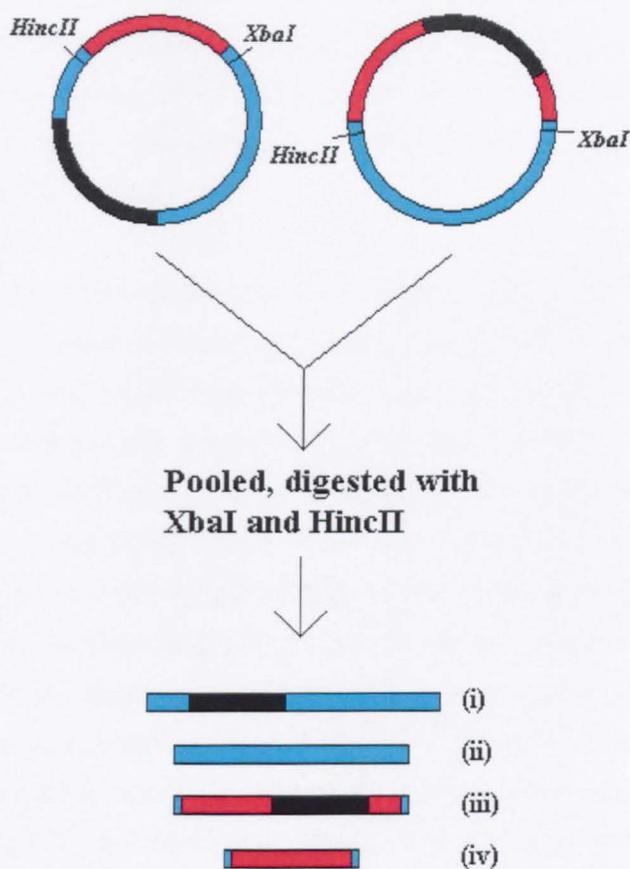


Figure 3.3: Digestion of pooled DNA with *HincII* and *XbaI* following insertional mutagenesis.

Digestion of pooled DNA should yield four bands on an agarose gel. (i) A 3817 bp fragment containing the transposon (black) inserted into the backbone of the plasmid (turquoise). (ii) A 2678 bp fragment containing the backbone of the plasmid alone. (iii) A 2630 bp fragment containing the transposon inserted into the UL38 ORF (red). (iv) A 1491 bp fragment containing the UL38 ORF alone. The desired products of the *HincII/XbaI* digest are type (iii) bands.

of 1491 bp, 2630 bp, 2678 bp, and 3817 bp. The 2630 bp band represents an insertion of the transposon sequence into the UL38 ORF. Therefore, the pooled DNA was digested with *XbaI* and *HincII*, run on 1xTAE gel, and the 2630 bp band extracted. On the gel it was not possible to distinguish the 2630 bp band from the 2678 bp band, but products from the 2678 bp band can be eliminated at a later stage of the experimental procedure, as described in the following paragraph.

To allow expression of the insertional mutants, the pool of mutated UL38 ORFs was transferred to a plasmid containing suitable promoter and enhancer elements. pFBpCI was made by inserting a *BglIII-EcoRI* fragment containing the HCMV promoter/enhancer from pCI-neo (Promega) into *BamHI/EcoRI* digested pFASTBAC Hta (Invitrogen) (pFBpCI supplied by Roger Everett). pFBpCI contains one *NotI* restriction site, which would prove a hindrance during digestion with *NotI* to remove the central portion of the transposon (as described below). The *NotI* site in pFBpCI was removed by digestion of the DNA with *NotI*, making the ends flush with T4 polymerase, and religation to produce pFBpCI<sup>NotI</sup>-. The *XbaI/HincII* digested fragments containing the mutated UL38 ORF as described above were ligated into *NheI/StuI* digested pFBpCI<sup>NotI</sup>- (*HincII* is compatible with *StuI*, while *NheI* is compatible with *XbaI*). The resulting mixture was used to transform DH5 $\alpha$  cells (Section 2.2.3.3). The transformed bacteria were spread over four LB/agar + amp + kan plates and incubated at 37°C overnight. Bacteria transformed with plasmids containing the 2678 bp DNA described in the previous paragraph would be unable to grow on these plates as they do not possess the kanamycin resistance gene contained within the transposon. Colonies obtained from these plates were pooled and a vacuum miniprep (Qiagen) used to obtain DNA. This DNA comprised a pool of plasmids that would allow the expression of VP19C mutants under the control of the HCMV immediate-early promoter. The DNA pool was digested with *NotI* and religated, thus removing the bulk of the transposon sequence and leaving a set of 15 bp insertions, which introduced five additional amino acids to the protein.

Following ligation of the *NotI* digested DNA pool, it was transformed into DH5 $\alpha$  cells (Section 2.2.3.3) and the resulting bacteria were spread over four LB/agar + amp plates and incubated at 37°C overnight. The four plates produced a total for 48 colonies. Each colony was picked and grown overnight at 37°C in 2.5ml of LB + amp. On the following day, a vacuum miniprep was used to obtain DNA from each culture. As the vector plasmid does not contain any *NotI* restriction sites, the position of the insertion in each of the 48 miniprep DNAs could be estimated by a series of restriction digests. The plasmid contains

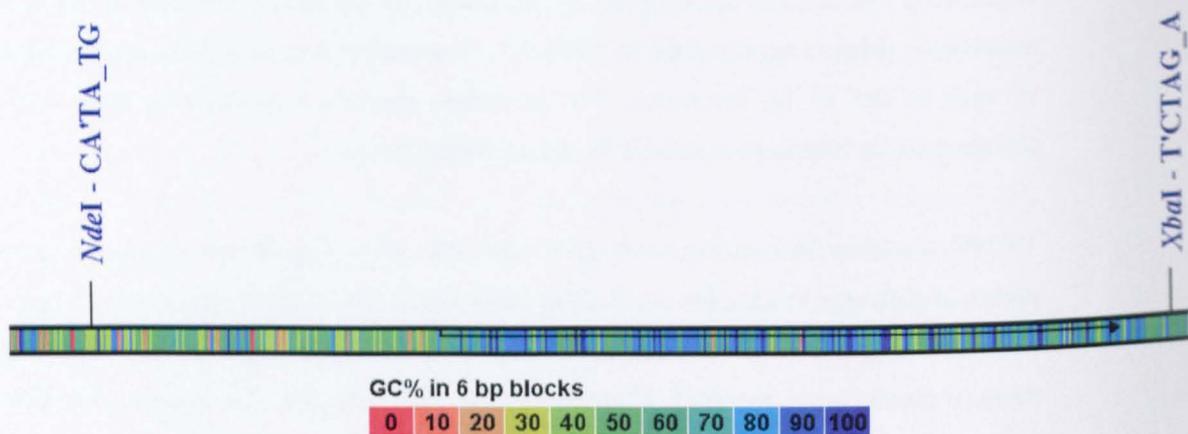
a single *NdeI* restriction site 722 bp upstream of the start of the UL38 coding sequence, and a single *XbaI* restriction site 110 bp downstream of the UL38 stop codon. Hence, *NdeI/NotI* and *NotI/XbaI* digests of each DNA allow an estimation to be made of the location of the *NotI* restriction site, as shown in Figure 3.4. Following this estimation, sequencing was carried out on each of the DNAs, as described in Section 2.2.2.9, using sequencing primers as described in Table 3.1. Sequencing determined the precise locations of each of the 15 bp insertions. The 48 DNAs examined produced a total of eleven different 15 bp insertions in the UL38 open reading frame.

VP19C contains 465 amino acids and therefore eleven insertional mutants equates to approximately one mutant for every 42 amino acids. It was decided that in order to produce a satisfactory analysis of the effects produced in VP19C by the five amino acid insertions, more mutants were required. Furthermore, it was felt that the number of different insertional mutants obtained from the first attempt had been reduced by the inclusion of pooling steps in the experimental procedure. Thus, 5  $\mu$ l of the original transposition reaction was transformed into DH5 $\alpha$  cells, the resulting bacteria were spread over four LB/agar + amp + kan plates and incubated at 37°C overnight.

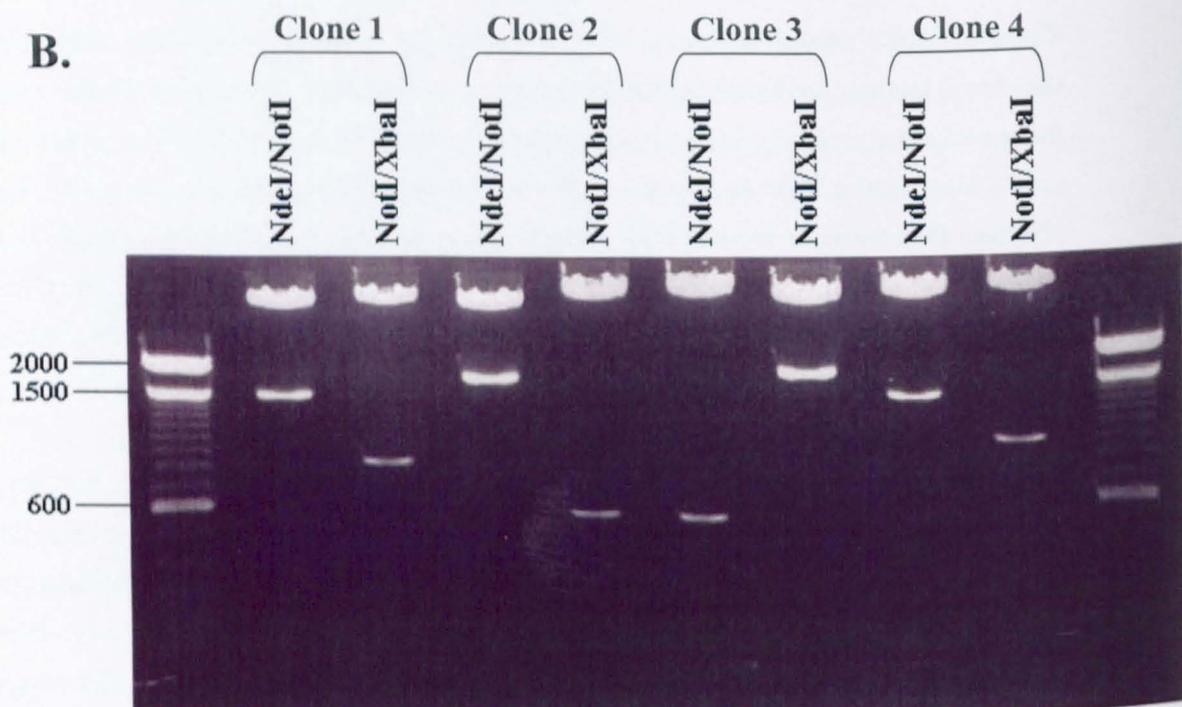
The four plates produced a total of 65 colonies. To ensure that the final number of insertional mutants produced did not decrease as a result of DNA pooling, each colony was treated as a distinct entity for the remainder of the procedure. Each colony was picked and grown overnight at 37°C in 200  $\mu$ l of LB + amp + kan. On the following day, DNA was obtained from each of the resulting cultures using the vacuum miniprep method. Each DNA was examined to identify whether it contained an insertion into or near the UL38 ORF (by carrying out an *XbaI/HincII* digest as described previously). Of the 65 DNAs examined, 23 contained an insertion into or near the UL38 ORF (approximately 33%).

Each of the 23 insertional mutants was cloned into the modified pFBpCI<sup>NotI</sup>- as described for the pooled clones above. Digestion with *NotI* and subsequent religation was carried out to remove the central portion of the transposon. The approximate location of the insertion was estimated and the sequence of each DNA was obtained as described previously. From these 23 mutants, 19 new, unique insertions into the UL38 ORF were identified (there were two cases where the transposon had inserted at the same position in two separate mutants, and two cases where the transposon had inserted between the *XbaI* and *HincII* restriction sites but outwith the UL38 ORF).

A.



B.



**Figure 3.4:** The positions of insertions into the UL38 ORF were estimated by carrying out *NdeI/NotI* and *NotI/XbaI* digests on each mutant.

**A.** In pUL38FbpCI, a unique *NdeI* site lies 722 bp upstream of the start of the UL38 ORF, while a unique *XbaI* site lies 110 bp downstream of the stop codon. Hence, *NdeI/NotI* and *NotI/XbaI* digests of each DNA allow an estimation to be made of the location of the inserted *NotI* restriction site. **B.** Gel photograph showing results of *NdeI/NotI* and *NotI/XbaI* digests. Band sizes on a 100 bp DNA ladder (New England Biolabs) are shown. Clone 1 has a *NotI* restriction site approximately 800 bp from the *XbaI* site (approximately 700 bp from the beginning of the UL38 ORF). Clone 2 has a *NotI* restriction site approximately 600 bp from the *XbaI* site (approximately 900 bp from the beginning of the UL38 ORF). Clone 3 has a *NotI* restriction site approximately 600 bp from the *NdeI* site (upstream of the start of the UL38 ORF). Clone 4 has a *NotI* restriction site approximately 900 bp from the *XbaI* site (approximately 600 bp from the start of the UL38 ORF). These estimations were used to select suitable sequencing primers for identifying the precise location of the insertion.

<b>Primer ID</b>	<b>Primer Sequence</b>	<b>Position within the UL38 ORF</b>
UL38 SEQ 2	5' TCGTTTAACTTCCTGGTGGC 3'	553-572
UL38 SEQ 3	5' ATGTGCGTTCGTGGACCTGG 3'	867-886
UL38 SEQ 4	5' AGTTCGCCGACAGGCTGTAC 3'	1184-1203
UL38 SEQ 5	5' CACGATCCGCGTTGGGTTGG 3'	376-357
UL38 SEQ 6	5' CACACGCCGCCACCAGGAAG 3'	580-561
UL38 SEQ 7	5' AGGTCCACGAACGCACATGC 3'	884-865

**Table 3.1:** Sequencing primers used in DNA sequence analysis of the UL38 insertional mutants.

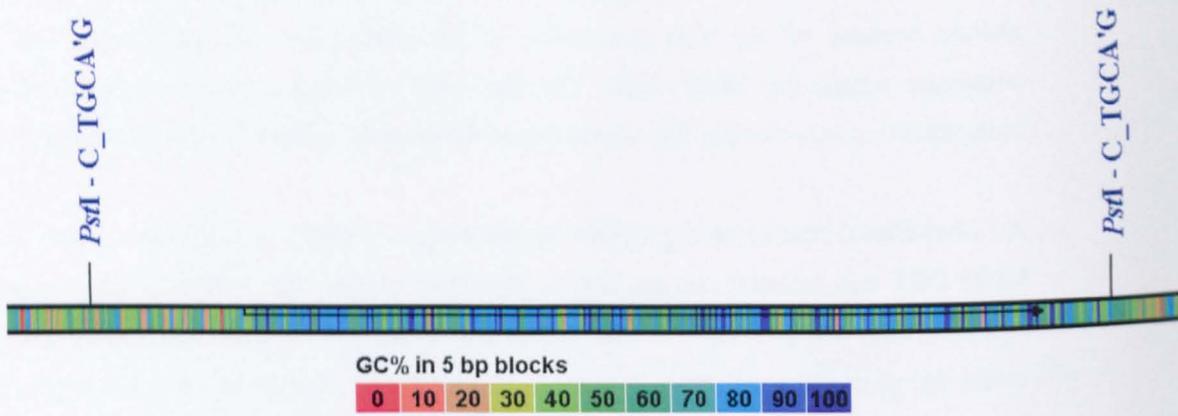
The first two cycles of the Mutation Generation System experimental procedure had produced a total of 30 unique 15 bp insertions into the UL38 ORF (one mutant for every 15.5 amino acids). In order to obtain more insertional mutants, the procedure was carried out for a third time. The starting plasmid for the previous two cycles (pBJ382) had been chosen because of the high proportion of the nucleotides of this plasmid that were contained within the UL38 ORF. The drawback of using pBJ382 was that following mutagenesis, a sub-cloning step was required to insert the mutated UL38 into pFBpCI.

An *XbaI/HincII* restriction digest was carried out on pBJ382. The fragment containing the UL38 ORF was isolated and purified as described above. This fragment was ligated into *NheI/StuI* digested pFBpCI<sup>NotI</sup>. The resulting plasmid (pUL38FbpCI) contains a gene that codes for ampicillin resistance and no *NotI* restriction sites. It is 7420 bp in length, of which 1395 bp comprises the UL38 ORF. Hence, approximately 19% of insertions should lie within the UL38 ORF.

A transposition reaction was carried out in accordance with the manufacturers' instructions and as described previously. A mass of approximately 300 ng of target DNA was used in the reaction. 1  $\mu$ l of the reaction mixture was diluted in 9  $\mu$ l of H<sub>2</sub>O and 5  $\mu$ l of the resulting mixture was used to transform DH5 $\alpha$  cells. The resulting bacteria were spread over four LB/agar + amp + kan plates and incubated at 37°C overnight. The four plates produced a total of approximately 138 colonies. Each colony was picked and grown overnight at 37°C in 200  $\mu$ l of LB + amp + kan. Vacuum minipreps were used to obtain plasmid DNA from each culture the following day.

In pUL38FbpCI, *PstI* restriction sites lie 251 bp upstream and 122 bp downstream of the UL38 ORF. Hence, digestion with *PstI* will produce a fragment of 2929 bp if the transposon has inserted within or near to the UL38 ORF, and 1790 bp if the insertion has occurred elsewhere in the plasmid. A typical diagnostic gel showing the results of *PstI* digests is shown in Figure 3.5. Out of the 138 DNAs, 25 were found to contain a plasmid that had inserted into or near the UL38 ORF. Digestion by *NotI* and subsequent religation was carried out on each of the 25 DNAs. The approximate location of each insertion was estimated and the sequence of each DNA obtained as described previously. Out of the 25 DNAs sequenced, two had insertions in identical locations to previously isolated mutants, and three had insertions that were outwith the UL38 ORF. The remaining 20 DNAs were unique insertions into the UL38 ORF.

A.



B.

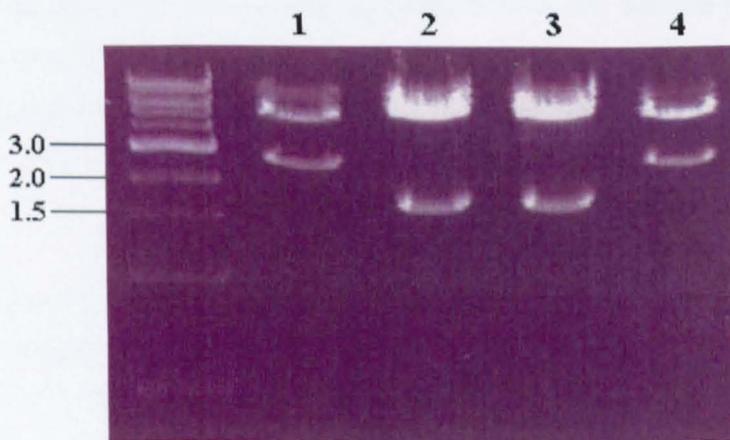


Figure 3.5: *Pst*I digests were used to identify clones with insertions lying in or near the UL38 ORF.

**A.** pUL38FbpCI contains two *Pst*I sites, lying 251 bp upstream and 122 bp downstream of the UL38 ORF. Thus, clones containing insertions lying in or near the UL38 ORF could be identified by digestion with *Pst*I. **B.** Gel photograph showing results of *Pst*I digests. Band sizes of a 1 kb DNA ladder (New England Biolabs) are shown. Digestion with *Pst*I will produce a fragment of 2929 bp if the transposon has inserted within or near to the UL38 ORF, and 1790 bp if the insertion has occurred elsewhere in the plasmid. Therefore, in clones 1 and 4 the transposon has inserted within or near the UL38 ORF. In clones 2 and 3 it has inserted elsewhere in the plasmid.

The three cycles of mutagenesis produced 50 clonally distinct mutants of the UL38 ORF, as summarised in Table 3.2. The insertion (in) number given in the name of each insertional mutant refers to the first amino acid in sequence that is altered by the introduction of extra amino acids, for example in in16, the first 15 amino acids of the sequence are the same as wild-type VP19C, but the 16<sup>th</sup> is a cysteine rather than a serine. The insertions were distributed throughout the ORF with an average spacing of approximately 9 amino acids in VP19C, and with a maximum separation of 38. Two pairs of insertions (in242A, in242B and in358A, in358B) were at different sites in the same codon. In these cases the mutant proteins differ only in the identity of the inserted amino acids (Table 3.2). One insertion (inSTOP) disrupted the TGA stop codon, resulting in the addition of five amino acids at the end of the protein.

### 3.3. Complementation of insertional mutants

To analyse the effects of the mutations on the properties of VP19C, their ability to complement the growth of the HSV-1 deletional mutant  $\nu\Delta 38$ YFP (Section 2.1.6; Thurlow *et al.*, 2005) was examined. 0.5  $\mu$ g of DNA from each insertional mutant and a wild-type UL19 control was transfected into BHK-21 cells, which were incubated at 37°C for five hours before being infected with 2 pfu/cell of  $\nu\Delta 38$ YFP. After one hour, the unabsorbed input virus was neutralised by an acid wash. Incubation was continued at 37°C for 40 hours, at which time the progeny virus was harvested and titrated on complementing UL38RSC cells, which express VP19C (Section 2.1.4). Figure 3.6 shows the relative yield obtained for each of the insertional mutants compared to that obtained with the wild-type control.

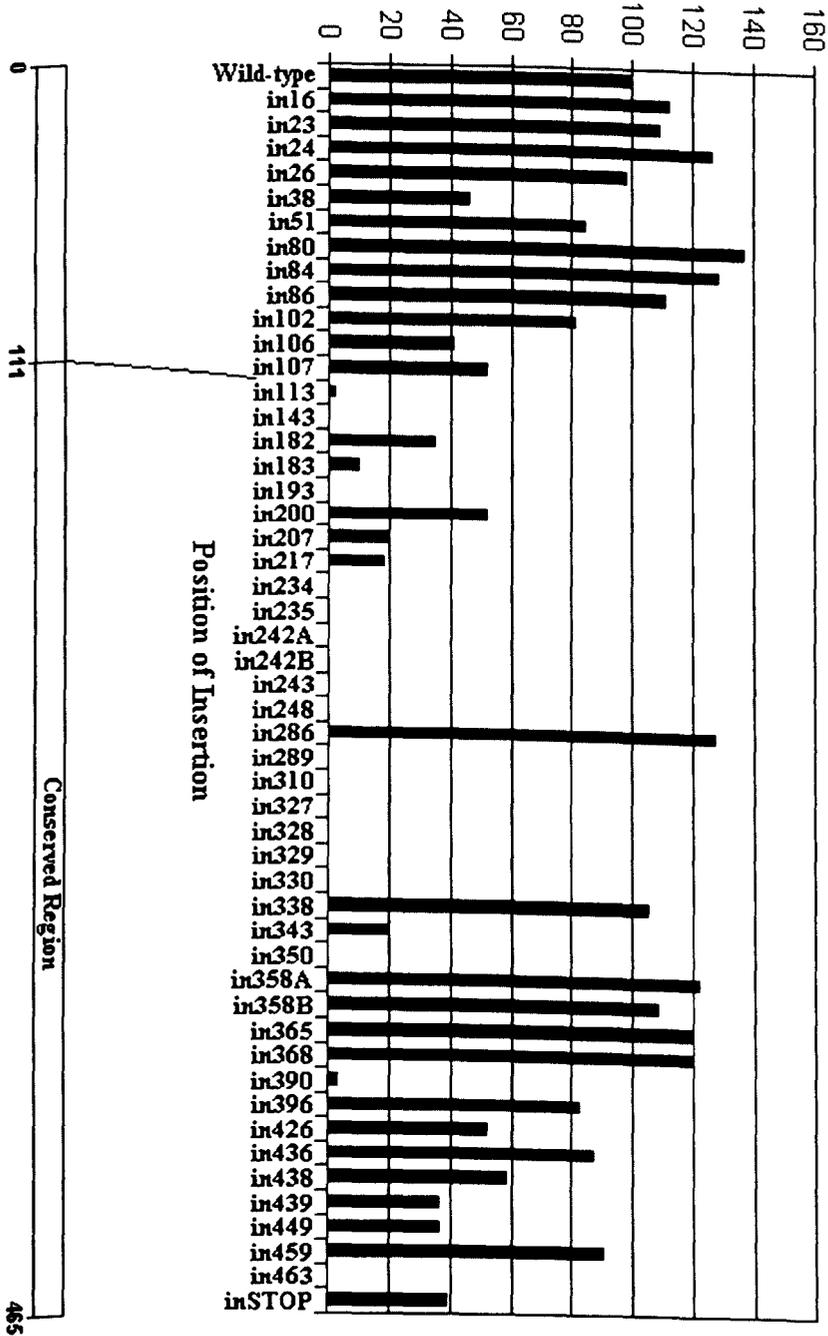
Eighteen of the 50 insertional mutants had little or no effect on the ability of VP19C to support virus growth, producing yields of over 80% of the wild-type VP19C control level. The remaining 32 showed lower levels of complementation, with two producing yields of less than 10%, and 16 being completely negative for virus growth. The results shown in Figure 3.6 suggest that VP19C (465 amino acids) can be divided into three broadly defined regions, with the majority of the severely disabling insertions (less than 10% of wild-type yield) occurring in the central region of the protein, between in113 and in350. The yield obtained with all twelve of the insertions into the first 107 amino acids was greater than 40% of the wild-type level – indeed, only in38 in the middle of this region – and in106 and in107 at its C-terminus gave yields of less than 80%. Similarly, all but two of the 14 insertions into the C-terminal 107 amino acids (in390 and in463) gave complementation

AA no.	AA sequence								
16	ggCGRSGst	106	spAAASPrh	234	amCGRTMvh	328	vyAAAVYvi	390	agCGRTGqv
23	ppMRPHPt	107	prAAAPRhh	235	mvAAAMVht	329	yvMRPHVik	396	pgCGRTQfa
24	ptAAAPTtr	113	rqCGRSQvi	242A	phAAAPHev	330	viNAAAIks	426	lgAAALGmm
26	trCGRTRdt	143	hlVRPHLah	242B	phCGRThev	338	rgCGRSGle	436	clLRPHLhr
37	vlLRPHLrp	182	raVRPHAg	243	heCGRNEvm	343	alVRPHLer	438	hrMRPHRte
51	lpSAAAPrg	183	agAAAAGlv	248	fgCGRIGgl	350	lrCGRIRit	439	rtDAAATer
80	tpCGRTPnd	193	aaALVAACA	286	avAAAAl	358A	hgCGRNGte	449	pvAAAPVvi
84	dtDAAATeq	200	daLRPHARd	289	lpVRPHPac	358B	hgAAAHGte	459	rpGAAAPge
86	eqVRPQqal	207	avAAAAl	310	flLRPHLyl	365	ppVRPHPap	463	wrVRPQRac
102	alMRPHLig	217	rgCGRSGtr	327	cvCGRSVyv	368	pnCGRTNrn	STOP	caCAAAA*

**Table 3.2:** Locations and sequences of five amino acid insertions into VP19C.

The number denotes the last unchanged amino acid before the insertion. The five inserted amino acids are shown in upper case and the unchanged flanking residues in lower case.

Virus Titre (% of Wild-type)



**Figure 3.6:** Functional analysis of VP19C insertional mutants. (Top) Complementation of growth of the VP19C null mutant v $\Delta$ 38YFP by transfected plasmids expressing the VP19C insertional mutants was carried out in BHK-21 cells. The progeny virus was titrated on UL38RSC cells. The results are shown as percentages of the titre obtained with the wild-type VP19C control plasmid (pUL38FBpCI). (Bottom) Scale drawing of the UL38 ORF showing the extent of the poorly-conserved N-terminal sequences (Figure 5.1A).

levels greater than 30%. In contrast, 16 of the 18 insertions giving complementation values of less than 10% mapped to the central region of the protein, between amino acids 113 and 350. Out of the 24 insertions in this region, only four gave values of greater than 30%.

### **3.4. Interaction of insertional mutants with VP23 and VP5**

In order to ensure the formation of functional capsids, and hence infectious virions, VP19C must interact correctly with the other triplex protein, VP23 and with the major capsid protein, VP5. Since these interactions are essential, they can be assumed to have formed correctly in the VP19C mutants that complemented  $\nu\Delta 38$ YFP. For this reason, only the eighteen mutants that gave less than 10% of the yield of wild-type VP19C and were considered to be severely disabled were examined.

Previous studies have shown that VP19C is a nuclear protein. In the absence of other virus proteins, VP23 and VP5 are uniformly distributed throughout the cell. However, they become predominantly nuclear in the presence of VP19C (Rixon *et al.*, 1996). To examine these interactions, BHK-21 cells were co-transfected with plasmids expressing the VP19C mutants, along with either pE18, which expresses VP23 (Nicholson *et al.*, 1994), or pE19, which expresses VP5 (Rixon *et al.*, 1996). The distributions of the proteins were determined by immunofluorescence (Section 2.2.4.5). Only one of the mutant VP19Cs (in330) was altered in its distribution and was no longer able to locate to the nucleus, making it impossible to determine its effect on the distribution of the other proteins (Figures 3.7 and 3.8).

The interaction between VP23 and the remaining 17 severely disabled VP19C mutants was examined by this assay. In a typical slide, the levels of fluorescence observed (and hence presumably the amounts of VP19C and VP23 transfected) varied greatly between cells. In general, the presence of VP19C and VP23 in the nuclei was more easily seen in cells exhibiting lower levels of fluorescence. Since there was variation in the patterns of fluorescence, each coverslip that demonstrated convincing examples of clear-cut nuclear localisation of the proteins under study was scored as positive for the nuclear location of that protein. In those that are scored as negative, no convincing examples of nuclear localisation were observed. Two distinct patterns were seen (Figure 3.7). Nine mutants (in143, in234, in235, in327, in328, in329, in350, in390, and in463) exhibited a wild-type pattern, with VP19C and VP23 co-localising to the nucleus. In the case of in390, the

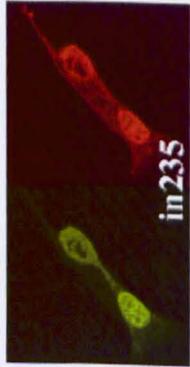
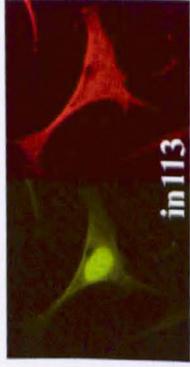
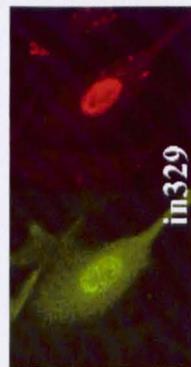
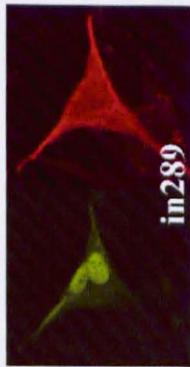
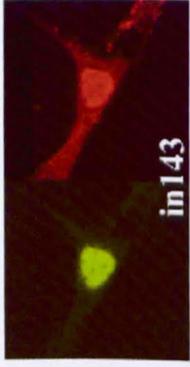
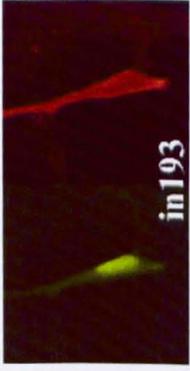
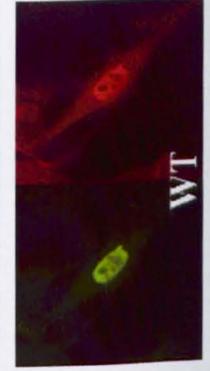
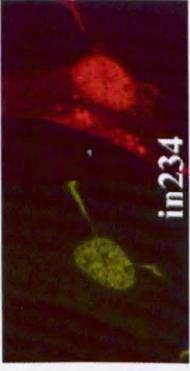
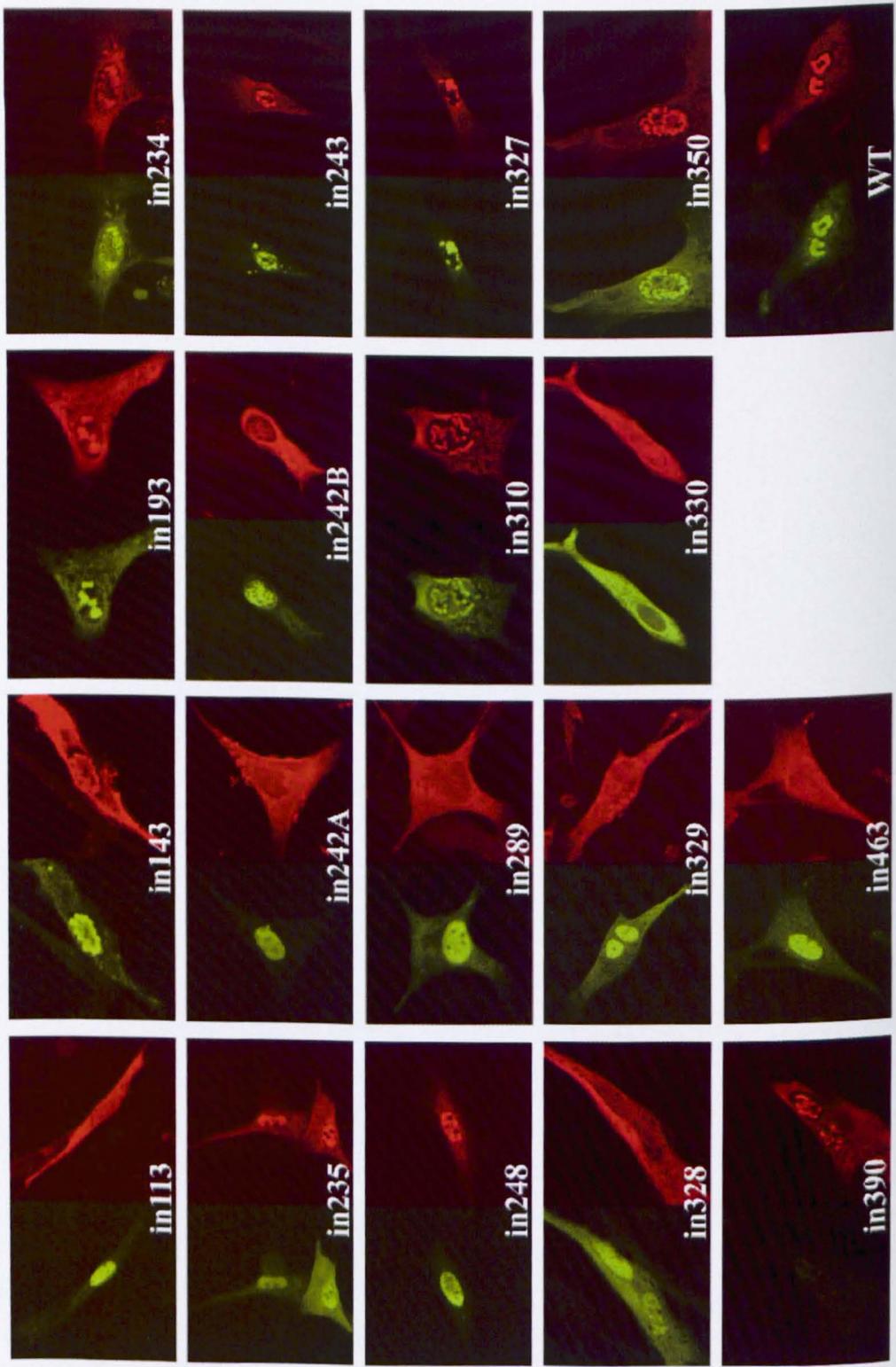


Figure 3.7: Influence of VP19C insertional mutants on the distribution of VP23.

BHK-21 cells were cotransfected with pE18 (expressing VP23) and either pUL38FBpCI (wild-type (WT) VP19C), or one of the VP19C insertional mutants as indicated. VP19C was detected with the monoclonal antibody mAb02040 and visualised using FITC GAM (green). VP23 was detected with the antiserum rAb186, and visualised using TRITC-conjugated GAR (red).



**Figure 3.8:** Influence of VP19C insertional mutants on the distribution of VP5.

BHK-21 cells were cotransfected with pE19 (expressing VP5) and either pUL38FBpC1 (wild-type (WT) VP19C), or one of the VP19C insertional mutants as indicated. VP19C was detected with the monoclonal antibody mAb02040, and visualised using FITC GAM (green). VP5 was detected with the antiserum rAb184, and visualised using TRITC-conjugated GAR (red).

VP19C antibody was unable to bind to the mutant protein, but VP23 was clearly transported to the nucleus, implying that the insertion had disrupted the antibody-binding site but had not interfered with the ability of VP19C to interact with VP23 or to localise to the nucleus. This would suggest that the binding site for the VP19C antibody is near amino acid 390. In the remaining eight mutants (in113, in193, in242A, in242B, in243, in248, in289, and in310), VP19C localised to the nucleus as expected but VP23 remained in the cytoplasm, implying that the inserted sequences had disrupted the interaction between VP19C and VP23.

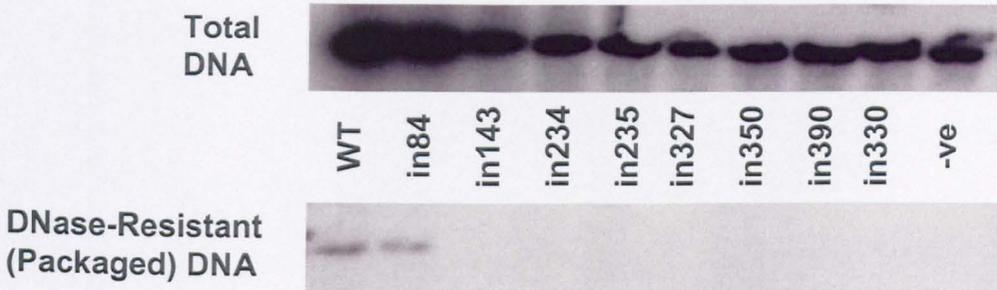
When the equivalent analysis was carried out for the interaction between VP19C and VP5, seven of the severely disabled VP19C mutants (in113, in242A, in242B, in289, in328, in329, and in463) were unable to transport VP5 to the nucleus (Figure 3.8). Out of these, four (in113, in242A, in242B and in289) had also failed to transport VP23 to the nucleus, while the other three (in328, in329 and in463) interacted successfully with VP23. The interaction between VP19C and VP5 was less clear-cut than the interaction between VP19C and VP23. Cells in which VP5 had been transported to the nucleus by VP19C were more difficult to identify. Even when assessing the interaction between wild-type VP19C and VP5, a significant fraction of the VP5 was frequently observed in the cytoplasm. Nevertheless, the same criteria were applied as for VP19C/VP23 co-localisation, and all coverslips scored as positive contained clear-cut examples of nuclear localisation of VP5. The interactions of VP19C mutants with VP23 and VP5 are summarised in Table 3.3.

### **3.5. Insertional mutants in DNA packaging**

Mutants in143, in234, in235, in327, in350, and in390 were shown to be able to transport both VP23 and VP5 to the nucleus, however they were unable to support virus growth. Since failure to package the viral genome correctly would be lethal, each of these mutants was analysed further by a plasmid-based DNA packaging assay. As described in Section 2.2.9, the presence of DNase-resistant DNA in the WT and in84 (a non-lethal insertional mutant) transfected samples (Figure 3.9) demonstrates that DNA packaging had taken place, while in each of the lethal insertional mutants tested the lack of packaged DNA suggests either that capsid assembly was disrupted or that the packaging mechanism was affected directly.

VP19C mutant	Localises to the nucleus	Takes VP23 to the nucleus	Takes VP5 to the nucleus
in113	+	-	-
in143	+	+	+
in193	+	-	+
in234	+	+	+
in235	+	+	+
in242A	+	-	-
in242B	+	-	-
in243	+	-	+
in248	+	-	+
in289	+	-	-
in310	+	-	+
in327	+	+	+
in328	+	+	-
in329	+	+	-
in330	-	-	-
in350	+	+	+
in390	+	+	+
in463	+	+	-

Table 3.3: Intracellular distribution of the severely disabled VP19C insertional mutants.



**Figure 3.9:** Ability of VP19C insertional mutants to package DNA.

BHK-21 cells were transfected with either pUL38FBpCI (wild-type VP19C), one of the VP19C insertional mutants, or a negative (no DNA) control as indicated. Cells were superinfected with  $\nu\Delta 38$ YFP, and total (top) and Dnase-resistant (bottom) DNAs were prepared. The DNAs were digested with *EcoRI* and *DpnI*, and the fragments were separated by electrophoresis through an agarose gel. The gel was blotted, the membrane was hybridised to  $^{32}\text{P}$ -labeled pAT153 DNA and, after being washed, the membrane was exposed to a phosphoimager screen.

## **CHAPTER 4**

# **MUTATIONAL ANALYSIS OF THE VP19C N-TERMINAL REGION**

## 4. Mutational analysis of the VP19C N-terminal region

### 4.1. The VP19C N-terminal

Studies examining HSV-1 capsid formation using a recombinant baculovirus-based model system had shown that the N-terminal 90 amino acids of VP19C were not absolutely required for capsid assembly (Spencer *et al.*, 1998). Because these sequences were considered dispensable for capsid assembly, it was decided to try and map the location of the VP19C N-terminus by difference mapping between capsid reconstructions made using N-terminally truncated and wild-type VP19C. Spencer *et al.* (1998) infected insect cells with recombinant baculoviruses expressing VP5, preVP22a, and VP23 along with one of the mutant VP19Cs. The three VP19C mutants that they analysed had the first 45, the first 90, and the first 105 amino acids removed. They observed capsids in cells infected with each of the two shorter deletions, although the efficiency of capsid formation was lower for the 90 amino acid deletion. No intact capsids were seen with the 105 amino acid deletion. Therefore, in order to map the N-terminus of VP19C, recombinant baculoviruses expressing two deletional mutants of VP19C were constructed – a deletion of the first 45 amino acids (which was expected to produce capsids efficiently) and a deletion of the first 63 amino acids (which was hoped to aid difference mapping by removing a larger section of the N-terminal region without reducing efficiency of capsid formation).

### 4.2. Cloning strategy for N-terminal deletion mutants

The nomenclature used in relation to N-terminal deletion mutants is as follows: in constructs with names containing ‘-45’ or ‘-63’, the first 45 or 63 amino acids respectively of VP19C have been removed. In constructs with names containing ‘NLS’, the SV40 large T antigen nuclear localisation signal (Kalderon *et al.*, 1984) has been attached to the N-terminus of VP19C. In constructs suffixed with an asterisk (such as pUL38-63FBpCI\*), the UL38 ORF contains a mutation at position 493 that was subsequently ‘repaired’ (Section 4.2.2). The cloning strategy for the construction of VP19C N-terminal deletion mutants is summarised in Figure 4.1. To make the recombinant baculoviruses, truncated UL38 ORFs were generated by PCR (Figure 4.1A) using the forward primers UL38-45F and UL38-63F, and the reverse primer UL38R (Table 4.1). pE38 (Nicholson *et al.*, 1994) was used as template DNA (Figure 4.2A). The PCR products were purified and cloned into

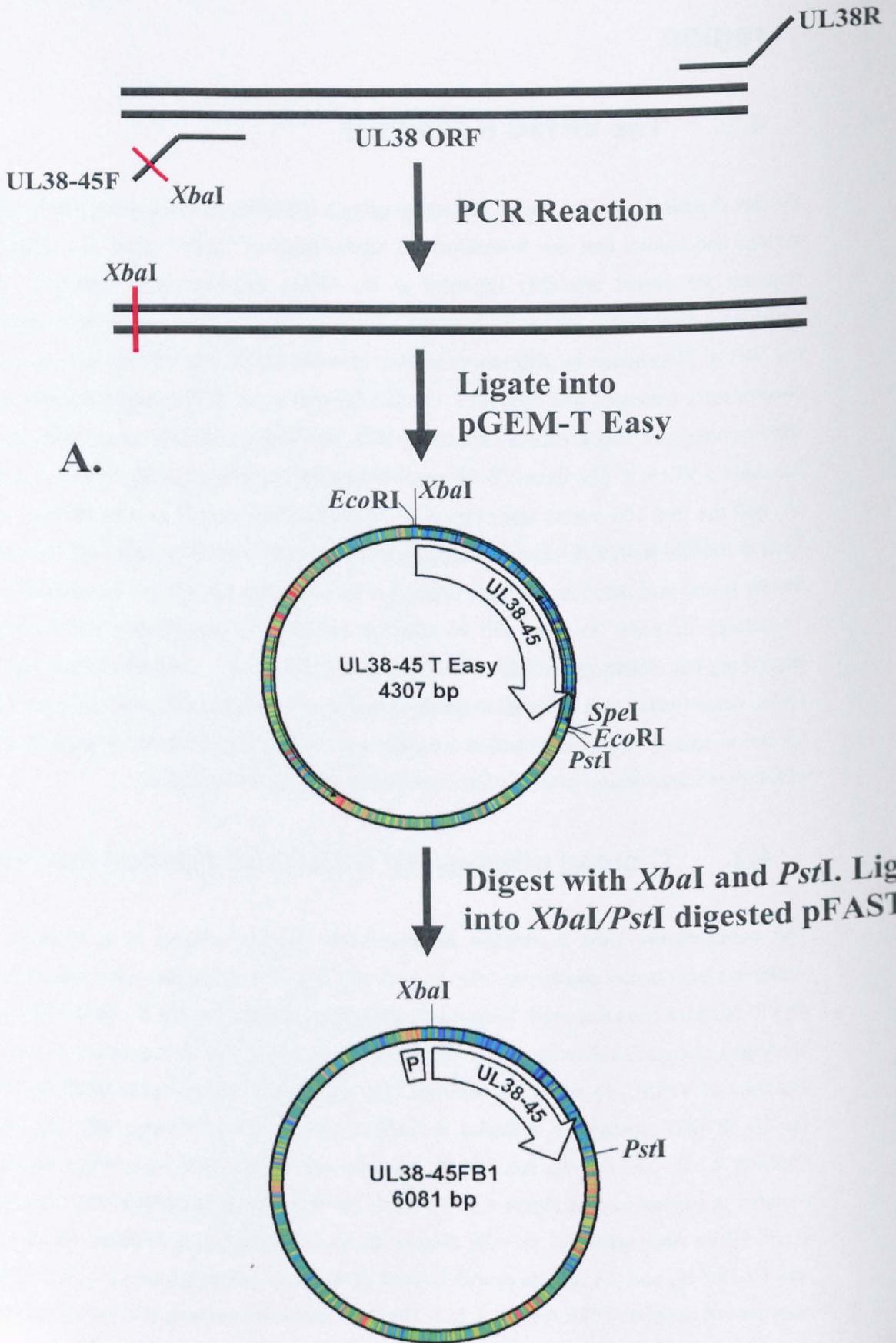
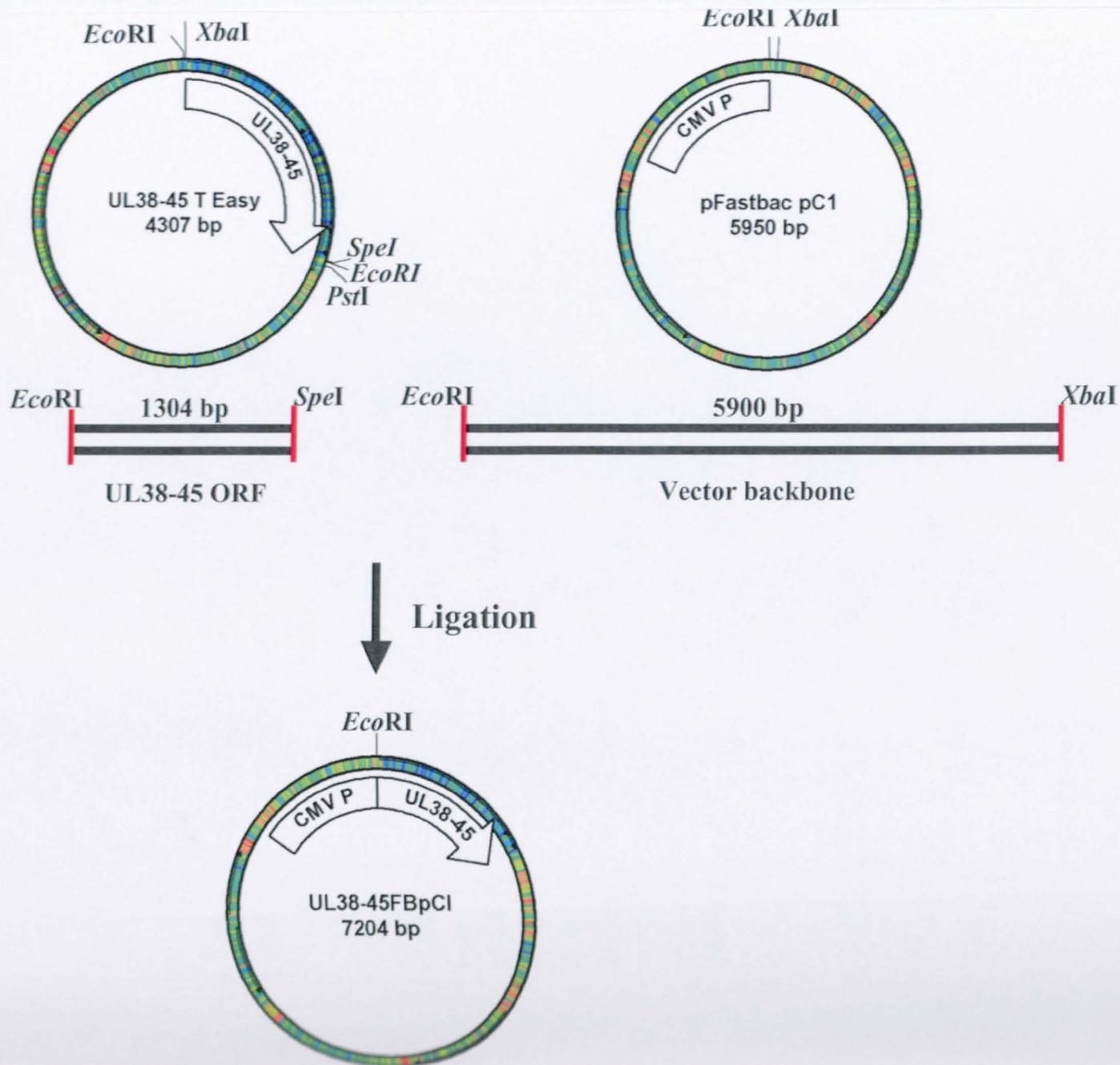


Figure 4.1: Cloning strategy for VP19C N-terminal deletion mutants.

A. Cloning strategy for producing pUL38-45FB1. A PCR reaction was carried out using pE38 as template DNA, UL38-45F as a forward primer, and UL38R as a reverse primer (Table 4.1). The PCR product was purified and cloned into pGEM-T Easy. The UL38 ORF was isolated by digestion with *Xba*I (which was introduced by UL38-45F) and *Pst*I which is contained on pGEM-T Easy, downstream of the point at which the PCR product inserts. This fragment was ligated into *Xba*I/*Pst*I digested pFASTBAC1 to generate pUL38-45FB1. The location of the polyhedrin promoter (P) is shown. A similar process was carried out to produce pUL38-63FB1\*, with the exception that the UL38-45F primer was replaced with UL38-63F.

B.



**Figure 4.1:** Cloning strategy for VP19C N-terminal deletion mutants.

**B.** Cloning strategy for producing pUL38-45FBpCI. An *EcoRI/SpeI* digest was carried out on pUL38-45T Easy and a 1304 bp fragment was isolated. An *EcoRI/XbaI* digest was carried out on pFASTBAC pCI and a 5900 bp fragment was isolated. These fragments were ligated together to produce pUL38-45FBpCI. A similar strategy was used to produce pUL38-63FBpCI\*, except that the *EcoRI/SpeI* fragment was purified from pUL38-63T Easy\*.

C.

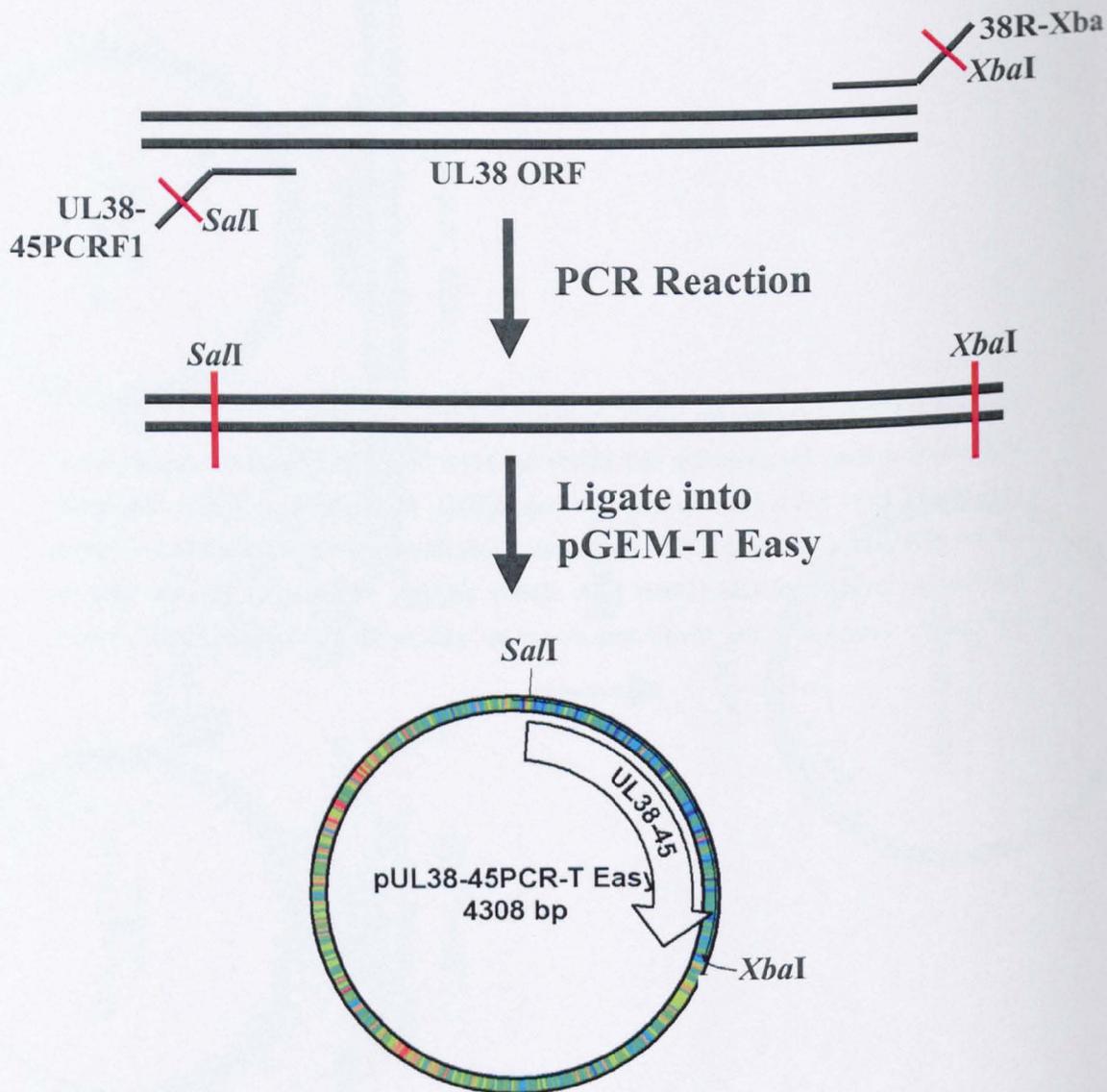
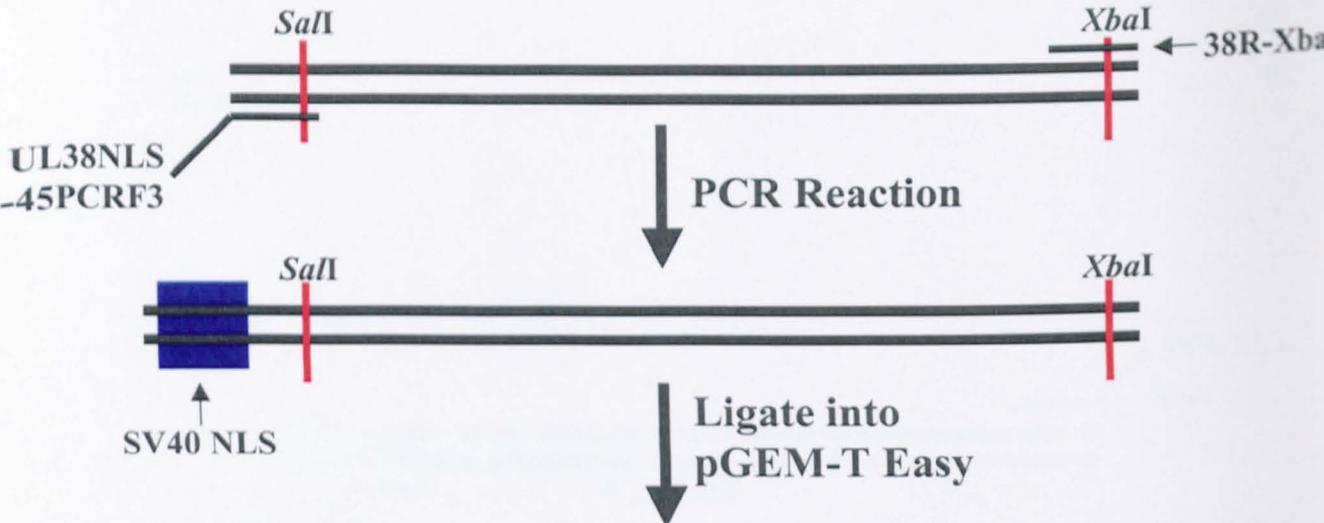
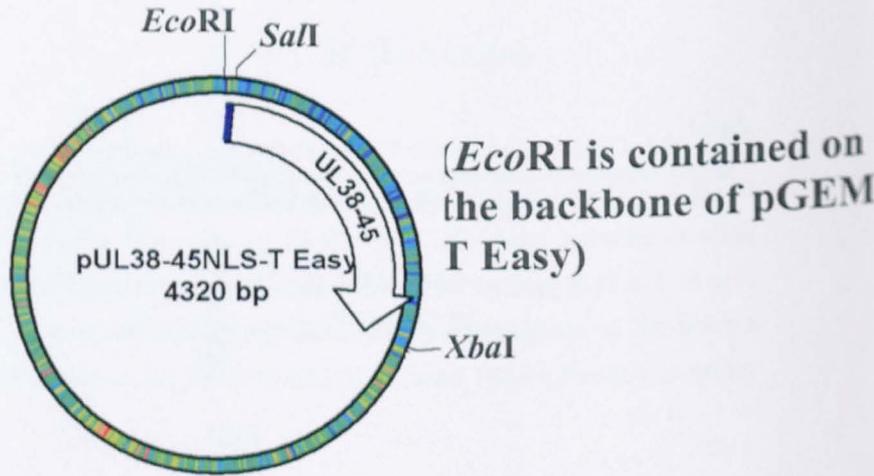


Figure 4.1: Cloning strategy for VP19C N-terminal deletion mutants.

C. Cloning strategy to produce pUL38-45PCR-T Easy. A PCR reaction was carried out using pE38 as template DNA, UL38-45PCR F1 as a forward primer, and 38R-Xba as a reverse primer. The PCR product was purified and cloned into pGEM-T Easy. A similar process was carried out to produce pUL38-63PCR-T Easy\*, with the exceptions being that pE38 was replaced with pUL38-45T Easy, and UL38-45PCR F1 was replaced with UL38-63PCR F1.



D.



Digest with *XbaI* and *EcoRI*.  
Ligate into *XbaI/EcoRI* digested pFASTBACpCI

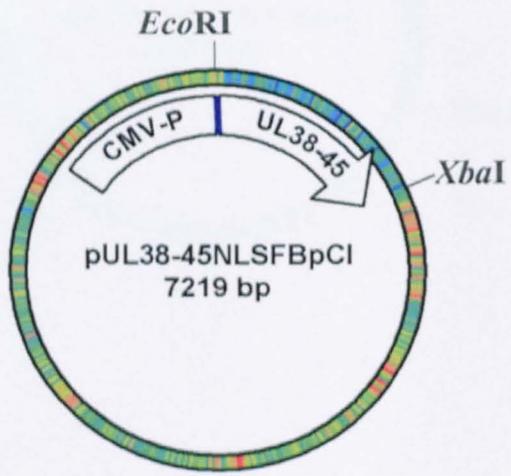
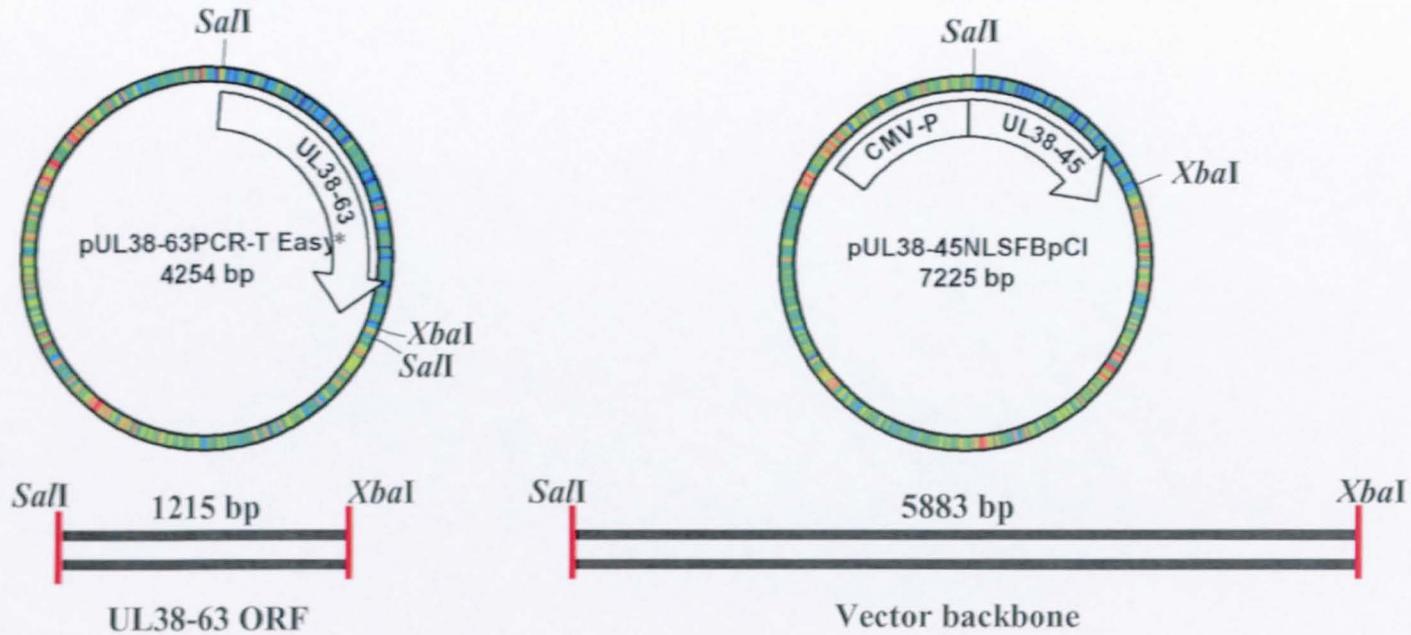


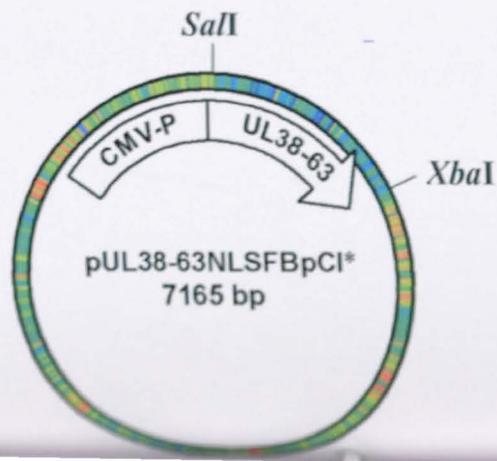
Figure 4.1: Cloning strategy for VP19C N-terminal deletion mutants.

**D.** Cloning strategy to produce pUL38-45NLSFBpCI. A PCR reaction was carried out using pUL38-45PCR-T Easy as template DNA, UL38NLS-45PCR F3 as a forward primer, and 38R-Xba as a reverse primer. UL38NLS-45PCR contains DNA coding for the SV40 NLS. The PCR product was purified and cloned into pGEM-T Easy. The location of the NLS in the plasmid is shown in blue. The UL38 ORF was isolated by digestion with *EcoRI* (which is contained on pGEM-T Easy, upstream of the point at which the PCR product inserts) and *XbaI* (which was introduced by 38R-Xba). This fragment was ligated into *EcoRI/XbaI* digested pFBpCI to generate UL38-45NLSFBpCI. A similar process was carried out to produce UL38-63NLSFBpCI\*, with the exceptions being that pUL38-45PCR-T Easy was replaced with pUL38-63PCR-T Easy\*, and that following the production of UL38-45NLSFBpCI, the plasmid was digested with *Sall* and *XbaI*, the fragment containing the UL38 ORF was removed and replaced with the UL38 ORF obtained from a *Sall/XbaI* digest of pUL38-63PCR-T Easy\*.

E.



Ligation



**Figure 4.1:** Cloning strategy for VP19C N-terminal deletion mutants.

**E.** Cloning strategy to 'repair' pUL38-63FBpCI\*. The mutation lies downstream of a *PvuII* restriction site. An *EcoRI/PvuII* digest was carried out on pUL38-63FBpCI\* and a 324 bp fragment was isolated. An *EcoRI/PvuII* digest was carried out on pUL38-45FBpCI and a 6836 bp fragment was isolated. These fragments were ligated together to produce pUL38-63FBpCI. A similar strategy was used to 'repair' pUL38-63NLSFBpCI. pUL38-63NLSFBpCI\* was digested with *PvuII* and *KpnI* and a band of 6198 bp was isolated. pUL38-45NLSFBpCI was also digested with *PvuII* and *KpnI* and a band of 967 bp was isolated. The two purified DNAs were ligated together to produce pUL38-63NLSFBpCI.

F.

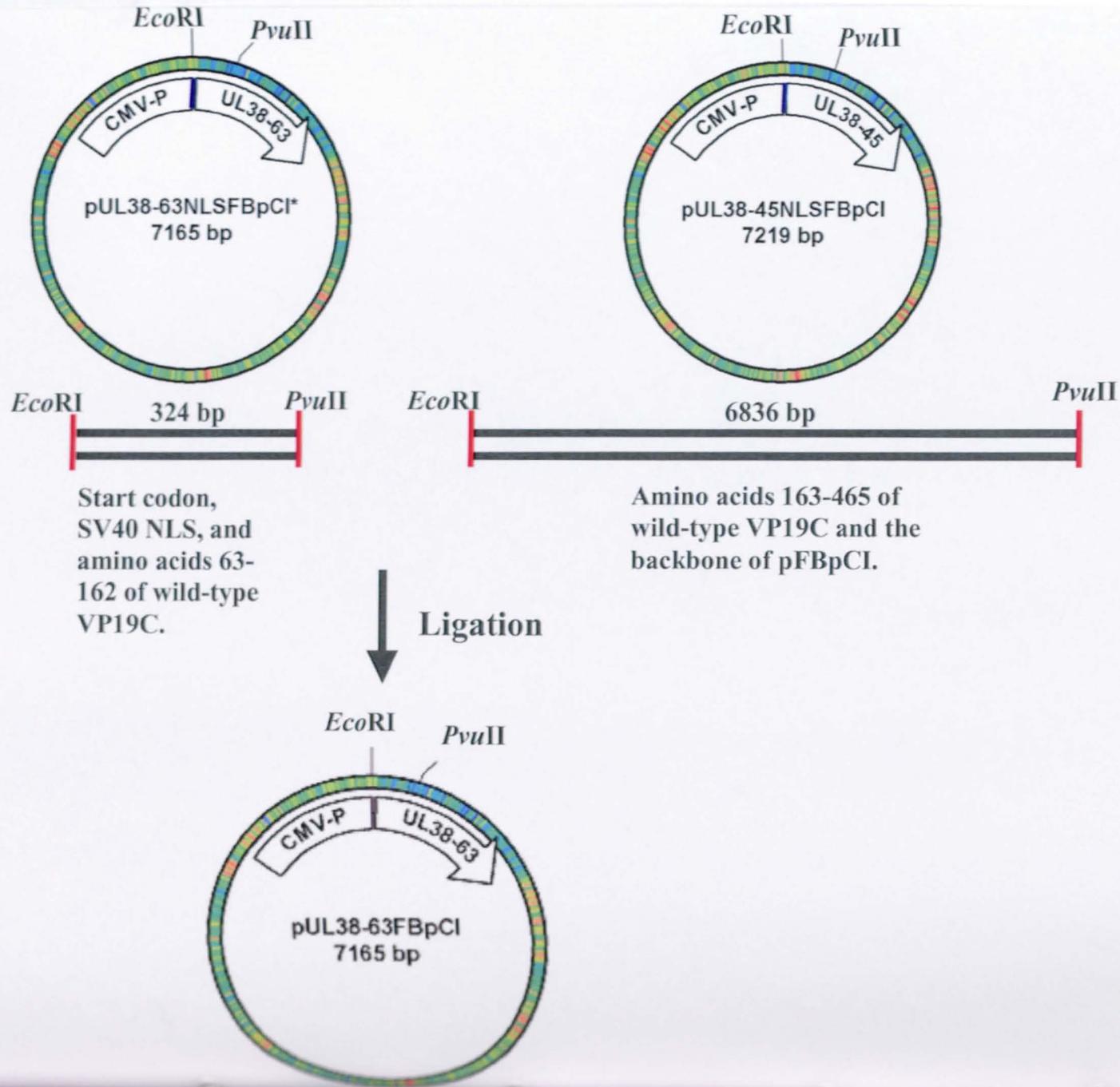
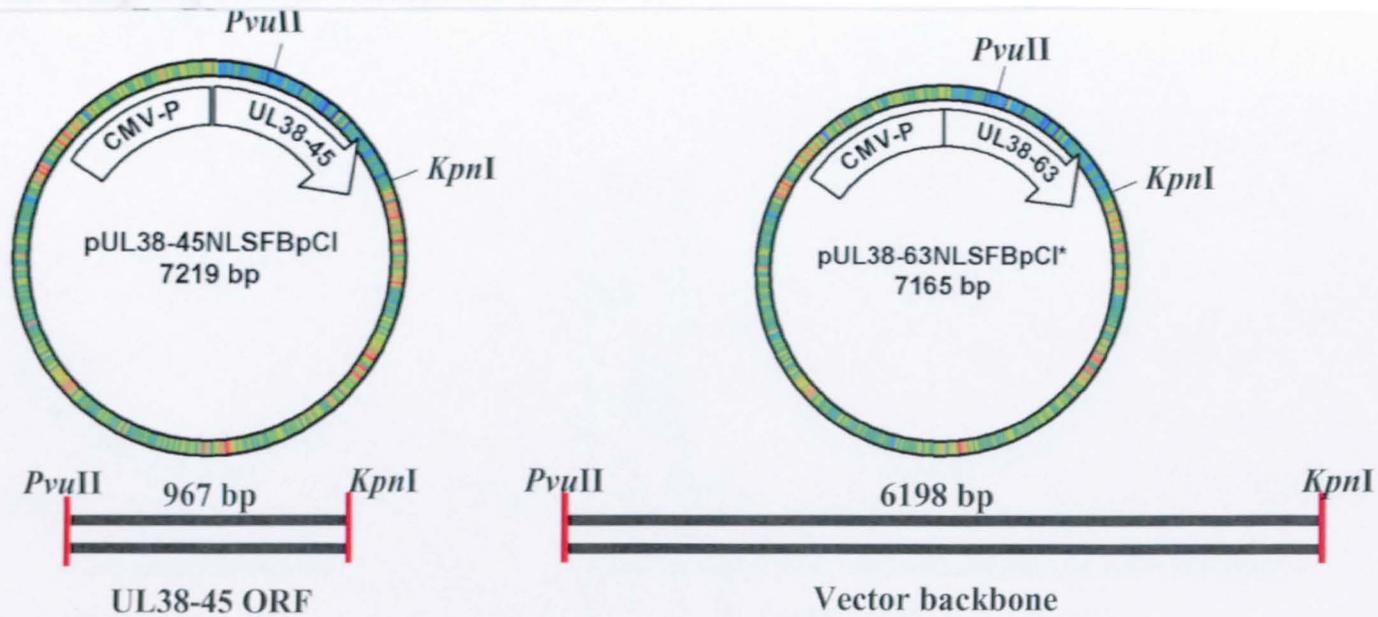


Figure 4.1: Cloning strategy for VP19C N-terminal deletion mutants.

**F.** Cloning strategy for producing pUL38-63NLSFBpCI\*. A *Sall/XbaI* digest was carried out on pUL38-63PCR-T Easy and a 1215 bp fragment was isolated. A *Sall/XbaI* digest was carried out on pUL38-45NLSFBpCI and a 5883 bp fragment was isolated. These fragments were ligated together to produce pUL38-63NLSFBpCI\*.

G.



Ligation

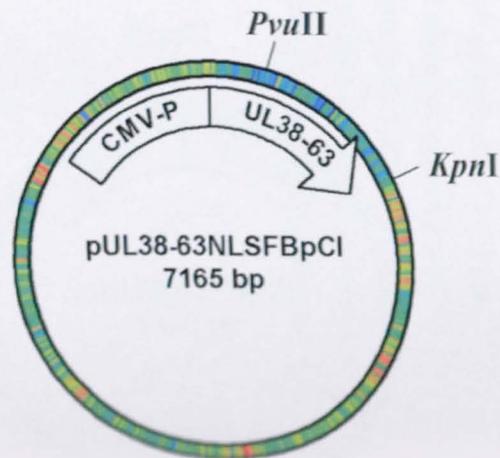
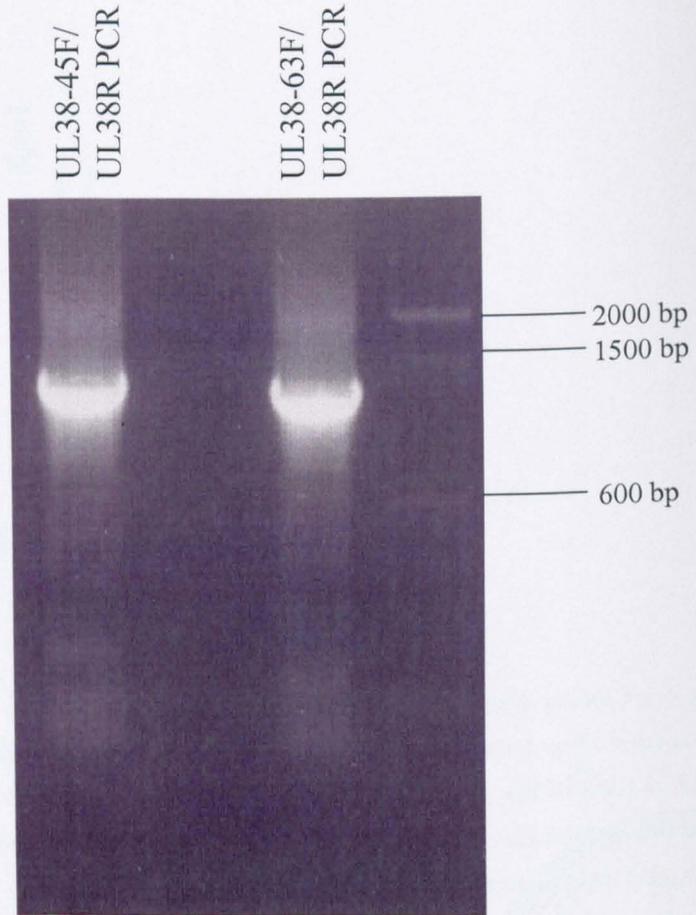


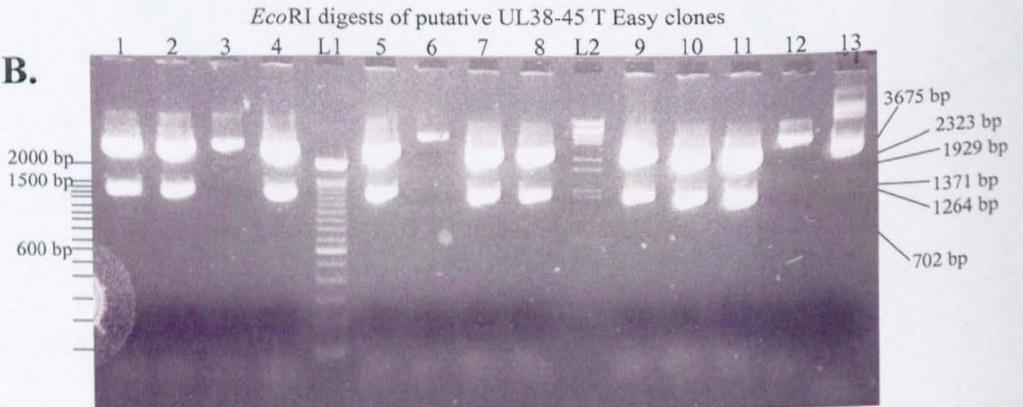
Figure 4.1: Cloning strategy for VP19C N-terminal deletion mutants.

**G.** Cloning strategy for producing pUL38-63NLSFBpCI. A *PvuII/KpnI* digest was carried out on pUL38-63NLSFBpCI\* and a 6198 bp fragment was isolated. A *PvuII/KpnI* digest was carried out on pUL38-45NLSFBpCI and a 967 bp fragment was isolated. These fragments were ligated together to produce pUL38-63NLSFBpCI.

**A.**



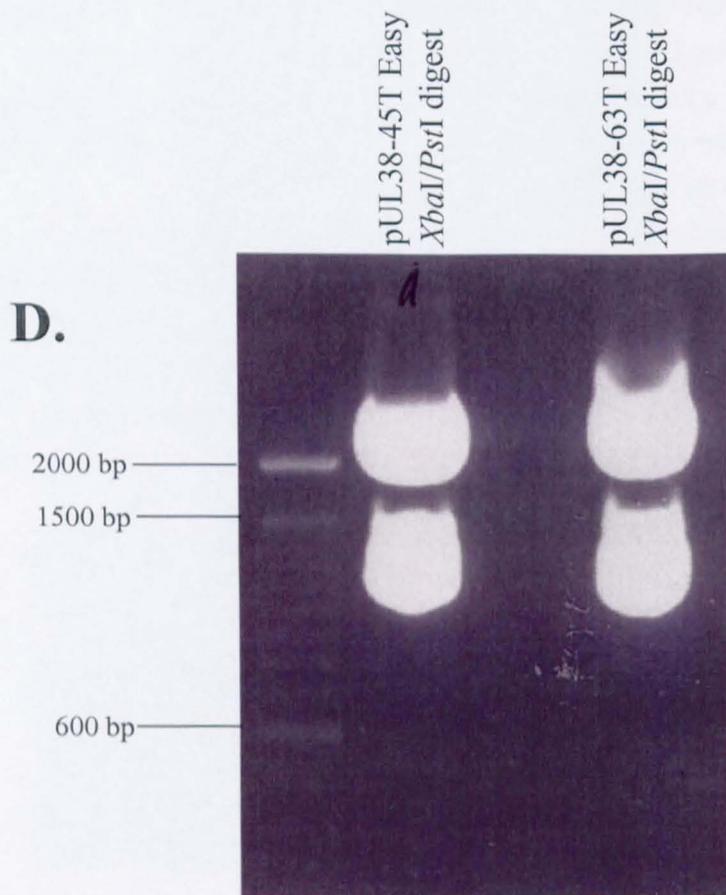
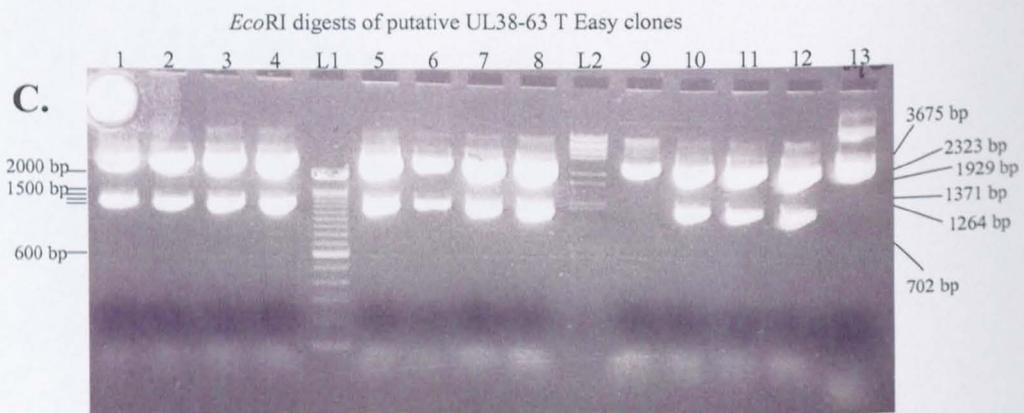
**B.**



**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**A.** Preparative agarose gel showing products from PCRs carried out using the forward primers UL38-45F and UL38-63F, and the reverse primer UL38R. pE38 was used as template DNA. The PCR using UL38-45F as a forward primer yields a product of 1290 bp. The PCR using UL38-63F as a forward primer yields a product of 1239 bp. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the right of the diagram.

**B.** Analytical agarose gel showing *EcoRI* digestion of the UL38-45F PCR products cloned into pGEM-T Easy. Bands of 1310 bp (insert) and 2997 bp (vector) indicate successful cloning (lanes 1, 3, 4, 5, 7, 8, 9, 10, and 11). Band sizes were estimated using a 100 bp ladder (New England Biolabs) (lane L1) and a  $\lambda$  DNA *BstEII* digest ladder (New England Biolabs) (lane L2). A reconstruction of the 100 bp ladder is shown to the left of the diagram, showing the sizes of the most prominent bands. Selected bands from the  $\lambda$  DNA *BstEII* digest ladder are shown to the right of the diagram.



**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**C.** Analytical agarose gel showing *EcoRI* digestion of the UL38-63F PCR product cloned into pGEM-T Easy. Bands of 1259 bp (insert) and 2997 bp (vector) indicate successful cloning (all numbered lanes other than lanes 9 and 13). Band sizes were estimated using a 100 bp ladder (New England Biolabs) (lane L1) and a  $\lambda$  DNA *BstEII* digest ladder (New England Biolabs) (lane L2). A reconstruction of the 100 bp ladder is shown to the left of the diagram, showing the sizes of the most prominent bands. Selected bands from the  $\lambda$  DNA *BstEII* digest ladder are shown to the right of the diagram.

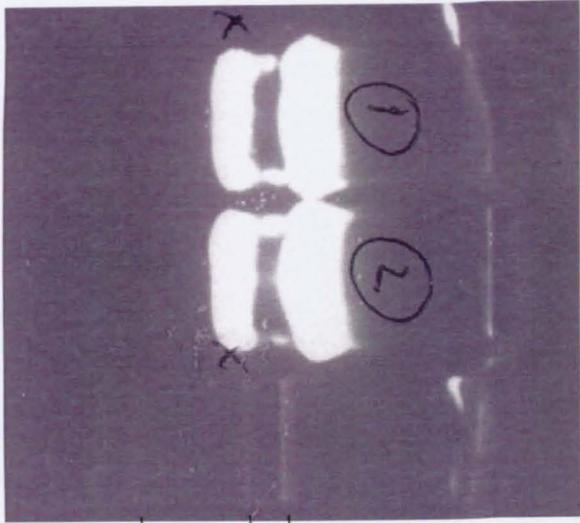
**D.** UL38 sequences were isolated from pUL38-45T Easy and pUL38-63T Easy by digestion with *XbaI* and *PstI* separated in a preparative agarose gel. This digest should yield bands of 1318 bp (insert) and 2989 bp (vector) in pUL38-45T Easy and bands of 1267 bp (insert) and 2989 bp (vector) in pUL38-63T-Easy. In each case the smaller band was isolated and purified. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the left of the diagram.

**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**E.** Analytical agarose gel showing *Xba*I/*Pst*I digestion of UL38-45 cloned into pFASTBAC1. This digest should yield bands of 1318 bp (insert) and 4763 bp (vector). This band pattern was observed in all clones examined (lanes 1-12). Band sizes were estimated using a 100 bp ladder (New England Biolabs) (lane L1) and a  $\lambda$  DNA *Bst*EII digest ladder (New England Biolabs) (lane L2). The sizes of the most prominent bands on the 100 bp ladder are shown to the left of the diagram. Selected bands from the  $\lambda$  DNA *Bst*EII digest ladder are shown to the right of the diagram.

**F.** Analytical agarose gel showing *Xba*I/*Pst*I digestion of UL38-63 cloned into pFASTBAC1. This digest should yield bands of 1267 bp (insert) and 4763 bp (vector). This band pattern was observed in all clones examined (lanes 1-12). Band sizes were estimated using a 100 bp ladder (New England Biolabs) (lane L1) and a  $\lambda$  DNA *Bst*EII digest ladder (New England Biolabs) (lane L2). The sizes of the most prominent bands on the 100 bp ladder are shown to the left of the diagram. Selected bands from the  $\lambda$  DNA *Bst*EII digest ladder are shown to the right of the diagram.

G.

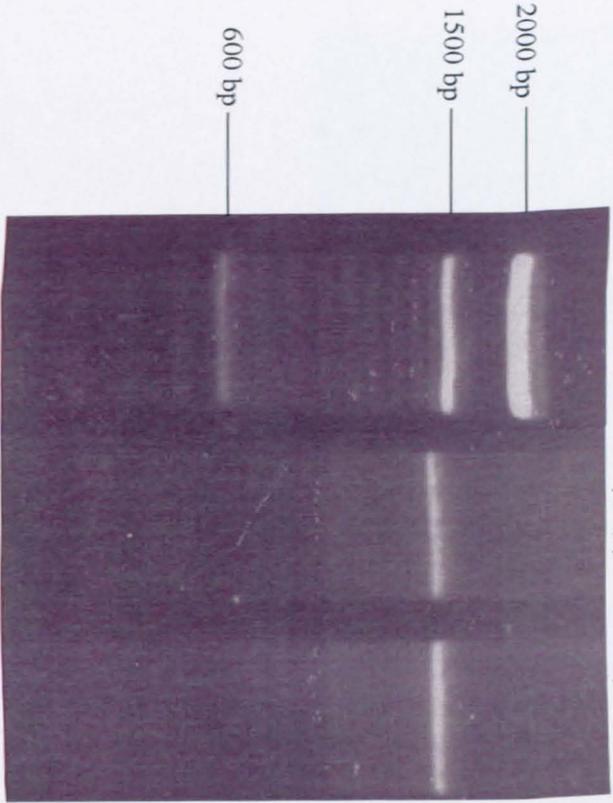


pUL38-45T Easy  
*EcoRI/SpeI* digest

pUL38-63T Easy  
*EcoRI/SpeI* digest

2000 bp  
1500 bp  
600 bp

H.



UL38-45PCR F1/  
38R-Xba PCR

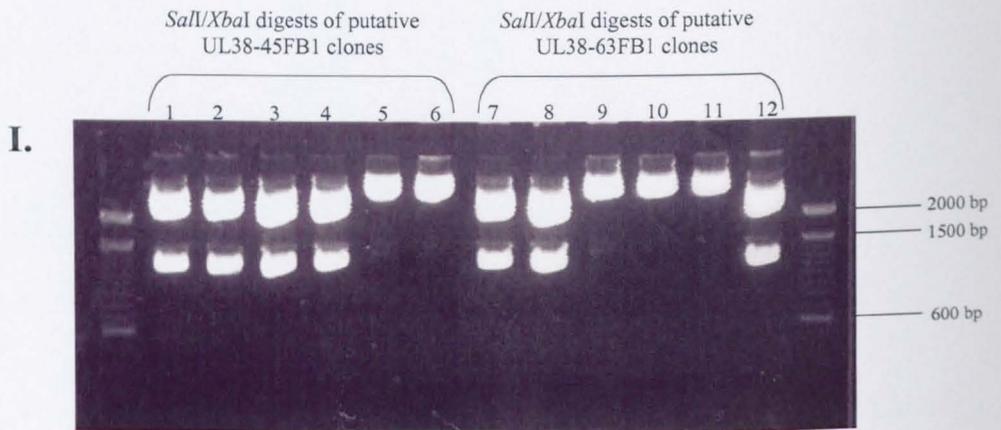
UL38-63PCR F1/  
38R-Xba PCR

2000 bp  
1500 bp  
600 bp

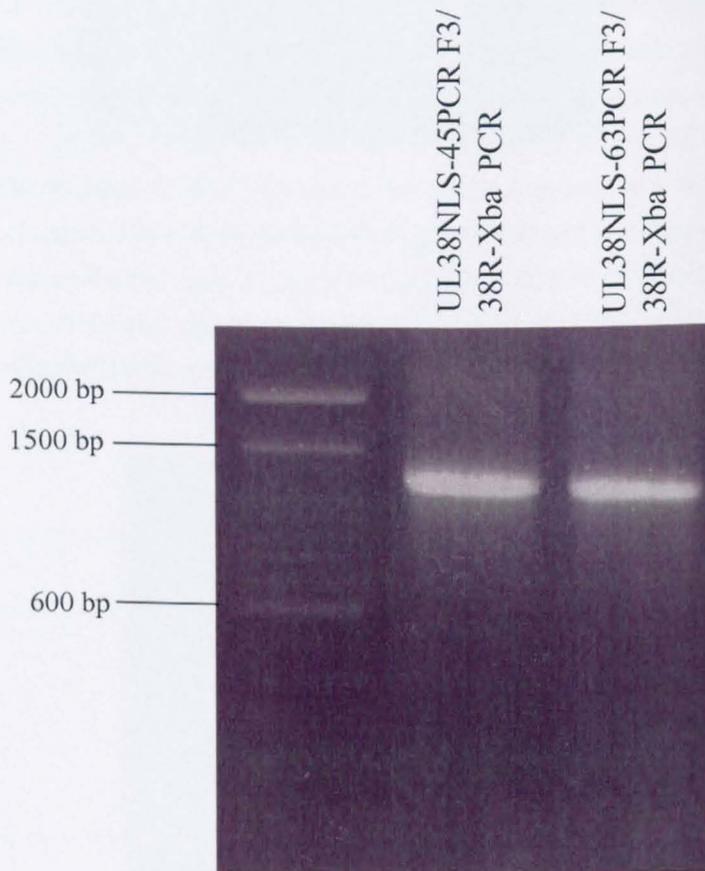
**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**G.** UL38 sequences were isolated from pUL38-45T Easy and pUL38-63T Easy by digestion with *EcoRI* and *SpeI*. This digest should yield bands of 1304 bp (insert) and 3003 bp (vector) in pUL38-45T Easy, and should yield bands of 1253 bp (vector) and 3003 bp (insert) in pUL38-63T Easy. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the right of the diagram.

**H.** Preparative agarose gel showing PCR reactions were carried out using the forward primers UL38-45PCR F1 and UL38-63PCR F1, and the reverse primer 38R-Xba. For the UL38-45 construct, pE38 was used as template DNA. For the UL38-63 construct, pUL38-45T Easy was used as a template. The PCR using UL38-45 PCR F1 as a forward primer should yield a product of 1290 bp. The PCR using UL38-63PCR F1 as a forward primer should yield a product of 1239 bp. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the left of the diagram.



**J.**



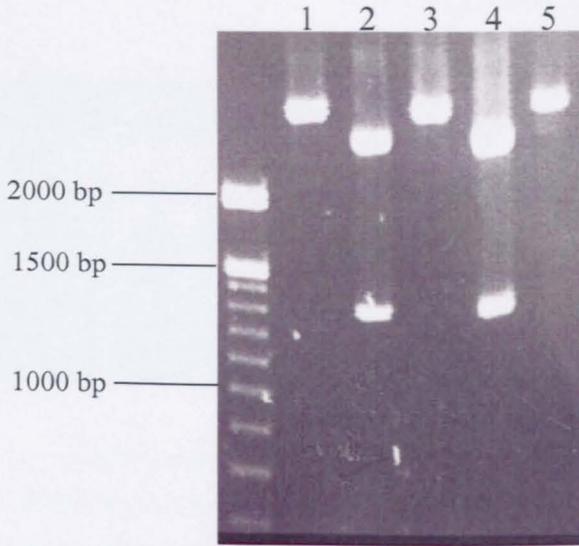
**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**I.** Analytical agarose gel showing *Sall/XbaI* digests of the UL38-45PCR F1 and UL38-63PCR F1 products cloned into pGEM-T Easy. A *Sall/XbaI* digest of pUL38-45PCR-T Easy should yield bands of 1305 bp and 3003 bp. A *Sall/XbaI* digest of pUL38-63PCR-T Easy should yield bands of 1256 bp and 3003 bp. Clones with inserts in the correct orientation for pUL38-45PCR-T are shown in lanes 1, 2, 3, and 4. Clones with inserts in the correct orientation for pUL38-63PCR-T are shown in lanes 7, 8, and 12). Band sizes were estimated using a 100 bp ladder (New England Biolabs) (shown on the outer lanes of the diagram). The sizes of the most prominent bands on the ladder are shown to the right of the diagram.

**J.** Preparative agarose gel showing products from PCR carried out using primers UL38NLS-45PCR F3 and 38R-Xba, with pUL38-45PCR-T Easy as template. This PCR reaction should yield a product of 1302 bp. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the left of the diagram.

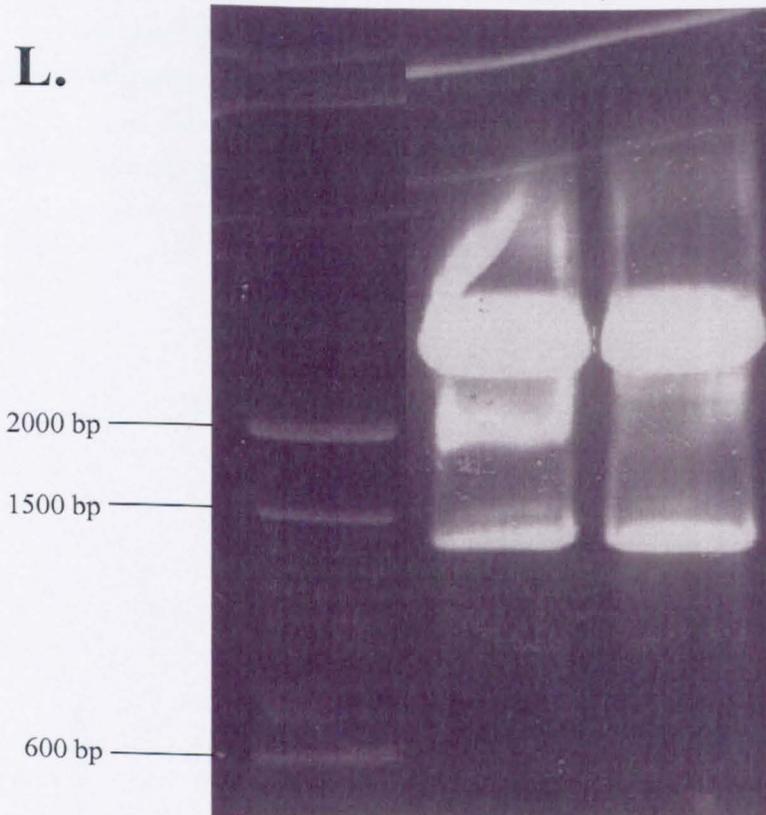
*SalI* digests of putative pUL38-45NLS-T Easy clones

**K.**



*EcoRI/XbaI* digests of pUL38-45NLS-T Easy

**L.**

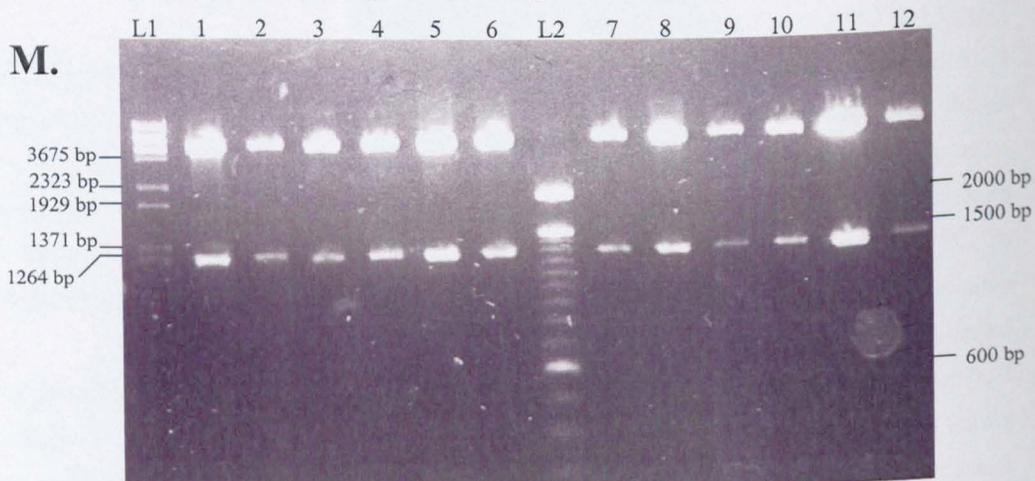


**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**K.** Analytical agarose gel showing *SalI* digestion of pUL38-45NLS-T Easy. A *SalI* digest should yield bands of 1305 bp and 3015 bp if the PCR product has inserted in the correct orientation (lanes 2 and 4). Band sizes were estimated using a 1 kb ladder. **L.** Preparative agarose gel showing an *EcoRI/XbaI* digest of pUL38-45NLS-T Easy. This digest yields bands of 1307 bp (insert) and 3013 bp (vector), as observed in both lanes shown. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the left of the diagram.

**L.** Preparative gel showing an *EcoRI/XbaI* digest of pUL38-45NLS-T Easy. This digest yields bands of 1307 bp (insert) and 3013 bp (vector), as observed in the two lanes shown. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The sizes of the most prominent bands on the ladder are shown to the left of the diagram.

*EcoRI/SphI* digests of putative pUL38-45NLSFBpCI clones



**Figure 4.2:** Cloning of the UL38 N-terminal truncation.

**M.** Analytical agarose gel showing *EcoRI/SphI* digestion of pUL38-45NLSFBpCI to determine whether cloning had been successful. This digest should yield bands of 23 bp, 33 bp, 1288 bp, and 5875 bp. Band sizes were estimated using a  $\lambda$  DNA *BstEII* digest ladder (New England Biolabs) (lane L1 and a 100 bp ladder (New England Biolabs) (lane L2). Selected bands from the  $\lambda$  DNA *BstEII* digest ladder are shown to the left of the diagram. The sizes of the most prominent bands on the 100 bp ladder are shown to the right of the diagram.

<b>Primer ID</b>	<b>Primer Sequence</b>
UL38-45F	5' GTCTAGAGGATCCATATGGGCCCCAGGGCTCCCCAGGGGGTTC 3'
UL38-63F	5' GTCTAGAGGATCCATATGTTGCTTGGCCTGGACGGCACAGAC 3'
UL38R	5' GACAAGATCTCACGCGCATGCCCGCCACTCGCCGGGG 3'
UL3845PCR F1	5' GAAGAAGAAGCGCAAGGTCGACGGCCCAGGGCTCCCCAGGGGGTTCG 3'
UL38-63 PCR F1	5' GAAGAAGAAGCGCAAGGTCGACTTGCTTGGCCTGGACGGCACAGACG 3'
38R-Xba	5' TCTAGATCACGCGCATGCCCGCCACTCGC 3'
UL38NLS-45PCR F3	5' CGCAATGGCTCCGAAGAAGAAGCGCAAGGTCGACGGCCCAGGGCTCCCCAGGGGGTTCGG 3'
UL38ATG	5' GACAGAATTTCGCAATGAAGACCAATCCGC 3'
UL38AA46	5' ACAGTCGACCCGTCGCGGCGAGAGATCGG 3'
UL38AA56	5' ACAGTCGACCGGGGTCCCACCCCTGGG 3'
UL38AA66	5' ACAGTCGACCCAAGCAACCACAGCGTGC 3'
UL38AA76	5' ACAGTCGACGCCCCAGGGGGCGCGTCTG 3'
UL38AA83	5' ACAGTCGACGTATCGTCGTTGGGGGTTCAGC 3'
UL38AA118	5' ACAGTCGACAGATCCGTCAGGATCACTTGG 3'
UL38-24 NTERM	5' GCAATGACACGCGATACCGCGGGACAG 3'
UL38-33 NTERM	5' GCAATGCTTCGGCGGTCCTGCGCCCC 3'
GFP UPSTREAM	5' GCTGATTATGATCTAGAGTTCG 3'
UL38NspeI	5' ACTAGTTCGCAATGGCTCCGAAGAAGAAGCGCAAGG 3'
UL38CbgIII	5' AGATCTTAGATCACGCGCATGCCCGCCACTCG 3'

Table 4.1: PCR primers used in cloning UL38 N-terminal mutants.

The recognition sequences for *Bam*HI (GGTACC), *Bgl*II (AGATCT), *Sal*I (GTCGAC), *Xba*I (TCTAGA), and *Spe*I (ACTAGT) used in cloning are singly or doubly underlined.

pGEM-T Easy (Promega) in accordance with the manufacturers' instructions. Individual clones were grown and screened by *EcoRI* digestion (Figures 4.2B and 4.2C). Single clones were selected and designated pUL38-45T Easy and pUL38-63T Easy (Figure 4.1A). For expression in insect cells, the UL38 sequences were isolated from pUL38-45T Easy and pUL38-63T Easy by digestion with *XbaI* and *PstI* (Figure 4.2D) and cloned downstream from the polyhedrin promoter in *XbaI/PstI* digested pFASTBAC1 (Invitrogen) (Figure 4.1A). Individual clones were grown and screened by digestion with *XbaI* and *PstI* (Figures 4.2E and 4.2F). Single clones were selected and designated pUL38-45FB1 and pUL38-63FB1. These plasmids were then used to generate recombinant baculoviruses using the FASTBAC protocol (Invitrogen). During the course of this work, immunolocalisation was carried out which suggested that the protein did not go to the nucleus efficiently in baculovirus infected SF21 cells. The work up to this point was carried out by David McNab. Because of the unexpected effects of these mutations on the behaviour of VP19C, it was decided to investigate the role of the N-terminal region further. Therefore, I carried out an investigation of their properties in mammalian cells where they could be examined in the context of HSV-1 infection.

#### **4.2.1. Nuclear Localisation of N-terminal Deletion Mutants**

For expression in mammalian cells, the UL38 sequences were isolated from pUL38-45T Easy and pUL38-63T Easy by digestion with *EcoRI* and *SpeI* (Figure 4.2G) and cloned downstream from the HCMV IE1 promoter in *EcoRI/XbaI* digested pFBpCI (see Section 2.1.9) to generate pUL38-45FBpCI and pUL38-63FBpCI\* (Figure 4.1B). Initial experiments with these constructs appeared to show that neither VP19C-45 nor VP19C-63 were efficiently localised to the nucleus. Failure to locate to the nucleus would be expected to interfere with capsid assembly. Therefore it was decided to insert a NLS into each of the mutant proteins. Consequently, the well-characterised SV40 large T antigen NLS (Kalderon *et al.*, 1984) was fused to the N-termini of VP19C-45 and VP19C-63.

##### **4.2.1.1. Construction of VP19C-NLS fusions**

Attempts to insert the SV40 NLS in front of the VP19C truncations were unsuccessful, so a two-step cloning strategy was adopted. PCR reactions were carried out (Figure 4.1C) using the forward primers UL38-45PCR F1 and UL38-63PCR F1 and the reverse primer 38R-Xba (Table 4.1). For the UL38-45 construct, pE38 (Nicholson *et al.*, 1994) was used as a template. For the UL38-63 construct, pUL38-63T Easy (described above) was used as a

template. UL38-45PCR F1 and UL38-63PCR F1 introduce *SalI* sites immediately upstream of amino acids 45 and 63 in the UL38 ORF. The PCR products (Figure 4.2H) were purified and cloned into pGEM-T Easy (Figure 4.1C). To determine whether the PCR products had inserted into pGEM-T Easy in the correct orientation, the plasmids were digested with *SalI* (Figure 4.2I). Single clones with inserts in the correct orientations were designated pUL38-45PCR-T Easy and pUL38-63PCR-T Easy\* respectively.

An oligonucleotide, UL38NLS-45PCR F3, was designed that specifies the nuclear localisation signal (NLS) of the SV40 large T antigen (sequence PKKKRKV (Kalderon *et al.*, 1984)) fused to the N-terminus of the UL38 ORF truncated at amino acid 45 (Table 4.1). A PCR reaction was performed (Figure 4.1D) using primers UL38NLS-45PCR F3 and 38R-Xba, with pUL38-45PCR-T Easy as template (Figure 4.2J). The PCR fragment was purified and cloned into pGEM-T Easy to generate pUL38-45NLS-T Easy (Figure 4.1D). To determine whether the PCR products had inserted into pGEM-T Easy in the correct orientation, the plasmids were digested with *SalI*. (Figure 4.2K). The UL38 insert was isolated by digestion with *EcoRI* and *XbaI* (Figure 4.2L) and ligated into *EcoRI/XbaI* digested pFBpCI to generate pUL38-45NLSFBpCI (Figure 4.1D). The insertions were identified by *EcoRI/SphI* digestion (Figure 4.2M). The *SalI/XbaI* fragment from pUL38-63PCR-T Easy\* was then purified and ligated into *SalI/XbaI* digested pUL38-45NLSFBpCI to generate pUL38-63NLSFBpCI\* (Figure 4.1E).

#### **4.2.2. 'Repairing' the Mutation in pUL38-63FBpCI\***

At this point the complete sequence of pUL38-63FBpCI\* became available and it was noticed that it contained an unexpected point mutation in the UL38 ORF: the guanine that appears at position 493 in the wild-type ORF had been substituted for an adenine. This resulted in an alteration in the amino acid sequence, with a glutamine residue (GAG) being substituted for a lysine (AAG) (Figure 4.3). Subsequent analysis of the sequence of the NLS-containing UL38-63 construct (pUL38-63NLSFBpCI\*) showed that it contained the same point mutation in the UL38 ORF. In order to ensure that any phenotypic differences were not due to this alteration, the UL38 ORFs in these plasmids had to be repaired.

To construct a 'repaired' version of pUL38-63FBpCI\*, pUL38-45NLSFBpCI (Section 4.2.1.1) was digested with *EcoRI* and *PvuII* and a 6836 bp band (containing the section of the UL38 ORF that codes for amino acids 163-465 of wild-type VP19C and the backbone of the pFBpCI plasmid) was isolated. At the same time, pUL38-63FBpCI\* was also

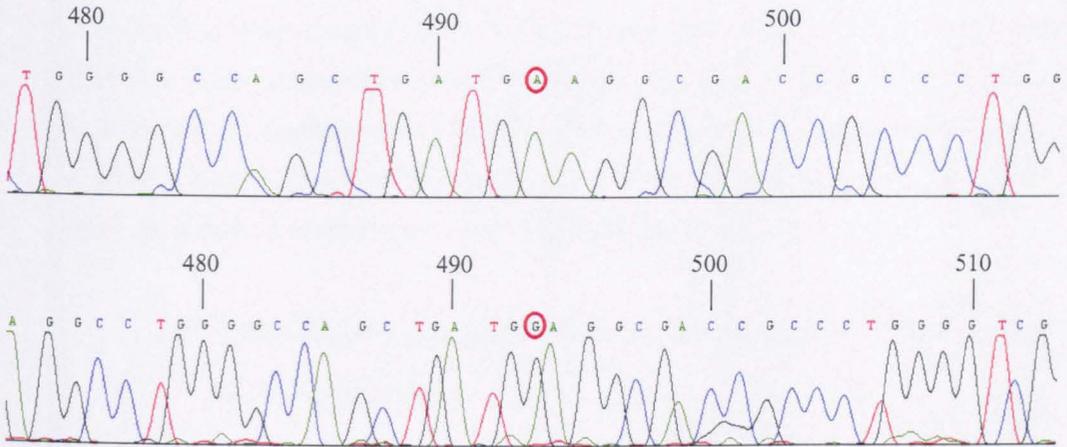


Figure 4.3: The mutation in pUL38-63FBpCI\*, and its subsequent correction.

**Above:** When pUL38-63FBpCI was sequenced, it was found to contain an unexpected point mutation in the UL38 ORF: the guanine at position 493 in the wild-type ORF had been substituted by an adenine (encircled). **Below:** This mutation was 'repaired' to make pUL38-63FBpCI. The guanine at position 493 is encircled.

Sequencing results were produced using the Chromas program ([www.technelysium.com.au/chromas.html](http://www.technelysium.com.au/chromas.html)).

digested with *EcoRI* and *PvuII* and a 324 bp band (containing the section of UL38 that codes for amino acids 63-162 of wild-type VP19C) was isolated. The two purified DNAs were ligated together to generate pUL38-63FBpCI (Figure 4.1F). The mutation occurs in amino acid 165 of the VP19C amino acid sequence, and as pUL38-45NLSFBpCI does not contain this mutation, pUL38-63FBpCI should not contain it either. Similarly, to remove the secondary mutation in pUL38-63NLSFBpCI\*, it was digested with *PvuII* and *KpnI*, and a band of 6198 bp was isolated. In addition, pUL38-45NLSFBpCI was also digested with *PvuII* and *KpnI*, and a band of 967 bp was isolated. The two purified DNAs were ligated together to produce pUL38-63NLSFCpCI (Figure 4.1G). The sequence of the resulting DNA was determined, and was shown to contain a guanine at position 493. The repaired versions of pUL38-63FBpCI and pUL38-63NLSFBpCI with the correct (wild-type) nucleotide sequences were used in all subsequent studies.

### **4.2.3. Nuclear Localisation and Complementation of VP19C-NLS Fusions**

To determine the effects of the N-terminal truncations on the behaviour of VP19C, the distributions of VP19C-45 and VP19C-63 in cells transfected with pUL38-45FBpCI and pUL38-63FBpCI were examined. In marked contrast to the pattern seen with wild-type VP19C (Figures 3.7 and 3.8), neither VP19C-45 (Figure 4.4B) nor VP19C-63 (Figure 4.4E) was specifically nuclear. These results confirm the initial finding from the baculovirus-infected cells (see above), and suggested that the N-terminal region of the protein is important for nuclear localisation.

In contrast to VP19C-45 and VP19C-63, VP19C-45NLS (Figure 4.4C) and VP19C-63NLS (Figure 4.4F) both localised efficiently to the nucleus, confirming that the exogenous NLS functioned as expected. Furthermore, when pE18 (expressing VP23) was co-transfected with pUL38-45FBpCI and pUL38-63FBpCI, VP19C-45NLS (Figure 4.4D) and VP19C-63NLS (Figure 4.4G) were also able to transport VP23 to the nucleus.

In a further attempt to analyse the effects of the N-terminal truncations on the properties of VP19C, their ability to complement the growth of the HSV-1 deletional mutant  $\nu\Delta 38\text{YFP}$  (Section 2.1.6; Thurlow *et al.*, 2005) was examined. pUL38-45FBpCI, pUL38-63FBpCI, and pE38 (encoding wild-type VP19C) DNAs were transfected into BHK-21 cells, which were incubated at 37°C for five hours before being infected with 2 pfu/cell of  $\nu\Delta 38\text{YFP}$ . After one hour, the unabsorbed input virus was neutralised by an acid wash (Section

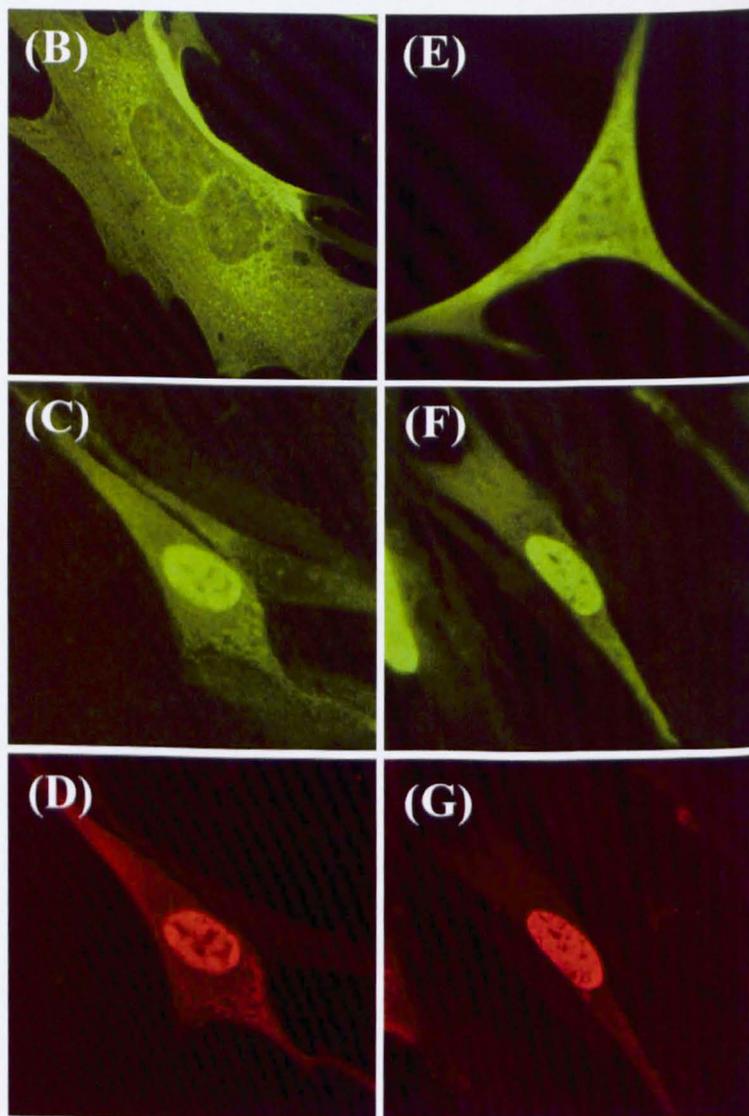
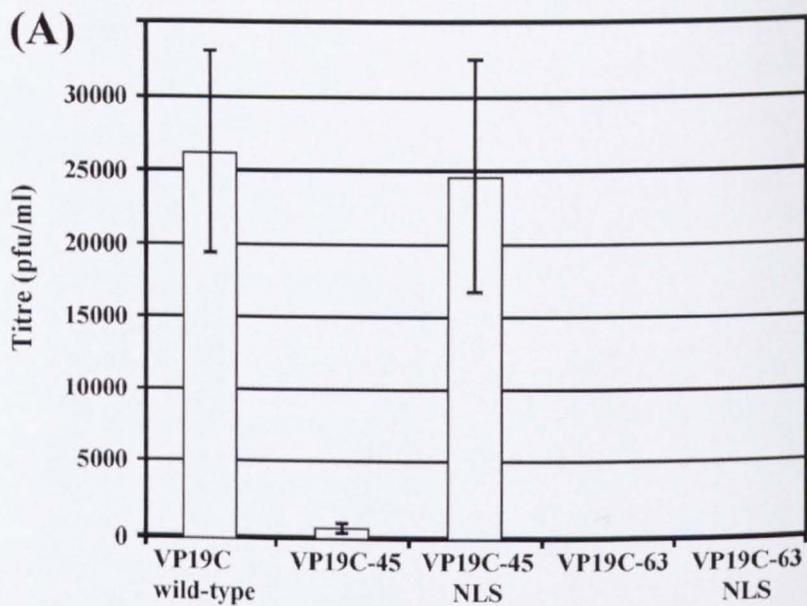


Figure 4.4: Roles of the VP19C N-terminal sequences.

(A) Complementation of growth of the VP19C null mutant v $\Delta$ 38YFP by transfected plasmid pUL38-45FBpCI (expressing VP19C-45), pUL38-63FBpCI (VP19C-63), pUL38-45NLSFBpCI (VP19C-45NLS), or pUL38-63NLSFBpCI (VP19C-63NLS) was carried out in BHK-21 cells. The progeny virus was titrated on UL38RSC cells. The error bars indicate the standard errors of the means. (B to G) Intracellular localisation of N-terminal mutants of VP19C. BHK-21 cells were transfected singly with plasmids (B) pUL38-45FBpCI (expressing VP19C-45) and (E) pUL38-63FBpCI (VP19C-63), or were cotransfected with (C and D) pUL38-45NLSFBpCI (VP19C-45NLS) and pE18 (VP23) or with (F and G) pUL38-63NLSFBpCI (VP19C-63NLS) and pE18. VP19C (B, C, E, and F) was detected with the monoclonal antibody mAb02040, and visualised using FITC-conjugated GAM (green). VP23 (D and G) was detected with the antiserum rAb186, and visualised using TRITC-conjugated GAR (red).

2.2.4.4). Incubation was continued at 37°C for 40 hours, at which time the progeny virus was harvested and titrated on complementing UL38RSC cells (Section 2.1.4).

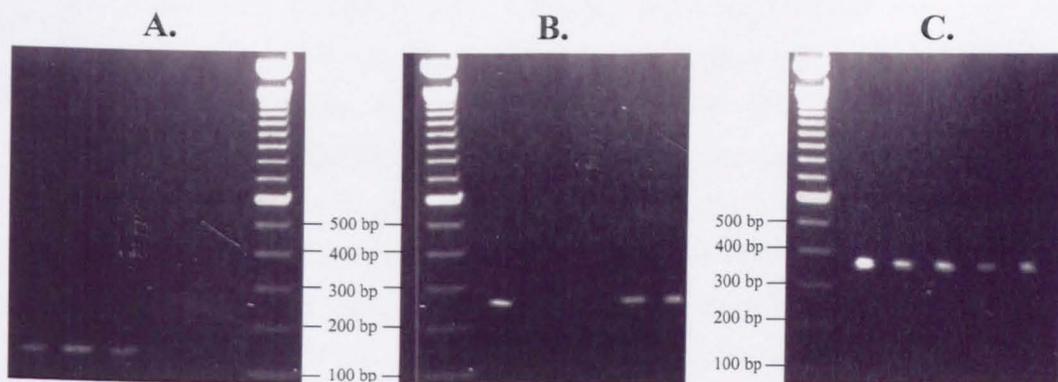
As shown in Figure 4.4A, the yield obtained with VP19C-45 was approximately 50 fold lower than for the wild-type VP19C control, while the removal of the first 63 amino acids completely abolished complementation, thereby demonstrating that the N-terminal sequences of VP19C are important for the production of wild-type levels of infectious virus.

When the ability of pUL38-45NLSFBpCI and pUL38-63NLSFBpCI to complement the growth of the HSV-1 deletional mutant  $\nu\Delta 38$ YFP was tested, VP19C-45NLS was found to complement  $\nu\Delta 38$ YFP as effectively as wild type VP19C (Figure 4.4A), thereby demonstrating that the reduction in growth caused by the removal of the N-terminal 45 amino acids was primarily as a result of the defect in nuclear localisation. In contrast, no complementation occurred in VP19C-63NLS (Figure 4.4A), suggesting that amino acids 45 to 63 have an additional role in the formation of infectious virus.

### 4.3. Mapping the VP19C NLS

Although VP19C is transported to the nucleus in the absence of other viral proteins, no NLS was identified when the complete amino acid sequence was submitted to the NLS-prediction programs PredictNLS and PSORT II (<http://cubic.bioc.columbia.edu/predictNLS/> and <http://psort.nibb.ac.jp/form2.html> respectively). To confirm that the N-terminal region of VP19C did contain an NLS, increasing amounts of sequence from the N-terminus were fused to the green fluorescent protein (GFP) and the intracellular distribution of the chimeric proteins were determined.

DNA fragments containing increasing numbers of residues from the 5' end of the VP19C ORF were produced by PCR using the forward primer UL38ATG, and the reverse primers UL38AA46, UL38AA56, UL38AA66, UL38AA76, UL38AA83, and UL38AA118 (Table 4.1). The appropriate PCR products were purified and cloned into pGEM-T Easy. Individual colonies were grown and screened by digestion with *EcoRI*, which makes two cuts flanking the insert in pGEM-T Easy. Single clones were selected and designated pGEM-AA46, pGEM-AA56, pGEM-AA66, pGEM-AA76, pGEM-AA83, and pGEM-AA118. The UL38 inserts were isolated by digestion with *EcoRI* and *SalI* (Figure 4.5) and ligated into *EcoRI/SalI* digested pEGFP-N1 (Clontech) to generate p46AA-UL38-GFP,



**Figure 4.5:** Analytical agarose gels showing *EcoRI/SalI* digests of representative pGEM-T Easy N-terminal VP19C clones.

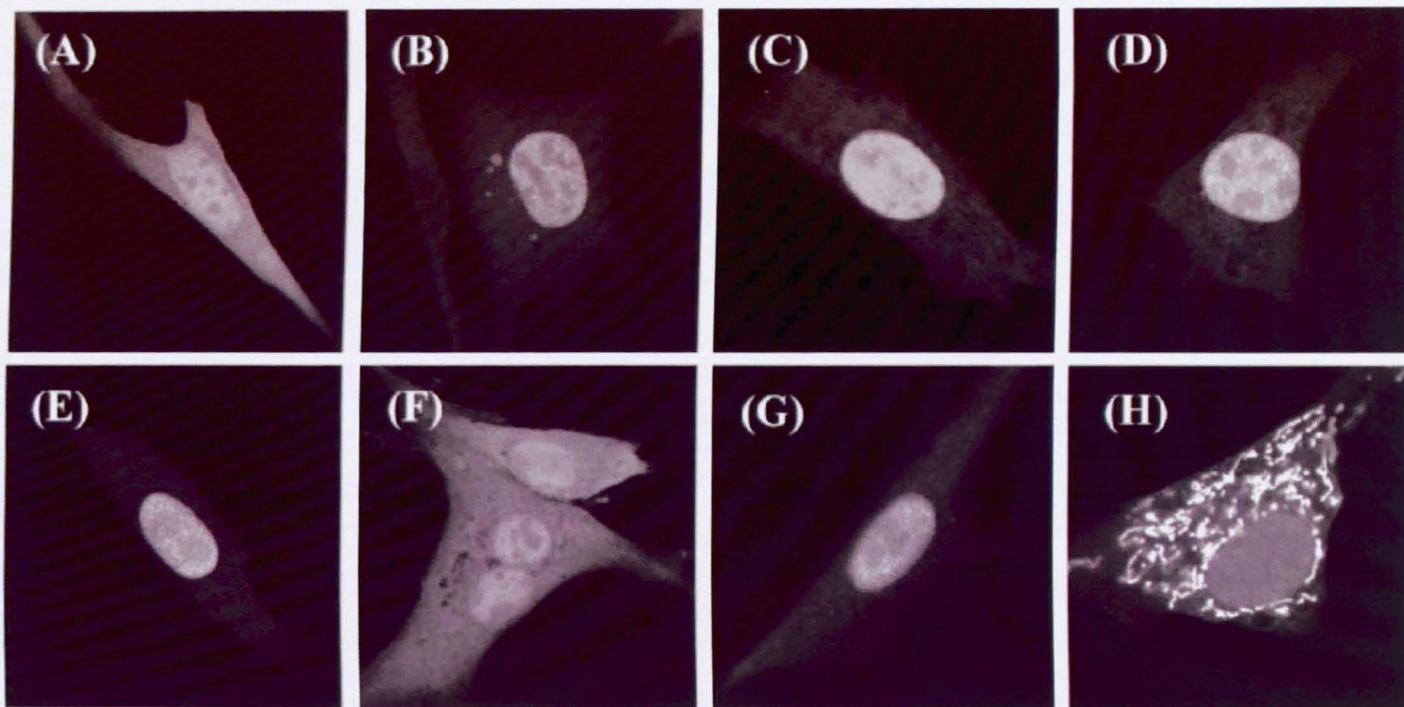
PCR was carried out using pE38 as template DNA, the forward primer UL38ATG, and suitable reverse primers as described in the text. The PCR products were ligated into pGEM-T Easy. *EcoRI/SalI* digests were carried out on the resulting constructs. **A.** Digestion of pGEM-AA46 with *EcoRI* and *SalI* produces a band of 146 bp. **B.** Digestion of pGEM-AA83 with *EcoRI* and *SalI* produces a band of 257 bp. **C.** Digestion of pGEM-AA118 with *EcoRI* and *SalI* produces a band of 362 bp. Band sizes were estimated using a 100 bp ladder (New England Biolabs). Selected band sizes in the 100 bp ladder are indicated.

p56AA-UL38-GFP, p66AA-UL38-GFP, p76AA-UL38-GFP, p83AA-UL38-GFP, and p118AA-UL38-GFP. All the plasmids were checked by diagnostic restriction enzyme digestion and DNA sequencing.

To determine their intracellular localisation properties, BHK-21 cells were transfected with the individual fusion protein constructs and incubated at 37°C for 16 hours before they were fixed with 4% formaldehyde. The distribution of green fluorescence was visualised on a confocal microscope. In cells expressing GFP alone, the fluorescence was uniformly distributed between the cytoplasm and the nucleus (Figure 4.6A). However, fusing GFP to the N-terminal 118, 83, 76, 66 or 56 amino acids of VP19C converted it to a nuclear protein (Figures 4.6B-4.6E). In contrast, the distribution seen with the fusion containing the N-terminal 46 amino acids resembled the GFP control (Figure 4.6F). Thus, it was clear that an NLS is located within the first 56 residues of VP19C. Examination of the amino acid sequence in this region revealed a high concentration of arginine residues extending from amino acids 25 to 57 (Figure 4.6I). Since stretches of basic amino acids are characteristic of many NLSs (Palmeri and Malim, 1999, and references therein), amino acid sequences from within this region were also fused to GFP.

To delete sequences from the N-terminus, PCR was carried out with p66AA-UL38-GFP as a template and using the forward primers UL38-24 NTERM and UL38-33 N-TERM and the reverse primer GFP UPSTREAM (Table 4.1). The resulting PCR products were cloned into pGEM-T Easy. Individual colonies were grown and screened by digestion with *EcoRI*. Inserts were isolated by digestion with *EcoRI* and *NotI* and ligated into *EcoRI/NotI* digested p66AA-UL38-GFP to generate p24-66AA-UL38-GFP and p33-66AA-UL38-GFP.

In p24-66AA-UL38-GFP, nuclear localisation of GFP was observed (Figure 4.6G), confirming that the NLS mapped within the region comprising residues 24 to 66. However, the intracellular distribution of GFP that was observed for p33-66AA-UL38-GFP was unexpected and distinctive: producing a cytoplasmic pattern that resembled the distribution expected for mitochondria (Figure 4.6H). Submission of the VP19C sequence to the subcellular localisation site prediction program iPSORT (<http://biocaml.org/ipsort/iPSORT/#predict>) identified amino acids 32 to 40 (LLRRVLRPP) as a potential mitochondrial localisation signal. Given what is known about the functions of VP19C, it seems unlikely that this mitochondrial localisation signal is functional in the context of the intact protein, so this was not investigated further.



(I)

24                      33                                      46                                      56                                      66  
 |                                      |    |    |    |  
 MKTNPLPATPSVWGGSTVELPPTTRDTAGQGLLRVLRPPISRRDGPGLPRGSGPRRAASTLWLLGLDGT

**Figure 4.6:** Mapping the VP19C NLS

BHK-21 cells were transfected with plasmids expressing the GFP protein fused in-frame to the N-terminal **(B)** 83, **(C)** 76, **(D)** 66, **(E)** 56, or **(F)** 46 amino acids of VP19C. **(G and H)** GFP is fused to amino acids 24 to 66, and 33 to 66 of VP19C. **(a)** GFP control. **(i)** N-terminal 70 amino acids of VP19C with the positions of various truncations indicated. Arginine residues are underlined.

Although further analysis would be required to determine the precise composition of the VP19C NLS, the results presented here map it to an arginine-rich sequence between residues 24 and 56.

#### **4.4. Nature of the growth defect associated with VP19C-63NLS**

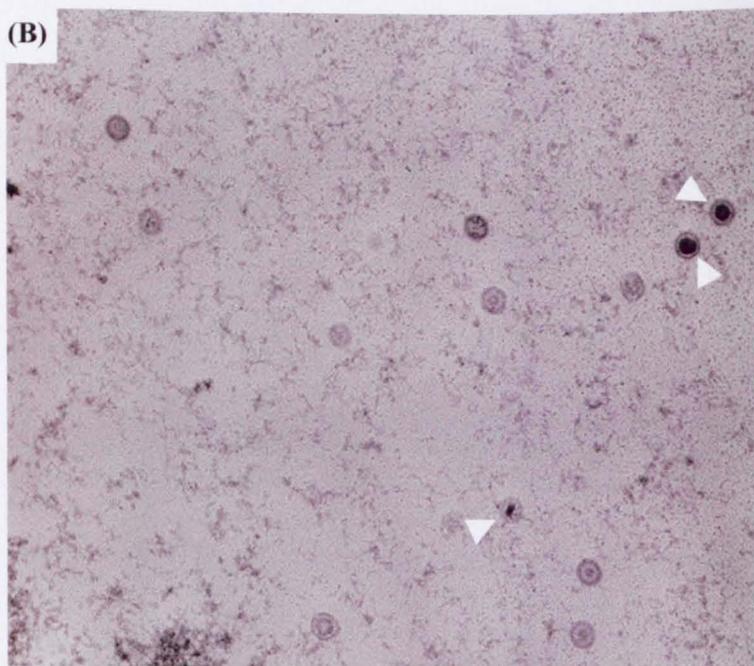
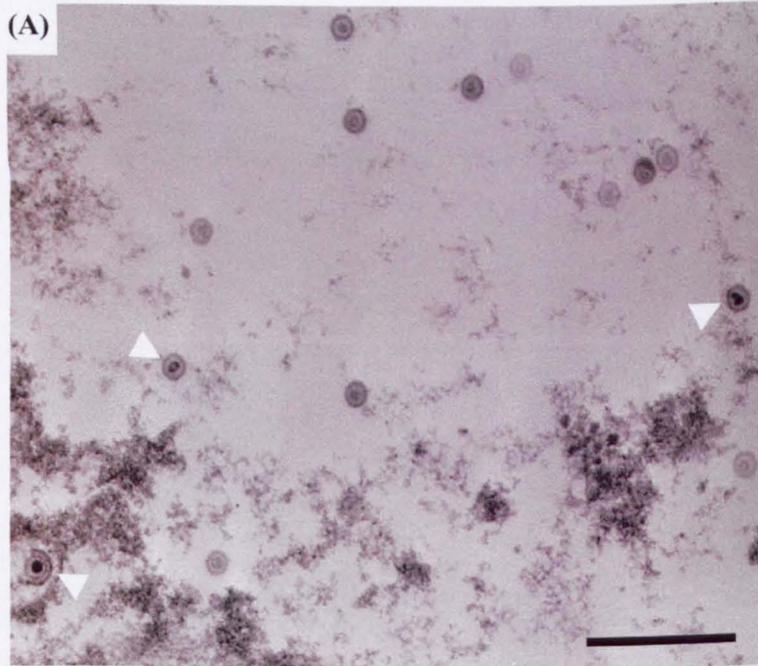
As described earlier, VP19C-63NLS was completely unable to substitute for wild-type VP19C in virus infection. To determine at what stage during infection the block on replication was occurring, electron microscopy was carried out. To ensure efficient complementation, cells were infected with recombinant baculoviruses expressing the UL38 N-terminal deletion mutants. Although not normally active in mammalian cells, genes cloned into baculovirus vectors can be expressed if their expression is driven by a mammalian promoter. U2OS cells were selected for this procedure due to their ability to be efficiently transduced with baculoviruses.

bUL38FBacpCI was produced during the course of the work for this thesis, and bUL38-45NLSFBacpCI were made by David McNab. UL38-63NLSFBacpCI was made in accordance with the protocol detailed in the BAC-TO-BAC Baculovirus Expression Systems manual (Invitrogen) and described in Section 2.2.7. As the UL38-63NLS ORF was already contained on a modified pFASTBAC plasmid (pFBpCI), no further cloning steps were required.

Duplicate 35 mm plates of U2OS cells (one containing a glass coverslip) were incubated with 50 pfu/cell of baculoviruses bUL38FBacpCI (expressing VP19C), bUL38-45NLSFBacpCI (VP19C-45NLS) and bUL38-63NLSFBacpCI (VP19C-63NLS) for one hour at 37°C. The baculovirus inoculum was removed and the cells were infected with 5pfu/cell of  $\nu$ Δ38YFP and incubation was continued at 37°C. After 16 hours, the coverslip from one of the duplicate plates was removed and processed for immunofluorescence microscopy, while the remaining cells were used for western blotting. The second plate was harvested and processed for electron microscopy. Immunofluorescence (Figure 4.7) showed similar numbers of cells expressing the three forms of VP19C, while western blotting (Figure 4.8) confirmed that all three proteins were expressed at similar levels. Productively infected cells are easily distinguished by the presence of characteristic herpesvirus capsids in their nuclei. Examination of thin sections revealed the presence of capsids in the nuclei of 48% of the bUL38FBacpCI and 57% of the bUL38-

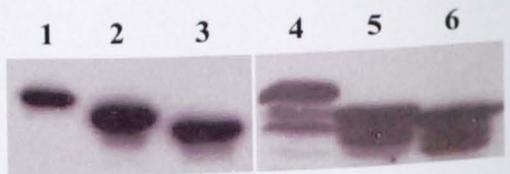
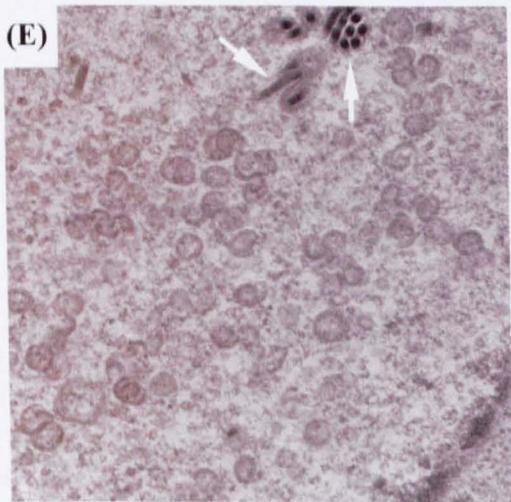
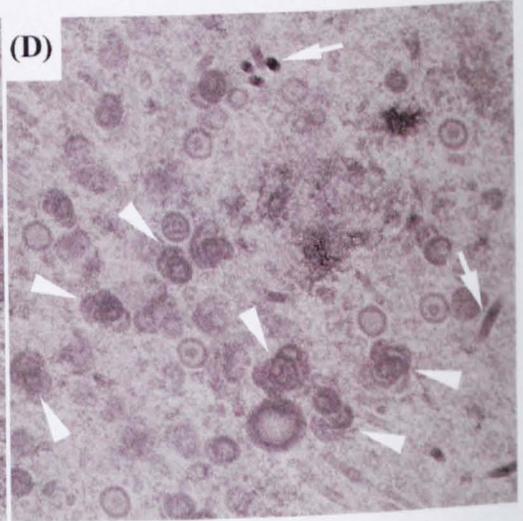
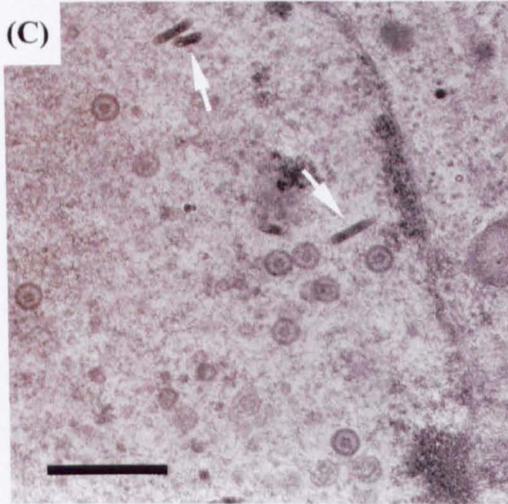


**Figure 4.7:** Expression of VP19C, VP19C-45NLS, and VP19C-63NLS in U2OS cells. U2OS cells were transduced separately with baculoviruses UL38FBacpCI (expressing VP19C), UL38-45NLSFBacpCI (VP19C-45NLS), and UL38-63NLSFBacpCI (VP19C-63NLS) for one hour at 37°C. All samples were also infected with vΔ38YFP. VP19C was detected with the monoclonal antibody mAb02040, and visualised using FITC GAM (green).



**Figure 4.8:** Effects of VP19C N-terminal mutations on capsid assembly.

**(A and B)** 35 mm plates of U2OS cells were transduced with 50 pfu/cell of baculovirus bUL38FBacpCI **(A)**, or bUL38-45NLSFBacpCI **(B)**, and incubated at 20°C for one hour before being infected with 5 pfu/cell of v $\Delta$ 38YFP. Infection was continued at 37°C for 16 hours. The cells were then harvested for electron microscopy. The majority of capsids shown in panels A and B are B-capsids and are not arrowed. C-capsids are indicated by triangles. Scale bar = 500 nm (applicable to both micrographs shown).



**(F)**

**Figure 4.8:** Effects of VP19C N-terminal mutations on capsid assembly.

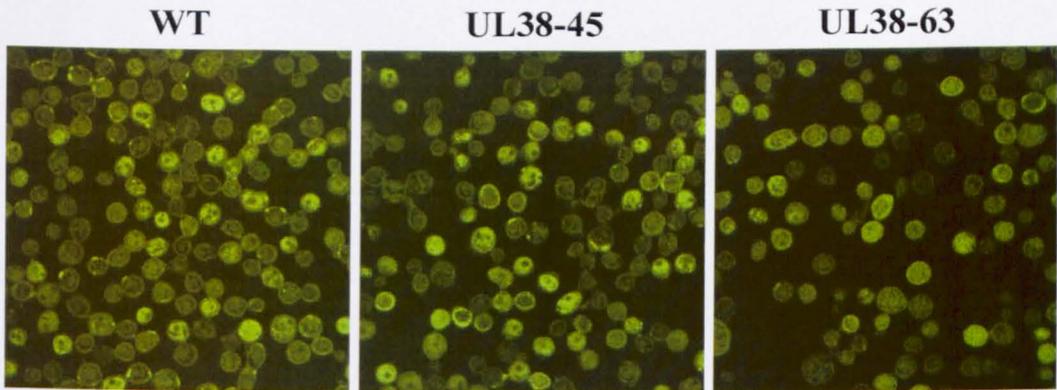
**(C to E)** 35 mm plates of Sf21 cells were infected with 5 pfu/cell each of baculoviruses AcUL19 (VP5), AcUL26.5 (preVP22a), and AcAB3.12 (VP23, VP26 and the UL26 protease), and with 5 pfu/cell of **(C)** AcUL38, **(D)** UL38-45FBac, or **(E)** UL38-63FBac, and incubated at 28°C for 48 hours. At the appropriate times after infection, the cells were harvested and prepared for electron microscopy. The majority of capsids in panels C to E are B capsids and are not arrowed. In panel D, arrowheads indicate concentrations of incomplete capsids. Only incomplete capsids are present in panel E, and are not arrowed. Baculovirus capsids (C to E) are indicated by arrows. Scale bar = 500 nm (applicable to all micrographs shown). **(F)** Expression of VP19C in the infected cells was confirmed by western blotting. Protein samples were separated on a 10% sodium dodecyl sulphate polyacrylamide gel and detected by enhanced chemiluminescence using mAb02040 and protein A-peroxidase. Lanes 1 and 2 show U2OS cell samples corresponding to panels A and B. Lane 3 shows a samples from U2OS cells transduced with UL38-63NLSFBacpCI and infected with vΔ38YFP. Lanes 4 to 6 show Sf21 cell samples corresponding to panels C to E.

45NLSFBacpCI transduced samples. As expected for VP19C proteins that support productive infection, all three capsid types (DNA containing (C-capsids), intermediate (B-capsids), and empty (A-capsids) were seen in cells expressing wild-type VP19C and VP19C-45NLS (Figure 4.8). In contrast, no virus-specified particles of any type were observed in over 100 nuclei of the bUL38-63NLSFBacpCI transduced samples. Published studies had reported that capsid assembly in a baculovirus model was not prevented by the removal of up to 90 amino acids from the N-terminus of VP19C, although it was less efficient than with full-length VP19C (Spencer *et al.*, 1998). Therefore, the failure to detect capsids in the bUL38-63NLSFBacpCI transduced U2OS cells was unexpected.

To investigate this further, capsid assembly was analysed using the more efficient baculovirus model system. Duplicate plates of Sf21 cells (one containing a glass coverslip) were infected separately with 5 pfu/cell of baculoviruses AcUL38 (expressing full-length VP19C), bUL38-45FBac (VP19C-45) or bUL38-63FBac (VP19C-63). All samples were also infected with 5 pfu/cell each of baculoviruses AcUL19 (VP5), AcUL26.5 (preVP22a) and AcAB3.12 (VP23, VP26 and the UL26 protease) to supply the remaining capsid proteins. As expected, characteristic herpesvirus B capsids were readily observed in cells expressing wild-type VP19C (Figure 4.8A), and in lower amounts in cells expressing VP19C-45 (Figure 4.8B). The reduced number of intact capsids seen with VP19C-45 is probably a result of inefficient nuclear localisation due to the absence of the NLS. However, in cells expressing VP19C-63, only incomplete capsid shells were seen (Figure 4.8C-E). Once again, immunofluorescence (Figure 4.9) and western blotting (Figure 4.8F) confirmed that all three proteins were being expressed at similar levels, thereby demonstrating that the block on particle formation was a direct result of the VP19C-63 deletion. Incomplete capsid shells will only be made if VP19C is present (Tatman *et al.*, 1994), therefore the presence of incomplete capsid shells in these cells confirms that VP19C-63 is being made.

## **4.5. VP19C-45NLS Difference Mapping**

The original intention in making the N-terminal deletions of VP19C was to locate the position of the N-terminus in capsid reconstructions. The inability of the VP19C-63 deletions to support capsid assembly meant that this could be attempted only with the shorter deletion. However, since VP19C-45NLS was as effective as wild-type VP19C in supporting virus growth, it was decided to construct a virus expressing this form of the protein and use it for capsid reconstruction. Plasmid pΔ38YFP, which was originally used



**Figure 4.9:** Expression of VP19C, VP19C-45, and VP19C-63 in Sf21 cells.

Sf21 cells were cotransfected separately with baculoviruses AcUL38 (expressing full-length VP19C), UL38-45Fbac (VP19C-45), or UL38-63Fbac (VP19C-63). All samples were also infected with baculoviruses AcUL19 (VP5), AcUL26.5 (preVP22a) and AcAB3.12 (VP23, VP26 and the UL26 protease) to supply the remaining capsid proteins. VP19C was detected with the monoclonal antibody mAb02040, and visualised using FITC GAM (green).

to make the UL38 minus mutant v $\Delta$ 38YFP (Thurlow *et al.* 2005), contains the YFP ORF in place of the UL38 ORF flanked by unique *NheI* and *BglII* restriction enzyme sites. Compatible restriction enzyme sites were introduced at either end of the UL38-45NLS ORF by PCR using UL38NspEI and UL38CbgII as primers, and pUL38-45NLS-T Easy as template. The resulting PCR product was ligated into pGEM-T Easy. Colonies were screened for the presence of the insert and one clone was selected and designated UL38-45NLSrescue pGEM (Figure 4.10A). A 1309 bp fragment containing the UL38-45NLS ORF was isolated into the *NheI/BglII* digested p $\Delta$ 38YFP. Colonies were screened for the presence of the insert and one clone was designated as UL38-45NLSrescue. (Figure 4.10B). To rescue the truncated UL38 back into virus, BHK cells were transfected with UL38-45NLSrescue, incubated for five hours at 37°C and infected with 2 pfu/cell of v $\Delta$ 38YFP. Progeny virus was harvested and titrated on non-complementing BHK cells. Individual plaques were screened for the absence of YFP. One isolate (designated v38-45NLS) was grown to high titre and used in subsequent experiments. To produce capsids for structural analysis, approximately  $5 \times 10^8$  BHK-21 cells were infected with 5 pfu/cell of v38-45NLS or HSV-1 strain 17. After 16 hours at 37°C, the cells were harvested and capsids purified by sucrose gradient sedimentation. The B-capsid samples were used in subsequent studies. Analysis of the protein composition of VP19C-45NLS capsids confirmed that they contained the truncated form of VP19C (Figure 4.11A). The purified capsids were then sent to Wah Chiu's laboratory at the Baylor College of Medicine, Houston, Texas, for electron cryomicroscopy. Electron cryomicroscopic images of the VP19C-45NLS B capsids (Figure 4.11B) showed that they were of similar size and conformation to wild-type B capsids, while 3D reconstruction confirmed that they conformed to the classical  $T=16$  structure of HSV-1 capsids. However, when the VP19C-45NLS capsids and wild-type B capsids were superimposed, differences became apparent (Figure 4.12). Surprisingly, the most obvious differences were at the tops of the hexons, where additional masses appeared to be present in the wild-type capsids, and in the tops of pentons, where the VP19C-45NLS capsids appeared to contain extra mass. These differences were not uniform across the capsids, but formed a gradient with the greatest apparent excesses of wild-type capsid mass towards the centres of the faces and the greatest deficiencies at the vertices. This pattern can be described most simply by a change in the overall curvature of the VP19C-45NLS capsid shell to give a more angular capsid with flatter faces and more prominent vertices.

To determine whether there had been any alteration in triplex appearance resulting from the N-terminal truncation of VP19C, the structures of triplexes computationally isolated

A.

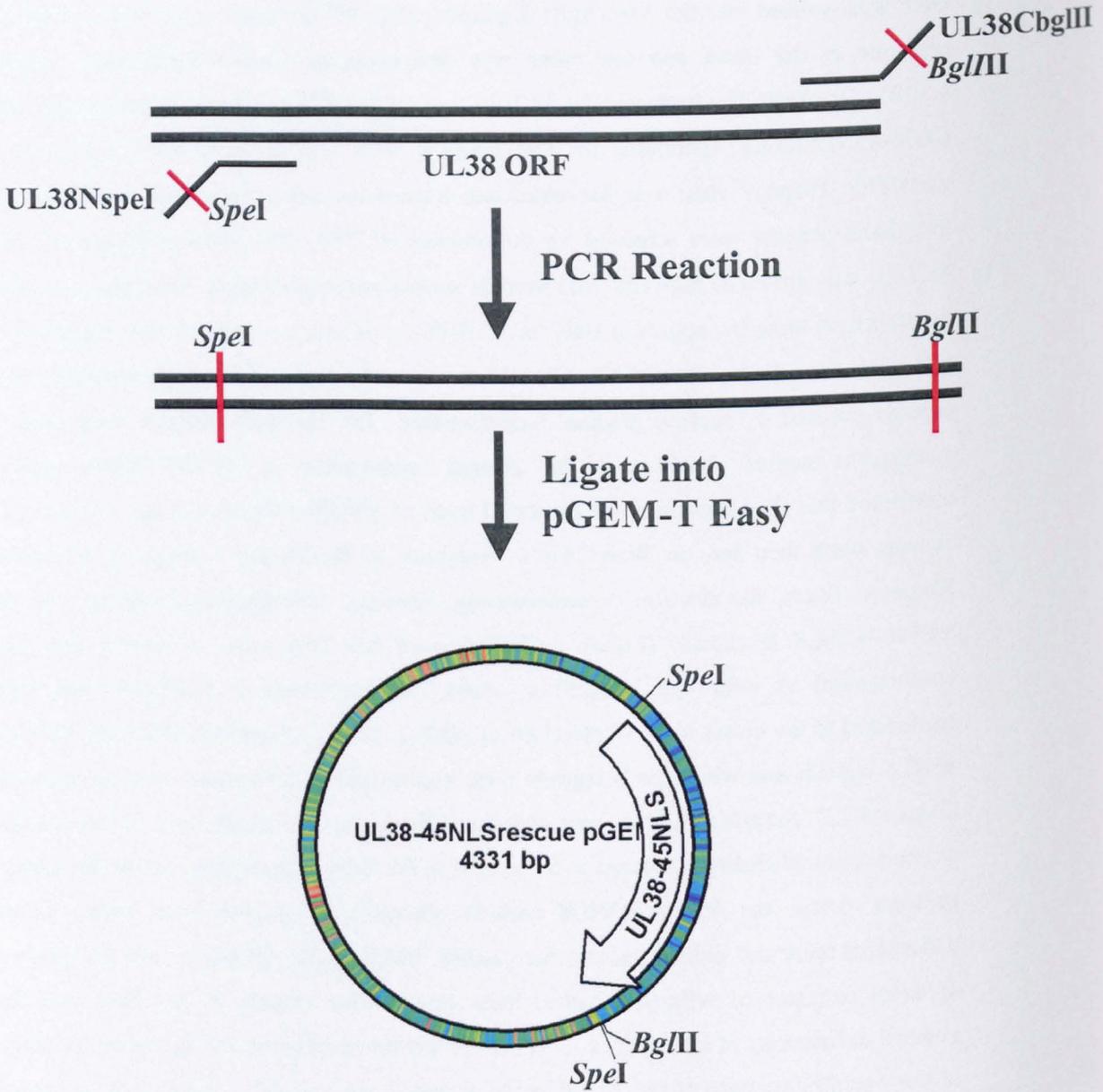
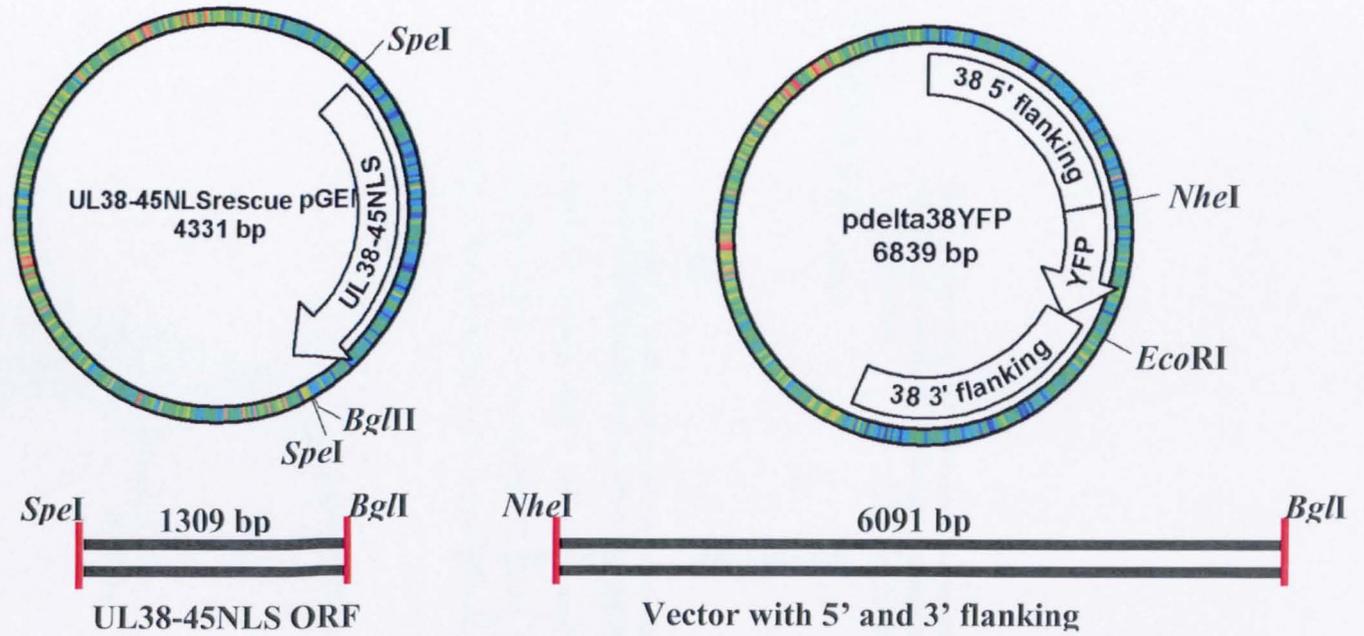


Figure 4.10: Cloning strategy for UL38-45NLSrescue.

A. Cloning strategy to produce UL38-45NLSrescue pGEM. A PCR reaction was carried out using pUL38-45NLS-T Easy as template DNA, UL38NspeI as a forward primer, and UL38CbgIII as a reverse primer. The PCR product was purified and cloned into pGEM-T Easy.

B.



Ligation

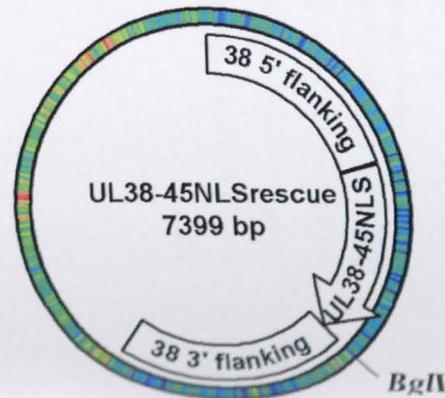
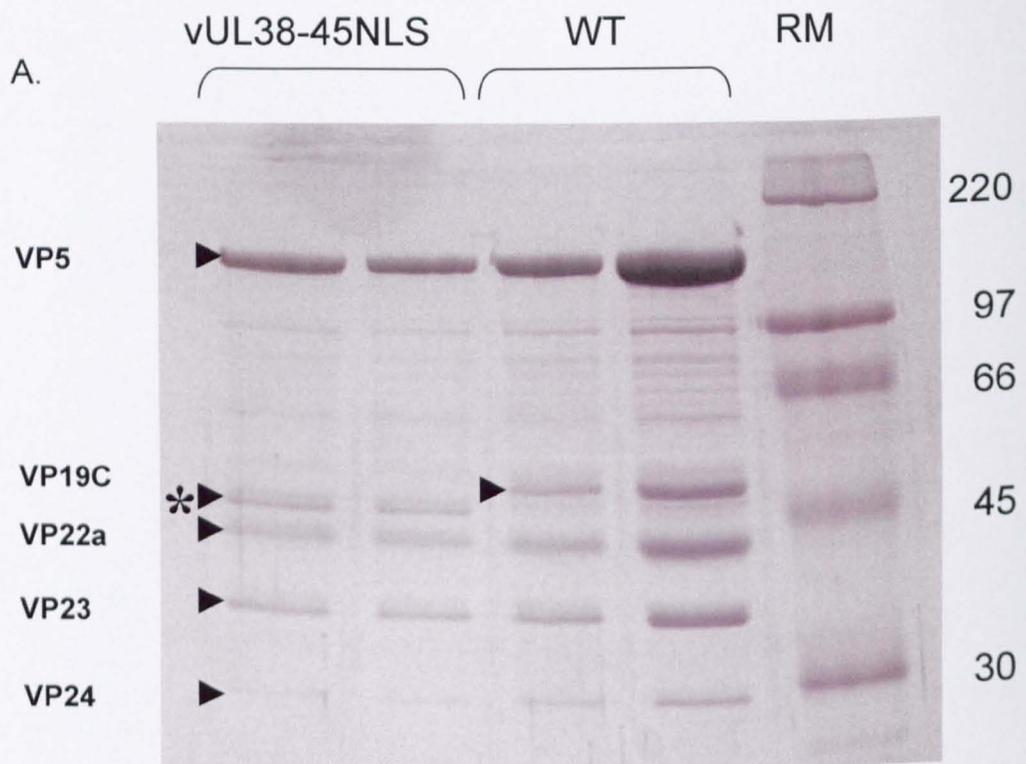
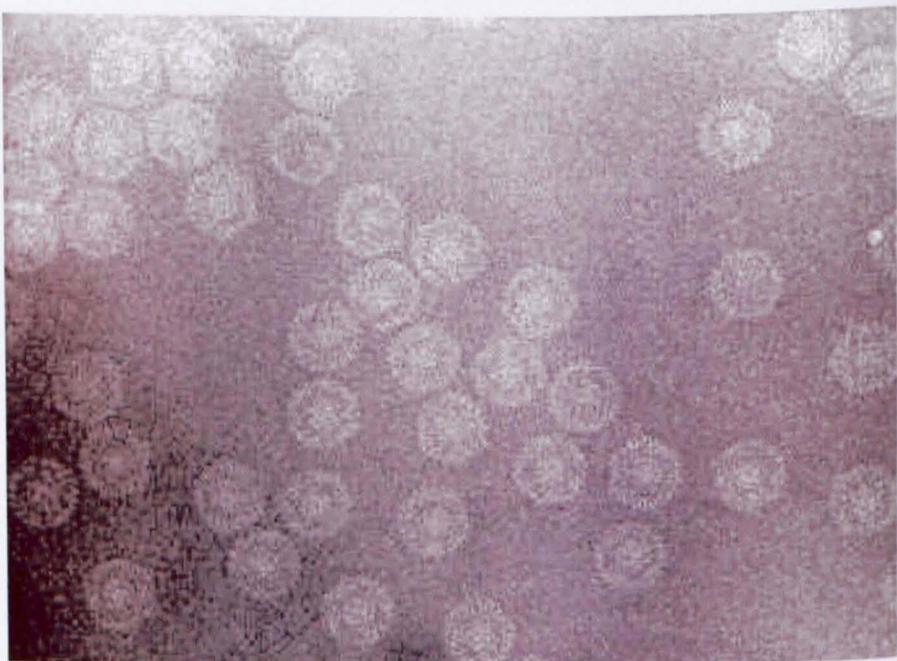


Figure 4.10: Cloning strategy for UL38-45NLSrescue.

**B.** An *SpeI/BglII* digest was carried out on pUL38-45NLSrescue pGEM and a 1309 bp fragment was isolated. An *NheI/EcoRI* digest was carried out on p $\Delta$ 38YFP and a 6091 bp fragment was isolated. These fragments were ligated together to produce pUL38-45NLSrescue.

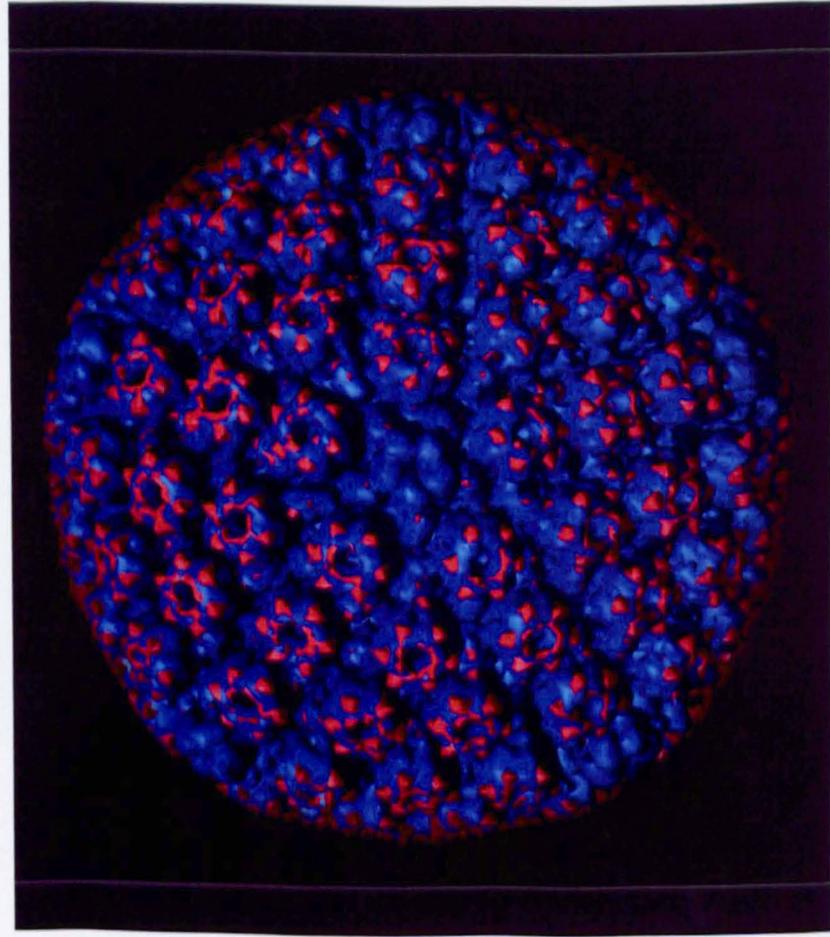


B.

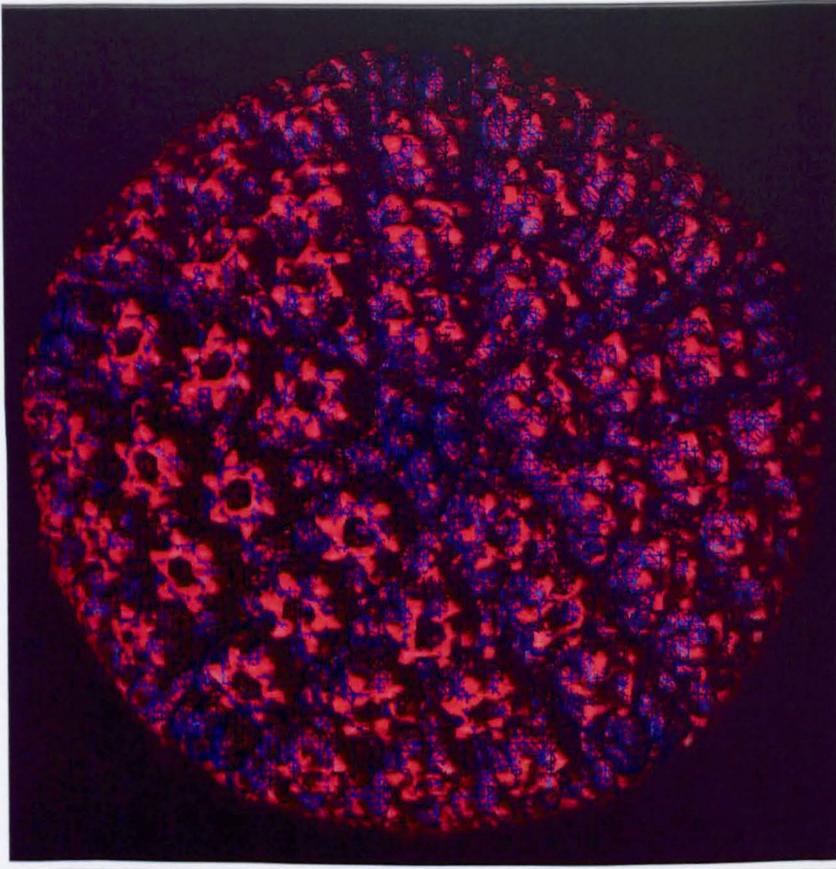


**Figure 4.11:** **A.** Protein composition of truncated VP19C capsids. B-capsids purified from BHK-21 cells infected with wild-type HSV-1, or with vUL38-45NLS, were analysed on a 10% SDS polyacrylamide gel. The positions of the capsid proteins are indicated on the left of the gel. The sizes of the major capsid protein, VP5, the triplex  $\beta$ -subunit, VP23, and the internal scaffolding proteins, VP24 and VP22a, are the same for both viruses. In contrast, VP19C is smaller in the vUL38-45NLS sample (denoted by \*) than in wild-type capsids due to the removal of the N-terminal residues. The sizes of the Rainbow Marker (RM) molecular weight standards are indicated to the right of the gel. **B.** Electron cryomicroscopic image of vUL38-45NLS capsids in vitreous ice.

A.



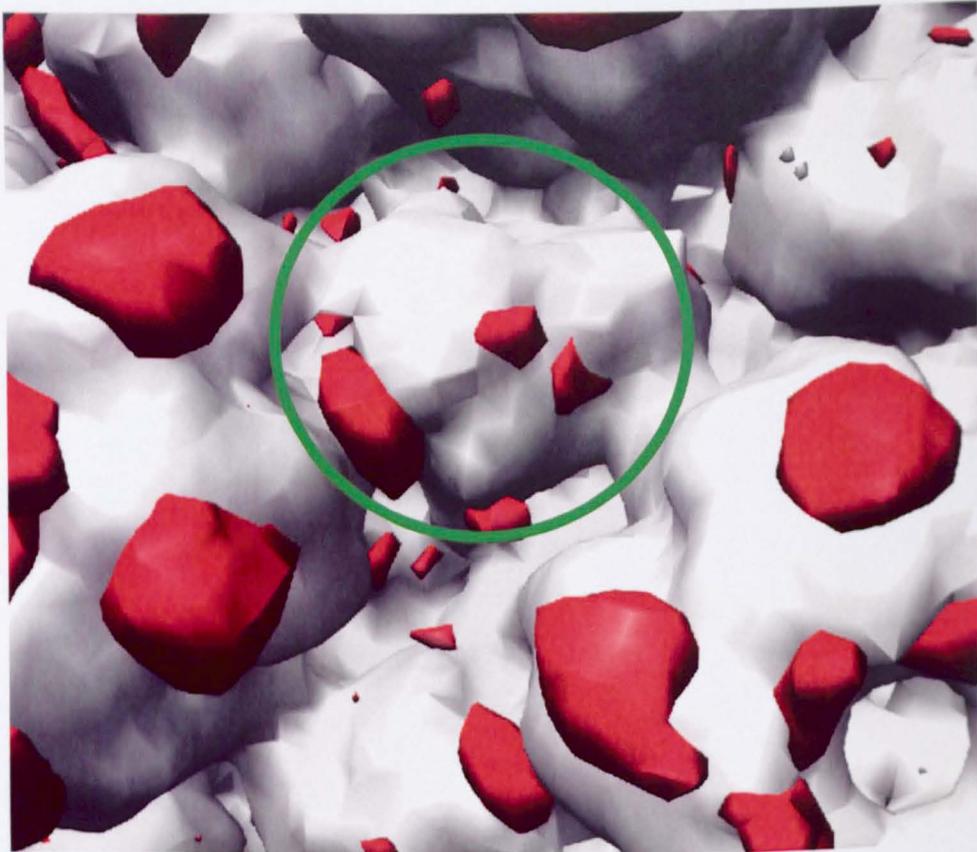
B.



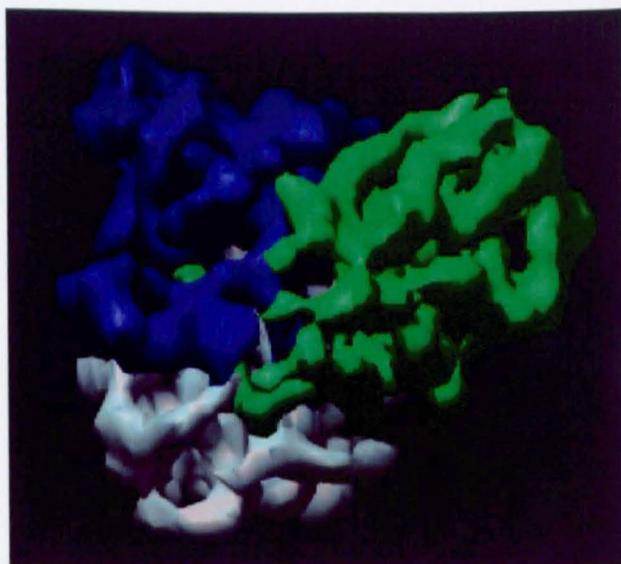
**Figure 4.12:** The UL38-45NLS capsid was reconstructed to an effective resolution of 20 Å and used to generate difference maps with the HSV-1 wild-type B-capsid reconstructed to the same resolution. The image is viewed along a fivefold axis of symmetry. The differences were not uniform across capsids. **A.** Additional mass seen in wild-type capsids is shown in red. The greatest excess of wild-type capsid mass was towards the centres of the faces. **B.** The wild-type capsid is shown as a solid red structure. The UL38-45NLS capsid is shown as a blue mesh. At the vertices, the blue mesh overlies the solid red structure indicating that this is the location of the greatest excess of UL38-45NLS capsid mass.

from the VP19C-45NLS and wild-type B capsid maps were compared. There are several differences in the conformation of the triplexes, but the greatest change was at the top, where the wild-type triplex appeared to have considerable extra mass (Figure 4.13).

A.



B.



**Figure 4.13:** 3D representations of wild-type and VP19C-45NLS triplexes.

**A.** Difference map between wild-type and VP19C-45NLS capsid reconstruction. In this enlarged region from the capsid comparison, a single triplex is shown (encircled). Regions in which the wild-type capsid contains extra mass are shown in red. **B.** 3D representation of a wild-type triplex shown in approximately the same orientation as in A. VP19C is shown in green. The two VP23s are shown in blue and white. This 8 Å resolution image was supplied by Matthew Baker. Segmentation of triplex proteins was carried out by Wen Jiang and Matthew Baker. Molecular graphics images were produced using the UCSF Chimera package from the Computer Graphics Laboratory, University of California, San Francisco (Pettersen *et al.*, 2004).

## **CHAPTER 5**

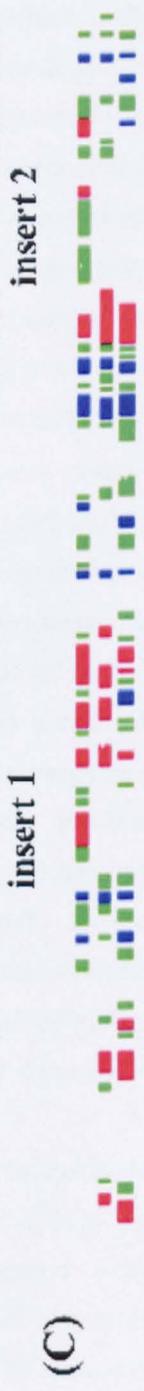
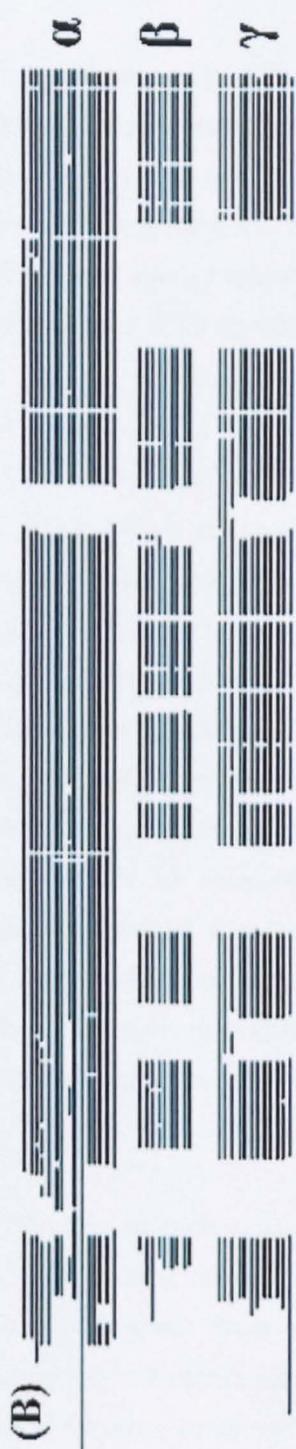
### **MUTATIONAL ANALYSIS OF VP19C – DISCUSSION**

## 5. Mutational analysis of VP19C – Discussion

### 5.1. Computational analysis of VP19C

Members of the three subfamilies of mammalian and avian herpesviruses (*Alpha*, *Beta* and *Gammaherpesvirinae*) share a common capsid architecture and in all examples analysed to date, the triplex is composed of one copy of a VP19C homologue (the  $\alpha$ -subunit) and two copies of a VP23 homologue (the  $\beta$ -subunit) (Gibson *et al.*, 1996; Chen *et al.*, 1999; Yu *et al.*, 2003). It is not known whether capsids from the fish and bivalve herpesviruses families are organised in this way, although CCV capsids contain suitably-sized candidate proteins with appropriate stoichiometries (Booy *et al.*, 1996). In HSV-1 and most other *Alphaherpesvirinae*, the  $\alpha$ -subunit is considerably larger than the  $\beta$ -subunit (465 amino acids compared with 318 in HSV-1). However, in *Beta* and *Gammaherpesvirinae*, the  $\alpha$ -subunit is significantly smaller than HSV-1 VP19C (Gibson *et al.*, 1996; Nealon *et al.*, 2001; O'Connor *et al.*, 2003) and is of similar size to the  $\beta$ -subunit. Since the large size of the  $\alpha$ -subunit in *Alphaherpesvirinae* appears anomalous, it seems likely that the additional sequences represent relatively recent evolutionary developments. Comparison of  $\alpha$ -subunit sequences reveals that the N-terminal 111 amino acids of VP19C are not conserved in the other *Alphaherpesvirinae* apart from HSV-2 (Figure 5.1A). This region is also highly divergent among different herpesvirus subfamilies. However, differences in the N-terminal region are not entirely responsible for the size variation among the subfamilies. An alignment of sequences from Alpha, Beta and Gammaherpesviruses is in broad agreement with earlier findings (Trus *et al.*, 2001) and indicates that the greater size of *Alphaherpesvirinae*  $\alpha$ -subunit proteins is largely accounted for by two internal insertions, each of approximately 40-50 residues (insert 1 and insert 2), as well as by extra sequences at the N-terminus (Figure 5.1B).

Predictions of secondary structure were obtained for the sequences of herpesvirus triplex  $\alpha$ -subunits (Figure 5.1C). The most striking observation was the apparent increase in secondary structure conservation towards the C-terminus. Perhaps indicating the importance of this region in the function of the protein. Also notable was the lack of any evidence of conservation of secondary structure in the N-terminal region of the triplex  $\alpha$ -subunits in *Alphaherpesvirinae*. Indeed, the secondary structure prediction program was unable to detect secondary structural elements in the VP19C N-terminal region to any significant degree of likelihood, with the most N-terminal secondary structural element predicted being the presence of a loop structure beginning at amino acid 102.



**Figure 5.1:** Sequence comparisons of VP19C homologues.

The sequences of fifteen Alphaherpesviruses, nine Betaherpesviruses, and twelve Gammaherpesviruses (identified by a BLAST search from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) using the HSV-1, HCMV, and KSHV herpesvirus triplex  $\alpha$ -subunit sequences as input) were aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). The CLUSTAL W output was used as the basis for subsequent analyses. **(a)** Levels of sequence conservation among the triplex  $\alpha$ -subunits in alphaherpesviruses were analysed and plotted using the JalView (<http://www.jalview.org/>) multiple sequence editor program (Clamp *et al.*, 2004). **(b)** Schematic representation of the consensus alignment between alpha, beta, and gammaherpesvirus triplex  $\alpha$ -subunit sequences, showing the locations of additional sequences (insert 1 and insert 2 comprising residues 140 to 176 and 357 to 405 respectively in HSV-1) in the alphaherpesvirus proteins. The consensus sequence was calculated using the Consensus program (<http://www.bork.embl-heidelberg.de/Alignment/consensus.html>). **(c)** Secondary structure predictions for triplex  $\alpha$ -subunits of (top) alpha, (middle) beta, and (bottom) gammaherpesviruses. Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). Structural elements present in >50% of examples (Figures 5A-G) were plotted relative to the alignments shown in (b) for each subfamily. The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in blue, and coil in green.

The nature of the triplex, as a non-symmetrical structure, occupying an icosahedrally symmetrical position in the capsid, is anomalous, and suggests the possibility that the progenitors of VP19C and VP23 arose as a result of gene duplication from a single trivalent ancestor. Although sequence analysis does not support this conjecture, evidence of structural relationships can endure in circumstances where sequence conservation is no longer apparent, as has been shown for the floor domain of VP5 and the capsid proteins of certain tailed bacteriophages (Baker *et al.*, 2005). However, comparisons of the structures of VP19C and VP23 from secondary structure prediction programs (Figure 5.2A-G), and from currently available 3D reconstructions (Figure 5.3) provide no evidence for relatedness between the two proteins. Thus, either they had independent origins, or the adaptation needed to fulfil their differing functions is sufficient to have removed any traces of a common original structure.

Trus *et al.*, (2001) discussed the possibility of an evolutionary relationship between VP19C and VP23 and were unable to rule it out, citing the alignments that they observed when analysing triplex  $\alpha$ - and  $\beta$ -subunits using Clustal W. An attempt to carry out a similar alignment for VP19C and VP23 using Clustal W produced similar results (Figure 5.2Hi). However, the resulting alignment gave rather poor identity and similarity scores, and contained numerous gaps. In order to test the validity of the alignment, the VP23 amino acid sequence was randomly shuffled using the Sequence Manipulation Suite Shuffle Protein program ([http://bioinformatics.org/sms2/shuffle\\_protein.html](http://bioinformatics.org/sms2/shuffle_protein.html)) (Stothard, 2000) and aligned to VP19C. This produced an alignment of similar quality (Figure 5.2Hii). A slightly less convincing alignment was observed when VP19C was aligned with a random sequence of 318 amino acids (the same length as VP23) (Figure 5.2Hiii) generated using the Sequence Manipulation Suite Random Protein Sequence program ([http://bioinformatics.org/sms2/random\\_protein.html](http://bioinformatics.org/sms2/random_protein.html)) (Stothard, 2000). Therefore, it appears questionable that Clustal W alignments alone provide evidence of an evolutionary relationship between such very poorly related proteins as VP19C and VP23.

## 5.2. VP19C insertional mutagenesis

Transposon-based insertional mutagenesis is a rapid and efficient method of screening a protein sequence for functional domains. The manufacturers claim that the insertions occur randomly. However, the panel of mutants obtained included subsets of insertions that lay in close proximity to each other – such as six insertions between amino acids 234 and 248

HSV-1 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 HSV-2 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 VZV -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 GaHV-2 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 MeHV-1 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 EHV-4 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 EHV-1 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 BoHV-1 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 CeHV-7 -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
 Consensus -----LSTLQVPTVSWGQTVVLEPTVETDQGLLRRLVLRPPIRRKLVLPSSGGE-----RSTLMLGLGQTDAPPGLTENDDTQALDKLRLRMRGG-AALISGRH  
M MM M M M M M M M M M

HSV-1 HLTROVITLTDLCQPNADRQCTALLARFPADLPFAHQRRPFGPCERLGEANQQLMEATALSSRAESGCTRAGLVSNFLVACAAASYDARDAAAVRAHVITANYRTRVGARLDRFSECLR  
 HSV-2 HLTROVITLTDLCQPNADRQCTALLARFPADLPFAHQRRPFGPCERLGEANQQLMEATALSSRAESGCTRAGLVSNFLVACAAASYDARDAAAVRAHVITANYRTRVGARLDRFSECLR  
 VZV QLCQVSLTDFPFDIAEMGSLIILRHPSDMI GEANTLQACRRPDVLLRGLRNLFACTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 GaHV-2 TTSRQVITLDFPCFPAEMGSLIILRHPSDMI GEANTLQACRRPDVLLRGLRNLFACTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 MeHV-1 HLTROVITLDFPCFPAEMGSLIILRHPSDMI GEANTLQACRRPDVLLRGLRNLFACTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 EHV-4 KLTROVITLDFPCFPAEMGSLIILRHPSDMI GEANTLQACRRPDVLLRGLRNLFACTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 EHV-1 KLTROVITLDFPCFPAEMGSLIILRHPSDMI GEANTLQACRRPDVLLRGLRNLFACTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 BoHV-1 HLTROVITLDFPCFPAEMGSLIILRHPSDMI GEANTLQACRRPDVLLRGLRNLFACTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 CeHV-7 DLIKQVSLTDLCEPFLVSPVFLRHCLDIVEDASVCPAGRRDENVEALRALSGCTAIVTVGSGGLRAYVTSLSPIAACRAEYVQAAADANRTAIVSAYGSRMRETLIRPSECLR  
 Consensus HLTROVITLTDLCQPNADRQCTALLARFPADLPFAHQRRPFGPCERLGEANQQLMEATALSSRAESGCTRAGLVSNFLVACAAASYDARDAAAVRAHVITANYRTRVGARLDRFSECLR  
M M MM M MM M M M M M M M M

HSV-1 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 HSV-2 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 VZV AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 GaHV-2 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 MeHV-1 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 EHV-4 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 EHV-1 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 BoHV-1 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 CeHV-7 AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
 Consensus AMVETHVPHVEMRPFGLVSWVTQDELASVTAVCAHQEAAHTGHPGPRSAVILPACAVDLDAELGLGEP--GAFLYLVLYYRQRRLDELCCVYVKSQHPRGLEPALKERLGRLRITN  
MM MM M M M M M M M M M M M M

HSV-1 TIHGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 HSV-2 TIHGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 VZV AAWGTGNTHQPNTRALVQLSDPTSPCSIGBITVNNVNLARQLYQWTDGDFGLTQVLSCHYAAVTLITLTPSESVYTRRMERFGYVNPVITILEGVVWSTIWN-  
 GaHV-2 AITGVDGLVPHGANNMPELGLTLLLRNDEKLG---AAASRDLNKADKWVDLGRTPKESCHYAAVYCRIGLHLS--SSVPVKKFERCCSLDIPVILIPGVVWNC-TWIEC-  
 MeHV-1 TIHGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 EHV-4 AIRGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 EHV-1 AIRGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 BoHV-1 VLRGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 CeHV-7 VMVGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
 Consensus TIHGTEDTTPKPNRNDPFLAGLANNPQTPRCSA---GCVTNPQPADRIYRWQPDLRGRTPARTCTYAAFAELMMPPDSRCLHRTERRFCAITVPVILEGVVWCPS-EWRACA-  
M M M M M M M M M M M M M M

■ = Alpha helix  
 ■ = Sheet  
 ■ = Loop

Figure 5.2A: Secondary structure predictions for alphaherpesvirus triplex  $\alpha$ -subunits .

Sequences were obtained from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) and aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. Structural elements present in >50% of examples were calculated and plotted below the CLUSTAL W alignment. The sequence of the triplex  $\alpha$ -subunit for HSV-1 was used to show the presence of conserved structural elements in the 'consensus' for secondary structure. Positions of insertions in VP19C insertional mutants are shown relative to the alignment and demarcated by a letter M. Non-lethal insertions are shown in green. Lethal insertions are shown in red. There appears to be no correlation between the predicted secondary structure of a region of VP19C the likelihood of an insertion into that region being lethal.

MuHV-2 -----L<sup>X</sup>Y<sup>X</sup>AT<sup>X</sup>SF<sup>X</sup>AK<sup>X</sup>MK<sup>X</sup>EL<sup>X</sup>LD<sup>X</sup>LD<sup>X</sup>ED<sup>X</sup>F-----PQVAK<sup>X</sup>K<sup>X</sup>K<sup>X</sup>IERA<sup>X</sup>IVRS<sup>X</sup>VYV<sup>X</sup>GHAA<sup>X</sup>ITAR<sup>X</sup>FF<sup>X</sup>RET<sup>X</sup>LMEV<sup>X</sup>E<sup>X</sup>Q<sup>X</sup>K<sup>X</sup>GL<sup>X</sup>LM<sup>X</sup>FR<sup>X</sup>LD<sup>X</sup>GI<sup>X</sup>EM<sup>X</sup>PKSV<sup>X</sup>LV<sup>X</sup>S<sup>X</sup>FF<sup>X</sup>YS<sup>X</sup>MA<sup>X</sup>VE<sup>X</sup>VG<sup>X</sup>AAT<sup>X</sup>KAV<sup>X</sup>LR<sup>X</sup>AA<sup>X</sup>Y<sup>X</sup>GE<sup>X</sup>-GEAV<sup>X</sup>LAW<sup>X</sup>I<sup>X</sup>RT<sup>X</sup>GV<sup>X</sup>-NR  
 MuHV-1 -----ANV<sup>X</sup>SS<sup>X</sup>F<sup>X</sup>PMK<sup>X</sup>REV<sup>X</sup>M<sup>X</sup>ED<sup>X</sup>Q<sup>X</sup>ED<sup>X</sup>F-----YKVS<sup>X</sup>Q<sup>X</sup>M<sup>X</sup>RLERS<sup>X</sup>IA<sup>X</sup>KGY<sup>X</sup>VYV<sup>X</sup>GHQA<sup>X</sup>ITAR<sup>X</sup>FF<sup>X</sup>VRES<sup>X</sup>LGE<sup>X</sup>VE<sup>X</sup>Q<sup>X</sup>KN<sup>X</sup>GV<sup>X</sup>LM<sup>X</sup>FR<sup>X</sup>LD<sup>X</sup>GI<sup>X</sup>EM<sup>X</sup>PST<sup>X</sup>VL<sup>X</sup>V<sup>X</sup>SL<sup>X</sup>FF<sup>X</sup>LS<sup>X</sup>M<sup>X</sup>AE<sup>X</sup>IV<sup>X</sup>SA<sup>X</sup>AT<sup>X</sup>KN<sup>X</sup>T<sup>X</sup>LA<sup>X</sup>AI<sup>X</sup>Y<sup>X</sup>GE<sup>X</sup>-GEA<sup>X</sup>IRT<sup>X</sup>WR<sup>X</sup>LD<sup>X</sup>GA<sup>X</sup>-NR  
 CaHV-1 -----ATKR<sup>X</sup>-SYGQ<sup>X</sup>IT-----GRIT<sup>X</sup>F-----E<sup>X</sup>REAK<sup>X</sup>GR<sup>X</sup>KMER<sup>X</sup>APAK<sup>X</sup>-YIYG<sup>X</sup>GH<sup>X</sup>EA<sup>X</sup>IA<sup>X</sup>IR<sup>X</sup>FF<sup>X</sup>Q<sup>X</sup>EGL<sup>X</sup>LEI<sup>X</sup>ET<sup>X</sup>V<sup>X</sup>H<sup>X</sup>Y<sup>X</sup>Q<sup>X</sup>CL<sup>X</sup>ML<sup>X</sup>RL<sup>X</sup>DT<sup>X</sup>SA<sup>X</sup>ES<sup>X</sup>PR<sup>X</sup>Y<sup>X</sup>VM<sup>X</sup>IS<sup>X</sup>LF<sup>X</sup>LA<sup>X</sup>I<sup>X</sup>RAAN<sup>X</sup>VC<sup>X</sup>SA<sup>X</sup>VR<sup>X</sup>TAL<sup>X</sup>SA<sup>X</sup>LY<sup>X</sup>TS<sup>X</sup>A<sup>X</sup>PG<sup>X</sup>AV<sup>X</sup>SS<sup>X</sup>LE<sup>X</sup>PSI<sup>X</sup>-A<sup>X</sup>  
 TuHV-1 L<sup>X</sup>KA<sup>X</sup>VAA<sup>X</sup>EAD<sup>X</sup>AA<sup>X</sup>VA<sup>X</sup>TT<sup>X</sup>VP<sup>X</sup>SA<sup>X</sup>RV<sup>X</sup>TRAT<sup>X</sup>AT<sup>X</sup>ED<sup>X</sup>GP<sup>X</sup>VP<sup>X</sup>TD<sup>X</sup>ED<sup>X</sup>L<sup>X</sup>ET<sup>X</sup>LR<sup>X</sup>AA<sup>X</sup>K<sup>X</sup>R<sup>X</sup>FK<sup>X</sup>IT<sup>X</sup>SG<sup>X</sup>VAP<sup>X</sup>-PVY<sup>X</sup>HAD<sup>X</sup>DA<sup>X</sup>Y<sup>X</sup>TR<sup>X</sup>FF<sup>X</sup>Q<sup>X</sup>EGL<sup>X</sup>LE<sup>X</sup>HE<sup>X</sup>AE<sup>X</sup>PG<sup>X</sup>ALL<sup>X</sup>FR<sup>X</sup>LD<sup>X</sup>GI<sup>X</sup>EM<sup>X</sup>PR<sup>X</sup>LL<sup>X</sup>IT<sup>X</sup>LS<sup>X</sup>W<sup>X</sup>PS<sup>X</sup>LAL<sup>X</sup>AA<sup>X</sup>IV<sup>X</sup>SA<sup>X</sup>AT<sup>X</sup>KN<sup>X</sup>T<sup>X</sup>LA<sup>X</sup>AI<sup>X</sup>Y<sup>X</sup>Q<sup>X</sup>PA<sup>X</sup>M<sup>X</sup>NA<sup>X</sup>AS<sup>X</sup>AW<sup>X</sup>LD<sup>X</sup>FA<sup>X</sup>R<sup>X</sup>-PR  
 HCMV L<sup>X</sup>KAR<sup>X</sup>VA<sup>X</sup>AK<sup>X</sup>PR<sup>X</sup>LD<sup>X</sup>PA<sup>X</sup>ED<sup>X</sup>N<sup>X</sup>EL<sup>X</sup>V<sup>X</sup>L<sup>X</sup>TALK<sup>X</sup>AK<sup>X</sup>REV<sup>X</sup>TI<sup>X</sup>SV<sup>X</sup>RY<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>GHQ<sup>X</sup>AL<sup>X</sup>TAR<sup>X</sup>FF<sup>X</sup>PE<sup>X</sup>GL<sup>X</sup>VE<sup>X</sup>FEA<sup>X</sup>Q<sup>X</sup>-PG<sup>X</sup>ALL<sup>X</sup>IR<sup>X</sup>ME<sup>X</sup>T<sup>X</sup>GC<sup>X</sup>DS<sup>X</sup>PR<sup>X</sup>HL<sup>X</sup>Y<sup>X</sup>IS<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>LL<sup>X</sup>GI<sup>X</sup>RAS<sup>X</sup>IV<sup>X</sup>SA<sup>X</sup>AT<sup>X</sup>KN<sup>X</sup>T<sup>X</sup>LA<sup>X</sup>AI<sup>X</sup>Y<sup>X</sup>Q<sup>X</sup>PA<sup>X</sup>M<sup>X</sup>NA<sup>X</sup>AS<sup>X</sup>AW<sup>X</sup>LD<sup>X</sup>FA<sup>X</sup>R<sup>X</sup>-HL  
 PaHV-2 L<sup>X</sup>QART<sup>X</sup>LAK<sup>X</sup>PR<sup>X</sup>LD<sup>X</sup>PA<sup>X</sup>ED<sup>X</sup>N<sup>X</sup>EL<sup>X</sup>V<sup>X</sup>L<sup>X</sup>TALK<sup>X</sup>AK<sup>X</sup>REV<sup>X</sup>TI<sup>X</sup>SV<sup>X</sup>RY<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>GHQ<sup>X</sup>AL<sup>X</sup>TAR<sup>X</sup>FF<sup>X</sup>PE<sup>X</sup>GL<sup>X</sup>VE<sup>X</sup>FEA<sup>X</sup>Q<sup>X</sup>-PG<sup>X</sup>SL<sup>X</sup>IL<sup>X</sup>R<sup>X</sup>ME<sup>X</sup>T<sup>X</sup>GC<sup>X</sup>DS<sup>X</sup>PR<sup>X</sup>HL<sup>X</sup>Y<sup>X</sup>IS<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>LL<sup>X</sup>GI<sup>X</sup>RAS<sup>X</sup>IV<sup>X</sup>SA<sup>X</sup>AT<sup>X</sup>KN<sup>X</sup>T<sup>X</sup>LA<sup>X</sup>AI<sup>X</sup>Y<sup>X</sup>Q<sup>X</sup>PA<sup>X</sup>M<sup>X</sup>NA<sup>X</sup>AS<sup>X</sup>AW<sup>X</sup>LD<sup>X</sup>FA<sup>X</sup>R<sup>X</sup>-WL  
 RhCMV L<sup>X</sup>KARE<sup>X</sup>IA<sup>X</sup>KR<sup>X</sup>PR<sup>X</sup>LD<sup>X</sup>PA<sup>X</sup>ED<sup>X</sup>N<sup>X</sup>EL<sup>X</sup>V<sup>X</sup>L<sup>X</sup>AV<sup>X</sup>KL<sup>X</sup>K<sup>X</sup>REV<sup>X</sup>NT<sup>X</sup>LAV<sup>X</sup>RY<sup>X</sup>LP<sup>X</sup>AE<sup>X</sup>HC<sup>X</sup>IT<sup>X</sup>SR<sup>X</sup>FF<sup>X</sup>PE<sup>X</sup>GL<sup>X</sup>VE<sup>X</sup>FEV<sup>X</sup>-PG<sup>X</sup>SL<sup>X</sup>IF<sup>X</sup>R<sup>X</sup>ME<sup>X</sup>T<sup>X</sup>GC<sup>X</sup>DS<sup>X</sup>PR<sup>X</sup>Y<sup>X</sup>LI<sup>X</sup>SI<sup>X</sup>Y<sup>X</sup>LM<sup>X</sup>AI<sup>X</sup>KAT<sup>X</sup>HS<sup>X</sup>MA<sup>X</sup>TK<sup>X</sup>Q<sup>X</sup>LE<sup>X</sup>GI<sup>X</sup>Y<sup>X</sup>TS<sup>X</sup>PA<sup>X</sup>KA<sup>X</sup>AL<sup>X</sup>AW<sup>X</sup>LD<sup>X</sup>FA<sup>X</sup>R<sup>X</sup>-KV  
 HHV-7 -----NEM<sup>X</sup>LQ<sup>X</sup>RK<sup>X</sup>HQ<sup>X</sup>TI<sup>X</sup>ETS<sup>X</sup>LN-----NL<sup>X</sup>IS<sup>X</sup>NE<sup>X</sup>K<sup>X</sup>K<sup>X</sup>IK<sup>X</sup>SL<sup>X</sup>LA<sup>X</sup>FR<sup>X</sup>VI<sup>X</sup>KE<sup>X</sup>CV<sup>X</sup>LS<sup>X</sup>RS<sup>X</sup>FT<sup>X</sup>PC<sup>X</sup>ELL<sup>X</sup>GF<sup>X</sup>RD<sup>X</sup>H<sup>X</sup>G<sup>X</sup>ML<sup>X</sup>PR<sup>X</sup>LE<sup>X</sup>DA<sup>X</sup>BS<sup>X</sup>PK<sup>X</sup>FL<sup>X</sup>IS<sup>X</sup>VL<sup>X</sup>PL<sup>X</sup>AM<sup>X</sup>AF<sup>X</sup>IV<sup>X</sup>SP<sup>X</sup>CTR<sup>X</sup>IS<sup>X</sup>LAN<sup>X</sup>LY<sup>X</sup>K<sup>X</sup>TS<sup>X</sup>IV<sup>X</sup>DS<sup>X</sup>INT<sup>X</sup>MI<sup>X</sup>DC<sup>X</sup>CS<sup>X</sup>--Y  
 HHV-6 G<sup>X</sup>SK<sup>X</sup>S<sup>X</sup>ARA<sup>X</sup>AI<sup>X</sup>VD<sup>X</sup>T<sup>X</sup>VE<sup>X</sup>AV<sup>X</sup>K<sup>X</sup>RR<sup>X</sup>KY<sup>X</sup>IS<sup>X</sup>AG<sup>X</sup>T<sup>X</sup>IN-----N<sup>X</sup>IV<sup>X</sup>E<sup>X</sup>K<sup>X</sup>ER<sup>X</sup>K<sup>X</sup>PL<sup>X</sup>Q<sup>X</sup>FL<sup>X</sup>S<sup>X</sup>-GREN<sup>X</sup>LR<sup>X</sup>IA<sup>X</sup>AR<sup>X</sup>V<sup>X</sup>TP<sup>X</sup>CELL<sup>X</sup>AP<sup>X</sup>ELEN<sup>X</sup>GL<sup>X</sup>M<sup>X</sup>Y<sup>X</sup>RF<sup>X</sup>ED<sup>X</sup>VD<sup>X</sup>N<sup>X</sup>PK<sup>X</sup>IL<sup>X</sup>FV<sup>X</sup>GL<sup>X</sup>FP<sup>X</sup>CS<sup>X</sup>NA<sup>X</sup>IV<sup>X</sup>SA<sup>X</sup>CV<sup>X</sup>RT<sup>X</sup>ALT<sup>X</sup>TM<sup>X</sup>Y<sup>X</sup>NS<sup>X</sup>M<sup>X</sup>VD<sup>X</sup>N<sup>X</sup>VL<sup>X</sup>SM<sup>X</sup>INT<sup>X</sup>CK<sup>X</sup>--Y  
 Consensus -----MD<sup>X</sup>AR<sup>X</sup>VA<sup>X</sup>KR<sup>X</sup>PR<sup>X</sup>PD<sup>X</sup>AD<sup>X</sup>EN<sup>X</sup>EL<sup>X</sup>V-----L<sup>X</sup>TALK<sup>X</sup>AK<sup>X</sup>REV<sup>X</sup>NT<sup>X</sup>IS<sup>X</sup>VR<sup>X</sup>Y<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>GHQ<sup>X</sup>AL<sup>X</sup>TAR<sup>X</sup>FF<sup>X</sup>PE<sup>X</sup>GL<sup>X</sup>VE<sup>X</sup>FEA<sup>X</sup>Q<sup>X</sup>-PG<sup>X</sup>ALL<sup>X</sup>IR<sup>X</sup>ME<sup>X</sup>T<sup>X</sup>GC<sup>X</sup>DS<sup>X</sup>PR<sup>X</sup>HL<sup>X</sup>Y<sup>X</sup>IS<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>LL<sup>X</sup>GI<sup>X</sup>RAS<sup>X</sup>IV<sup>X</sup>SA<sup>X</sup>AT<sup>X</sup>KN<sup>X</sup>T<sup>X</sup>LA<sup>X</sup>AI<sup>X</sup>Y<sup>X</sup>Q<sup>X</sup>PA<sup>X</sup>M<sup>X</sup>NA<sup>X</sup>AS<sup>X</sup>AW<sup>X</sup>LD<sup>X</sup>FA<sup>X</sup>R<sup>X</sup>-HL

MuHV-2 L<sup>X</sup>HO<sup>X</sup>I<sup>X</sup>I<sup>X</sup>HL<sup>X</sup>PG<sup>X</sup>CT<sup>X</sup>NS<sup>X</sup>IT<sup>X</sup>ST<sup>X</sup>CL<sup>X</sup>I<sup>X</sup>TC<sup>X</sup>SM<sup>X</sup>LGH<sup>X</sup>SY<sup>X</sup>N<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>PF<sup>X</sup>V<sup>X</sup>PK<sup>X</sup>E<sup>X</sup>LY<sup>X</sup>LD<sup>X</sup>DE<sup>X</sup>FT<sup>X</sup>SEE<sup>X</sup>---L<sup>X</sup>RF<sup>X</sup>AY<sup>X</sup>P<sup>X</sup>V<sup>X</sup>I<sup>X</sup>T<sup>X</sup>Y<sup>X</sup>DY<sup>X</sup>NS<sup>X</sup>R<sup>X</sup>Q<sup>X</sup>SR<sup>X</sup>SA<sup>X</sup>F<sup>X</sup>VA<sup>X</sup>TR<sup>X</sup>VS<sup>X</sup>PH<sup>X</sup>L<sup>X</sup>LD<sup>X</sup>AL<sup>X</sup>RY<sup>X</sup>R<sup>X</sup>FR<sup>X</sup>AS<sup>X</sup>R<sup>X</sup>FF<sup>X</sup>LN<sup>X</sup>TV<sup>X</sup>SG<sup>X</sup>Y<sup>X</sup>QS<sup>X</sup>-AG<sup>X</sup>CL<sup>X</sup>GT<sup>X</sup>L<sup>X</sup>Q<sup>X</sup>RL<sup>X</sup>GW<sup>X</sup>FC<sup>X</sup>-S<sup>X</sup>DA<sup>X</sup>RS<sup>X</sup>GV<sup>X</sup>VAC  
 MuHV-1 L<sup>X</sup>HR<sup>X</sup>V<sup>X</sup>VH<sup>X</sup>PL<sup>X</sup>GC<sup>X</sup>TS<sup>X</sup>NS<sup>X</sup>IT<sup>X</sup>ST<sup>X</sup>CL<sup>X</sup>I<sup>X</sup>TC<sup>X</sup>SM<sup>X</sup>RG<sup>X</sup>HS<sup>X</sup>YN<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>PL<sup>X</sup>VP<sup>X</sup>KE<sup>X</sup>TY<sup>X</sup>LD<sup>X</sup>DE<sup>X</sup>FT<sup>X</sup>SEE<sup>X</sup>---L<sup>X</sup>RF<sup>X</sup>AY<sup>X</sup>P<sup>X</sup>V<sup>X</sup>I<sup>X</sup>T<sup>X</sup>Y<sup>X</sup>DY<sup>X</sup>NS<sup>X</sup>R<sup>X</sup>Q<sup>X</sup>SR<sup>X</sup>SA<sup>X</sup>F<sup>X</sup>V<sup>X</sup>VS<sup>X</sup>R<sup>X</sup>ITH<sup>X</sup>R<sup>X</sup>HT<sup>X</sup>L<sup>X</sup>IN<sup>X</sup>VL<sup>X</sup>RY<sup>X</sup>R<sup>X</sup>FR<sup>X</sup>VS<sup>X</sup>R<sup>X</sup>FF<sup>X</sup>LN<sup>X</sup>NS<sup>X</sup>I<sup>X</sup>SG<sup>X</sup>Y<sup>X</sup>GP<sup>X</sup>-AG<sup>X</sup>CL<sup>X</sup>ET<sup>X</sup>L<sup>X</sup>Q<sup>X</sup>RL<sup>X</sup>GW<sup>X</sup>FC<sup>X</sup>-S<sup>X</sup>DS<sup>X</sup>RS<sup>X</sup>GI<sup>X</sup>VAS  
 CaHV-1 L<sup>X</sup>HK<sup>X</sup>V<sup>X</sup>VS<sup>X</sup>AV<sup>X</sup>GC<sup>X</sup>M<sup>X</sup>NS<sup>X</sup>V<sup>X</sup>SP<sup>X</sup>GI<sup>X</sup>TS<sup>X</sup>L<sup>X</sup>TC<sup>X</sup>VM<sup>X</sup>RG<sup>X</sup>HAY<sup>X</sup>N<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>PF<sup>X</sup>PK<sup>X</sup>DM<sup>X</sup>V<sup>X</sup>KL<sup>X</sup>D<sup>X</sup>Y<sup>X</sup>GH<sup>X</sup>SE<sup>X</sup>-PL<sup>X</sup>YAL<sup>X</sup>Y<sup>X</sup>AY<sup>X</sup>L<sup>X</sup>V<sup>X</sup>IV<sup>X</sup>N<sup>X</sup>HAG<sup>X</sup>-EP<sup>X</sup>PN<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>IV<sup>X</sup>VS<sup>X</sup>SV<sup>X</sup>R<sup>X</sup>NR<sup>X</sup>SL<sup>X</sup>IN<sup>X</sup>LL<sup>X</sup>QR<sup>X</sup>Q<sup>X</sup>RV<sup>X</sup>TER<sup>X</sup>FL<sup>X</sup>Y<sup>X</sup>LD<sup>X</sup>R<sup>X</sup>K<sup>X</sup>IV<sup>X</sup>EP<sup>X</sup>ND<sup>X</sup>A<sup>X</sup>-SAC<sup>X</sup>CV<sup>X</sup>V<sup>X</sup>Q<sup>X</sup>RL<sup>X</sup>GW<sup>X</sup>FC<sup>X</sup>ES<sup>X</sup>GI<sup>X</sup>L<sup>X</sup>KT<sup>X</sup>GV<sup>X</sup>LG<sup>X</sup>Y  
 TuHV-1 L<sup>X</sup>HEAL<sup>X</sup>Q<sup>X</sup>PV<sup>X</sup>SLQ<sup>X</sup>AV<sup>X</sup>AP<sup>X</sup>GL<sup>X</sup>TS<sup>X</sup>L<sup>X</sup>VT<sup>X</sup>CV<sup>X</sup>M<sup>X</sup>Q<sup>X</sup>GN<sup>X</sup>V<sup>X</sup>N<sup>X</sup>DL<sup>X</sup>K<sup>X</sup>Q<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>PL<sup>X</sup>L<sup>X</sup>V<sup>X</sup>L<sup>X</sup>K<sup>X</sup>D<sup>X</sup>Y<sup>X</sup>LL<sup>X</sup>DE<sup>X</sup>RR<sup>X</sup>EG<sup>X</sup>---VM<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>WL<sup>X</sup>I<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>PK<sup>X</sup>E<sup>X</sup>-HP<sup>X</sup>SP<sup>X</sup>FF<sup>X</sup>V<sup>X</sup>GT<sup>X</sup>AV<sup>X</sup>RR<sup>X</sup>KT<sup>X</sup>L<sup>X</sup>ALL<sup>X</sup>RR<sup>X</sup>RF<sup>X</sup>AS<sup>X</sup>AV<sup>X</sup>TV<sup>X</sup>LER<sup>X</sup>QL<sup>X</sup>R<sup>X</sup>HAL<sup>X</sup>GG<sup>X</sup>Y<sup>X</sup>CA<sup>X</sup>AV<sup>X</sup>V<sup>X</sup>Q<sup>X</sup>RL<sup>X</sup>GW<sup>X</sup>FC<sup>X</sup>-EG<sup>X</sup>VT<sup>X</sup>AG<sup>X</sup>VL<sup>X</sup>TR  
 HCMV L<sup>X</sup>HR<sup>X</sup>RE<sup>X</sup>LET<sup>X</sup>GL<sup>X</sup>CV<sup>X</sup>K<sup>X</sup>TV<sup>X</sup>SL<sup>X</sup>IT<sup>X</sup>SL<sup>X</sup>LL<sup>X</sup>TC<sup>X</sup>VM<sup>X</sup>RG<sup>X</sup>LY<sup>X</sup>NT<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>V<sup>X</sup>FAL<sup>X</sup>MI<sup>X</sup>PK<sup>X</sup>DM<sup>X</sup>Y<sup>X</sup>L<sup>X</sup>TWE<sup>X</sup>-ES<sup>X</sup>KGR<sup>X</sup>---L<sup>X</sup>Q<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>LI<sup>X</sup>V<sup>X</sup>I<sup>X</sup>DY<sup>X</sup>DE<sup>X</sup>GE<sup>X</sup>ET<sup>X</sup>PG<sup>X</sup>I<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>VT<sup>X</sup>SS<sup>X</sup>IA<sup>X</sup>HW<sup>X</sup>Q<sup>X</sup>TL<sup>X</sup>VD<sup>X</sup>V<sup>X</sup>ARG<sup>X</sup>K<sup>X</sup>FA<sup>X</sup>RE<sup>X</sup>RC<sup>X</sup>S<sup>X</sup>F<sup>X</sup>V<sup>X</sup>NR<sup>X</sup>IT<sup>X</sup>EP<sup>X</sup>Q<sup>X</sup>I<sup>X</sup>-PL<sup>X</sup>CT<sup>X</sup>GV<sup>X</sup>I<sup>X</sup>Q<sup>X</sup>KL<sup>X</sup>GW<sup>X</sup>CL<sup>X</sup>-AD<sup>X</sup>I<sup>X</sup>HT<sup>X</sup>S<sup>X</sup>FL<sup>X</sup>VH  
 PaHV-2 L<sup>X</sup>HR<sup>X</sup>RE<sup>X</sup>LET<sup>X</sup>VG<sup>X</sup>CV<sup>X</sup>K<sup>X</sup>AV<sup>X</sup>SL<sup>X</sup>MT<sup>X</sup>SL<sup>X</sup>LL<sup>X</sup>TC<sup>X</sup>VM<sup>X</sup>RG<sup>X</sup>LY<sup>X</sup>NT<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>V<sup>X</sup>FAL<sup>X</sup>MI<sup>X</sup>PK<sup>X</sup>DM<sup>X</sup>Y<sup>X</sup>L<sup>X</sup>TWE<sup>X</sup>-ES<sup>X</sup>KGR<sup>X</sup>---L<sup>X</sup>Q<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>LI<sup>X</sup>V<sup>X</sup>I<sup>X</sup>DY<sup>X</sup>DE<sup>X</sup>GE<sup>X</sup>ET<sup>X</sup>PG<sup>X</sup>I<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>VT<sup>X</sup>SS<sup>X</sup>IA<sup>X</sup>HW<sup>X</sup>Q<sup>X</sup>TL<sup>X</sup>VD<sup>X</sup>V<sup>X</sup>ARG<sup>X</sup>K<sup>X</sup>FA<sup>X</sup>RE<sup>X</sup>RC<sup>X</sup>S<sup>X</sup>F<sup>X</sup>V<sup>X</sup>NR<sup>X</sup>IT<sup>X</sup>EP<sup>X</sup>Q<sup>X</sup>I<sup>X</sup>-PL<sup>X</sup>CT<sup>X</sup>GV<sup>X</sup>I<sup>X</sup>Q<sup>X</sup>KL<sup>X</sup>GW<sup>X</sup>CL<sup>X</sup>-AD<sup>X</sup>I<sup>X</sup>HT<sup>X</sup>S<sup>X</sup>FL<sup>X</sup>VH  
 RhCMV L<sup>X</sup>HR<sup>X</sup>RL<sup>X</sup>RT<sup>X</sup>IG<sup>X</sup>CV<sup>X</sup>K<sup>X</sup>VS<sup>X</sup>LG<sup>X</sup>IT<sup>X</sup>SL<sup>X</sup>LL<sup>X</sup>TC<sup>X</sup>VL<sup>X</sup>SG<sup>X</sup>Y<sup>X</sup>L<sup>X</sup>Y<sup>X</sup>NT<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>I<sup>X</sup>FS<sup>X</sup>WL<sup>X</sup>I<sup>X</sup>PK<sup>X</sup>DM<sup>X</sup>Y<sup>X</sup>TL<sup>X</sup>TWE<sup>X</sup>-ES<sup>X</sup>KGR<sup>X</sup>---L<sup>X</sup>Q<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>LI<sup>X</sup>V<sup>X</sup>I<sup>X</sup>DY<sup>X</sup>DE<sup>X</sup>GE<sup>X</sup>ET<sup>X</sup>PG<sup>X</sup>I<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>VT<sup>X</sup>SS<sup>X</sup>IA<sup>X</sup>HW<sup>X</sup>Q<sup>X</sup>TL<sup>X</sup>VD<sup>X</sup>V<sup>X</sup>ARG<sup>X</sup>K<sup>X</sup>FA<sup>X</sup>RE<sup>X</sup>RC<sup>X</sup>S<sup>X</sup>F<sup>X</sup>V<sup>X</sup>NR<sup>X</sup>IT<sup>X</sup>EP<sup>X</sup>Q<sup>X</sup>I<sup>X</sup>-PL<sup>X</sup>CT<sup>X</sup>GV<sup>X</sup>I<sup>X</sup>Q<sup>X</sup>KL<sup>X</sup>GW<sup>X</sup>CL<sup>X</sup>-AD<sup>X</sup>I<sup>X</sup>HT<sup>X</sup>SV<sup>X</sup>LT<sup>X</sup>H  
 HHV-7 L<sup>X</sup>E<sup>X</sup>K<sup>X</sup>VS<sup>X</sup>L<sup>X</sup>FG<sup>X</sup>VT<sup>X</sup>NY<sup>X</sup>IS<sup>X</sup>ET<sup>X</sup>NC<sup>X</sup>L<sup>X</sup>ISC<sup>X</sup>MQ<sup>X</sup>GH<sup>X</sup>V<sup>X</sup>D<sup>X</sup>VR<sup>X</sup>KEN<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>GL<sup>X</sup>IL<sup>X</sup>K<sup>X</sup>D<sup>X</sup>LL<sup>X</sup>LE<sup>X</sup>TP<sup>X</sup>EN<sup>X</sup>---AI<sup>X</sup>Q<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>L<sup>X</sup>CY<sup>X</sup>I<sup>X</sup>Y<sup>X</sup>RO<sup>X</sup>EN<sup>X</sup>-K<sup>X</sup>IEY<sup>X</sup>G<sup>X</sup>Y<sup>X</sup>I<sup>X</sup>V<sup>X</sup>LS<sup>X</sup>ET<sup>X</sup>HE<sup>X</sup>EV<sup>X</sup>L<sup>X</sup>IDL<sup>X</sup>LL<sup>X</sup>EN<sup>X</sup>K<sup>X</sup>F<sup>X</sup>SK<sup>X</sup>ER<sup>X</sup>F<sup>X</sup>MF<sup>X</sup>NY<sup>X</sup>L<sup>X</sup>IN<sup>X</sup>Q<sup>X</sup>EN<sup>X</sup>---NY<sup>X</sup>FG<sup>X</sup>SL<sup>X</sup>Q<sup>X</sup>RI<sup>X</sup>G<sup>X</sup>HC<sup>X</sup>CK<sup>X</sup>-T<sup>X</sup>MI<sup>X</sup>K<sup>X</sup>SG<sup>X</sup>IL<sup>X</sup>DL  
 HHV-6 L<sup>X</sup>ET<sup>X</sup>K<sup>X</sup>VS<sup>X</sup>L<sup>X</sup>FG<sup>X</sup>VT<sup>X</sup>SL<sup>X</sup>V<sup>X</sup>SC<sup>X</sup>SS<sup>X</sup>CL<sup>X</sup>LC<sup>X</sup>V<sup>X</sup>Q<sup>X</sup>GN<sup>X</sup>V<sup>X</sup>D<sup>X</sup>ANKEN<sup>X</sup>I<sup>X</sup>H<sup>X</sup>GL<sup>X</sup>TV<sup>X</sup>L<sup>X</sup>KE<sup>X</sup>I<sup>X</sup>FL<sup>X</sup>EP<sup>X</sup>EN<sup>X</sup>---ST<sup>X</sup>Q<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>V<sup>X</sup>V<sup>X</sup>H<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>KE<sup>X</sup>V<sup>X</sup>LS<sup>X</sup>-AL<sup>X</sup>Q<sup>X</sup>Y<sup>X</sup>GI<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>VT<sup>X</sup>LS<sup>X</sup>ET<sup>X</sup>HE<sup>X</sup>EV<sup>X</sup>L<sup>X</sup>VD<sup>X</sup>IL<sup>X</sup>R<sup>X</sup>Q<sup>X</sup>Y<sup>X</sup>E<sup>X</sup>KER<sup>X</sup>FL<sup>X</sup>PL<sup>X</sup>NY<sup>X</sup>LI<sup>X</sup>EN<sup>X</sup>TL<sup>X</sup>---SY<sup>X</sup>FG<sup>X</sup>SV<sup>X</sup>Q<sup>X</sup>RI<sup>X</sup>GR<sup>X</sup>CA<sup>X</sup>-T<sup>X</sup>MI<sup>X</sup>K<sup>X</sup>SG<sup>X</sup>FL<sup>X</sup>Q<sup>X</sup>Y  
 Consensus L<sup>X</sup>HR<sup>X</sup>RE<sup>X</sup>LET<sup>X</sup>GL<sup>X</sup>CV<sup>X</sup>K<sup>X</sup>TV<sup>X</sup>SL<sup>X</sup>IT<sup>X</sup>SL<sup>X</sup>LL<sup>X</sup>TC<sup>X</sup>VM<sup>X</sup>RG<sup>X</sup>LY<sup>X</sup>NT<sup>X</sup>L<sup>X</sup>K<sup>X</sup>TE<sup>X</sup>V<sup>X</sup>FAL<sup>X</sup>MI<sup>X</sup>PK<sup>X</sup>DM<sup>X</sup>Y<sup>X</sup>L<sup>X</sup>TWE<sup>X</sup>-ES<sup>X</sup>KGR<sup>X</sup>---L<sup>X</sup>Q<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>LI<sup>X</sup>V<sup>X</sup>I<sup>X</sup>DY<sup>X</sup>DE<sup>X</sup>GE<sup>X</sup>ET<sup>X</sup>PG<sup>X</sup>I<sup>X</sup>V<sup>X</sup>Y<sup>X</sup>VT<sup>X</sup>SS<sup>X</sup>IA<sup>X</sup>HW<sup>X</sup>Q<sup>X</sup>TL<sup>X</sup>VD<sup>X</sup>V<sup>X</sup>ARG<sup>X</sup>K<sup>X</sup>FA<sup>X</sup>RE<sup>X</sup>RC<sup>X</sup>S<sup>X</sup>F<sup>X</sup>V<sup>X</sup>NR<sup>X</sup>IT<sup>X</sup>EP<sup>X</sup>Q<sup>X</sup>I<sup>X</sup>-PL<sup>X</sup>CT<sup>X</sup>GV<sup>X</sup>I<sup>X</sup>Q<sup>X</sup>KL<sup>X</sup>GW<sup>X</sup>CL<sup>X</sup>-AD<sup>X</sup>I<sup>X</sup>HT<sup>X</sup>S<sup>X</sup>FL<sup>X</sup>VH

MuHV-2 RGGQLP<sup>X</sup>V<sup>X</sup>VR<sup>X</sup>LEK<sup>X</sup>FY<sup>X</sup>VD<sup>X</sup>L<sup>X</sup>PL<sup>X</sup>VE<sup>X</sup>F<sup>X</sup>  
 MuHV-1 RAGQLS<sup>X</sup>V<sup>X</sup>VR<sup>X</sup>LEK<sup>X</sup>FY<sup>X</sup>VD<sup>X</sup>L<sup>X</sup>PL<sup>X</sup>VE<sup>X</sup>F<sup>X</sup>  
 CaHV-1 KSAQIP<sup>X</sup>V<sup>X</sup>VR<sup>X</sup>LEK<sup>X</sup>FY<sup>X</sup>TD<sup>X</sup>ST<sup>X</sup>MF<sup>X</sup>EF<sup>X</sup>  
 TuHV-1 KAVQLS<sup>X</sup>V<sup>X</sup>FR<sup>X</sup>LEK<sup>X</sup>FY<sup>X</sup>VD<sup>X</sup>L<sup>X</sup>SPACE<sup>X</sup>F<sup>X</sup>  
 HCMV KELKLS<sup>X</sup>V<sup>X</sup>VR<sup>X</sup>LD<sup>X</sup>N<sup>X</sup>FS<sup>X</sup>VEL<sup>X</sup>DF<sup>X</sup>PRE<sup>X</sup>F<sup>X</sup>  
 PaHV-2 KETKLS<sup>X</sup>V<sup>X</sup>FR<sup>X</sup>LD<sup>X</sup>N<sup>X</sup>FC<sup>X</sup>VE<sup>X</sup>I<sup>X</sup>EP<sup>X</sup>RE<sup>X</sup>F<sup>X</sup>  
 RhCHV QEVKLS<sup>X</sup>V<sup>X</sup>LE<sup>X</sup>HF<sup>X</sup>N<sup>X</sup>VE<sup>X</sup>EP<sup>X</sup>TE<sup>X</sup>F<sup>X</sup>  
 HHV-7 QGISL<sup>X</sup>T<sup>X</sup>IR<sup>X</sup>L<sup>X</sup>K<sup>X</sup>DV<sup>X</sup>F<sup>X</sup>VEL<sup>X</sup>ER<sup>X</sup>K<sup>X</sup>IF<sup>X</sup>  
 HHV-6 RGITL<sup>X</sup>P<sup>X</sup>VI<sup>X</sup>K<sup>X</sup>LE<sup>X</sup>NI<sup>X</sup>F<sup>X</sup>VD<sup>X</sup>LE<sup>X</sup>KK<sup>X</sup>V<sup>X</sup>F<sup>X</sup>  
 Consensus KELKLS<sup>X</sup>V<sup>X</sup>VR<sup>X</sup>LD<sup>X</sup>N<sup>X</sup>FS<sup>X</sup>VEL<sup>X</sup>DF<sup>X</sup>PRE<sup>X</sup>F<sup>X</sup>

X = Alpha helix  
 Y = Sheet  
 Z = Loop

**Figure 5.2B:** Secondary structure predictions for betaherpesvirus triplex  $\alpha$ -subunits .

Sequences were obtained from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) and aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. Structural elements present in >50% of examples were calculated and plotted below the CLUSTAL W alignment. The sequence of the triplex  $\alpha$ -subunit for HCMV was used to show the presence of conserved structural elements in the 'consensus' for secondary structure.

AtHV-3 -----MPTLNENAVYKIERDYLKLEPPRPHKINLEEGPKFQKSIKQALSKYVTV-SVNPINFLSNQMLDVE-----FRQSYFADFLIQAKTIHDK  
 SaHV-2 -----MPTLNENAVYKIERDYLKLEPPRPHKIEISGGQKFAKSIKQALSKYVTV-SVNPIDPFPSNQMLDVE-----FRQELYADFLIQAKTIHDK  
 BoHV-4 -----MPTLNENAVYKIERDYLKLEPPRPHKINLEEGPKFQKSIKQALSKYVTV-SITFGLDMFRQDVLLEV-----FRQELYGEFLVGSSTLBNK  
 EHV-2 -----MPTLNENAVYKIERDYLKLEPPRPHKINLEEGPKFQKSIKQALSKYVTV-SITFGLDMFRQDVLLEV-----FRQELYGEFLVGHAKTFHPC  
 KSHV -----MKVQENAAARLGRQVGLGLPEPPHRSVLETPPEFARGVDRLLSKYAA-STPTFTVGSLSHHEALRQAE-----FRQPTYGDFLVYSQTFEPQ  
 McHV-5 -----MKTRDANVHKLNDISMRLLDPPPHRVLSRGRDFKGVDRLLSKYVTV-STPTFTVGSLSHHEALRQAE-----FRQPTYGDFLVYSQTFMSSC  
 HHV-4 -----MKVQGV-SVDKRRIRRIAGLPPPEARRLNISSSEPTFDVDRVGLVEHAQASSTAAAVRAGLLAPGEVAVAAGSGGGGSFSSGSGGPPVFPVDFLIIHASSFNNA  
 CeHV-15 -----MKVQGV-SVDKRRIRRIAGLPPPEARRLNISSCAEFARDVDRVGLVEHAQASSTAAAVRAGLLAPGEVAVAAGSGGGGSFSSGSGGPPVFPVDFLIIHASSFNNA  
 CaHV-3 -----MKVVSAINKVRVHRQVGLGLPFAARRLNITDKDFVEDVDRVGLIDDMSSSLIGAFSIIKKL-----VTDLLEPR-----SVYSDFLIQASSFNNA  
 AlHV-1 -----MPTLNENAVYKIERDYLKLEPPRPHKINLEEGPKFQKSIKQALSKYVTV-SVNPINFLSNQMLDVE-----FRQELYADFLIQAKTIHDK  
 MuHV-4 -----MPTLNENAVYKIERDYLKLEPPRPHKINLEEGPKFQKSIKQALSKYVTV-SITFGLDMFRQDVLLEV-----FRQELYGEFLVGSSTLBNK  
 Consensus -----MKVQENAAARLGRQVGLGLPEPPHRSVLETPPEFARGVDRLLSKYAA-STPTFTVGSLSHHEALRQAE-----FRQPTYGDFLVYSQTFSEPQ

AtHV-3 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENINSTFITRIKSVAKGEVLNKKSPPENIKVLPSDLFLDDESEITD  
 SaHV-2 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 BoHV-4 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 EHV-2 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 KSHV -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 McHV-5 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 HHV-4 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 CeHV-15 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 CaHV-3 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 AlHV-1 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 MuHV-4 -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD  
 Consensus -----GTPAFAFAPSVHRISPVWYGGNCIEDTVPNLAELIEEGRPHKFLTVGGLPVLENISTFTVTRIKSVVKGVELNKKSPPENIKVLPSDLFLDDESEITD

AtHV-3 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIDMKQDFYINKMTFGALCKVGSSSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 SaHV-2 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 BoHV-4 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 EHV-2 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 KSHV -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 McHV-5 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 HHV-4 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 CeHV-15 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 CaHV-3 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 AlHV-1 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 MuHV-4 -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL  
 Consensus -----LAPKPHLLSFFYACVTVYF-LHNVPVMTLTFIKENKSRDIPCHLKVIVYDVRTKYIAMQDFYINKMSFGALCKLGSSEITVE RRGNYHFRGSSLPIVEVPDFVADPGPWMTIL

X = Alpha helix  
 Y = Sheet  
 Z = Loop

**Figure 5.2C:** Secondary structure predictions for gammaherpesvirus triplex  $\alpha$ -subunits .

Sequences were obtained from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) and aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. Structural elements present in >50% of examples were calculated and plotted below the CLUSTAL W alignment. The sequence of the triplex  $\alpha$ -subunit for KSHV was used to show the presence of conserved structural elements in the 'consensus' for secondary structure.

GaHV-2 **NS**SMITVVEILFPLKSLTIDANLLIQKCEKRVLFLPFVRRVRLKDVVDKSYFYIAGTEPDITLGLLSLTKTRFAAVITRALPGRMSAVVLGMCSTPNG-LALQNTGPFDDLNGTIVCLMPPIFFVCCRIRLESDELLFPVTVPTRLANE  
 MeHV-1 **NS**SMITVVEILFPLKSLTIDANLLIQKCEKRVLFLPFVRRVRLKDVVDKSYFYIAGTEPDITLGLLSLTKTRFAAVITRALPGRMSAVVLGMCSTPNG-LALQNTGPFDDLNGTIVCLMPPIFFVCCRIRLESDELLFPVTVPTRLANE  
 CeHV-7 ---NSFTEIILLPQGLTFRHETSNLQKCEKRIITLTVLRRHRASLSEIATSSYYVAGAPDVLISLEAYRVRPFAVITQVLPKGLMVAWVGVCFNAPG-LFIQNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 VZV ---NSMFEIRVLLPAGELSPARTSALQKCEBKIITPSTLRRHRASLVDIALSSYYIAGAPDITLISLEAYRMRPFAVITRVLPGKLLAHAGVSTFPEG-LFIQNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 EHV-4 ---NSMFAFIDITLIPGDLSPADLSALQKCEKRVLFLPTALRRRVMLSSVTLASYYVAGAPDITLISMAAPRRRFPATITQVLPKMKIAAALGVAVLPEG-TFIQNTSPFDLTNGDSVICALPPILDV-EDKLRKLSNGEELFPLTTPLAQARE  
 EHV-1 ---NSMFAFIDITLIPGDLSPADLSALQKCEKRVLFLPTALRRRVMLSSVTLASYYVAGAPDITLISMAAPRRRFPATITQVLPKMKIAAALGVAVLPEG-TFIQNTSPFDLTNGDSVICALPPILDV-EDKLRKLSNGEELFPLTTPLAQARE  
 BoHV-1 **NS**SMITVVEILFPLKSLTIDANLLIQKCEKRVLFLPFVRRVRLKDVVDKSYFYIAGTEPDITLGLLSLTKTRFAAVITRALPGRMSAVVLGMCSTPNG-LALQNTGPFDDLNGTIVCLMPPIFFVCCRIRLESDELLFPVTVPTRLANE  
 PRV ---NSVDIALIPGDLSPADLSALQKCEKRVVLETLRRHATLREVALP----CGSDVLAAMARRRFAAVITRVTFRMLATPLGVSGRQS-LVQNTSPFDLTNGDSVICALPPILDV-EDKLRKLSNGEELFPLTTPLAQARE  
 HSV-2 ---NSITCFRADIAIPSGISRPDAALQKCEKRVVLPPTIRRQLADVAHESFVSGVSPDITLGLLILAYRRRFPVITRVLTRIVACPVLDLTHAGTVNLRNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 HSV-1 ---NSLADGFTDIAIPSGISRPDAALQKCEKRVVLPPTIRRQLADVAHESFVSGVSPDITLGLLILAYRRRFPVITRVLTRIVACPVLDLTHAGTVNLRNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 CeHV-1 ---NSLADGFTDIAIPSGISRPDAALQKCEKRVVLPPTIRRQLADVAHESFVSGVSPDITLGLLILAYRRRFPVITRVLTRIVACPVLDLTHAGTVNLRNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 SaHV-1 ---NSLADGFTDIAIPSGISRPDAALQKCEKRVVLPPTIRRQLADVAHESFVSGVSPDITLGLLILAYRRRFPVITRVLTRIVACPVLDLTHAGTVNLRNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 GaHV-1 ---NSLADGFTDIAIPSGISRPDAALQKCEKRVVLPPTIRRQLADVAHESFVSGVSPDITLGLLILAYRRRFPVITRVLTRIVACPVLDLTHAGTVNLRNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD  
 Consensus **NS**SMITVVEILFPLKSLTIDANLLIQKCEKRVVLPPTIRRQLADVAHESFVSGVSPDITLGLLILAYRRRFPVITRVLTRIVACPVLDLTHAGTVNLRNTSPVDLCKGDIYICLLPPIRGR-VDCIKLDSKLELVFPLTTFQQLSRD

GaHV-2 **IA**KTILGRALEALIGR--GMPFPTSRESEITMYNGRSYITISITL-HLDDAAESTVTRLLLNHFAINECEMILYTIMIPTLLTLGASDGYNVALVGLLETATRAVQGLIIFINPFLQAWR--RFPVYEALSAWITMTLNLGNVLSLHLLK  
 MeHV-1 **IA**KTILGRALEALIGR--GMPFPTSRESEITMYNGRSYITISITL-HLDDAAESTVTRLLLNHFAINECEMILYTIMIPTLLTLGASDGYNVALVGLLETATRAVQGLIIFINPFLQAWR--RFPVYEALSAWITMTLNLGNVLSLHLLK  
 CeHV-7 **IA**KTIVARSLETIAIGS-AEHPASRGPETVTYKSYITITRSLDAAEBAARTLLAMFVTEGNCMILYTIMIPTLLTLGASDGYNVALVGLLETATRAVQGLIIFINPFLQAWR--RFPVYEALSSWINMATHLGDILSRSLLR  
 VZV **IA**KTIVARAVEITAGIQIPHEVLRGIDVICYNRRYELTETN--CHRDGSAARTITLVLLMFSINEGCLLLALIPITLLVQAGHDGYNLLIQANCVRETTQGLINIPFMPRIQDHR--RFPVYETISSWITSSRLGDTLGFRAILR  
 EHV-4 **IA**RLVARAVCALIPN----AQQHGADVMFYNGRKYNTDF-RHRDANGVARSIVNMI FAMNEGSLVLLSIPNLLTLGTDGDFVNAIIQIGSATREVGQQLIHQQFVPPQDGAR--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 EHV-1 **IA**RLVARAVCALIPN----AQQHQAEVMFYNGRKYNTDF-RHRDANGVARSIVNMI FAMNEGSLVLLSIPNLLTLGTDGDFVNAIIQIGSATREVGQQLIHQQFVPPQDGAR--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 BoHV-1 **IA**RVVVARAVEALGDE---GARGALADVMYHNRERYQVTPDV-LCRGADAAARTITLVLMNPNVNEGSLVLLSIPNLLTLGTDGDFVNAITVQLGASREAGQLLLEPAEPQDGR--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 PRV **IA**RVVVARAVEALGDE---GARGADVFNSGRRYQLDF--HRINAQAARTISLVLMNPIILNEGAVILLSLIPNLLTLGTDGDFVNAITVQLGASREAGQLLLEPAEPQDGR--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 HSV-2 **IA**RVLVARGIRDLNPS--PRTGELPDLNLVLYNRSRLSLLAV--QQLGSPVNAELRSVLMNVYSITEGTTIITLIPRLLEALSADGYNVALLQMSVTRAAQLIHPAPALMGDGER--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 HSV-1 **IA**RVLVARGIRDLNPS--PRTGELPDLNLVLYNRSRLSLLAV--QQLGSPVNAELRSVLMNVYSITEGTTIITLIPRLLEALSADGYNVALLQMSVTRAAQLIHPAPALMGDGER--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 CeHV-1 **IA**RVLVARGIRDLNPS--PRTGELPDLNLVLYNRSRLSLLAV--QQLGSPVNAELRSVLMNVYSITEGTTIITLIPRLLEALSADGYNVALLQMSVTRAAQLIHPAPALMGDGER--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 SaHV-1 **IA**RVLVARGIRDLNPS--PRTGELPDLNLVLYNRSRLSLLAV--QQLGSPVNAELRSVLMNVYSITEGTTIITLIPRLLEALSADGYNVALLQMSVTRAAQLIHPAPALMGDGER--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 GaHV-1 **IA**RVLVARGIRDLNPS--PRTGELPDLNLVLYNRSRLSLLAV--QQLGSPVNAELRSVLMNVYSITEGTTIITLIPRLLEALSADGYNVALLQMSVTRAAQLIHPAPALMGDGER--RFPVYDALMSHIGVARSRLGDVVEGKPLVR  
 Consensus **IA**RVLVARGIRDLNPS--PRTGELPDLNLVLYNRSRLSLLAV--QQLGSPVNAELRSVLMNVYSITEGTTIITLIPRLLEALSADGYNVALLQMSVTRAAQLIHPAPALMGDGER--RFPVYDALMSHIGVARSRLGDVVEGKPLVR

GaHV-2 **VCTFDG**AVNIGSGLDLCPIANW  
 MeHV-1 **ACTFDG**AVNVVSDMCPVISNW  
 CeHV-7 **ICTFDG**AVTVRPGERISVIQ--  
 VZV **VCVFDG**AVTVHPQDRTAVIQ--  
 EHV-4 **ICTFEG**AVTIRGEGKAPVIQTL  
 EHV-1 **ICTFEG**AVTIRGEGKAPVIQTL  
 BoHV-1 **VCTFDG**AVSVVRVSGKPIVVP--  
 PRV **VCFDGA**AVTVFEGKAAVVE--  
 HSV-2 **VCTFDG**AVVCGDMAPVIRI--  
 HSV-1 **VCTFDG**AVVSGDMAPVIRI--  
 CeHV-1 **VCTFDG**AVTVRSGDLAPVIRI--  
 SaHV-1 **VCTFDG**AVIKSELAPVIRI--  
 GaHV-1 **VVVPYFDG**AVTIRGEGDLNAIVE--  
 Consensus **VCTFDG**AVVSGDMAPVIRI--

■ = Alpha helix  
 ■ = Sheet  
 ■ = Loop

Figure 5.2D: Secondary structure predictions for alphaherpesvirus triplex  $\beta$ -subunits .

Sequences were obtained from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) and aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. Structural elements present in >50% of examples were calculated and plotted below the CLUSTAL W alignment. The sequence of the triplex  $\beta$ -subunit in HSV-1 was used to show the presence of conserved structural elements in the 'consensus' for secondary structure.

HCMV **MA**MEANT**FC**T**PH**IG**LS**I**AD**VG**KL**T**LV**AV**PI**P**RL**HL**L**IKHY**QL**GL**HQ**F**V**D**HT**-**RC****V**RL**R**GL**L**R**N**MT**L****FL**RR**V**EG**N**Q**IL**LE**VP**PH**GL**LY**VL**NT**GP**VT**WE**K**GL**AL**CV**LP**PL**PH**GL**LARE**NL**TL**Q**W**EL**VL**PW**I**VP****PL**A**LE**IN**QR**I  
 PaHV-2 **MA**ME**ST**I**FC**T**PH**KL**LI**AD**L**GL**KL**AN**LT**AV**VP**IP**RL**HL**L**IK**HL**QL**GL**H**Q**Y**T**DE**T**-**RC****V**PR**NR**SL**LR**N**MT**L**TL**RR**V**EG**N**Q**IL**LE**VP**PH**GL**LY**VL**NT**GP**VT**WE**K**GL**AL**CV**LP**PL**PH**GL**MARE**NL**TL**Q**W**EL**VL**PW**I**VP****PL**A**Q**EIN**QR**I  
 CeHV-8 ---**MD**ST**I**FC**T**PH**KL**LI**AD**VG**KL**AN**LT**SA**IP**I**RR**HL**L**IK**HH**HL**GL**H**Q**Y**ID**ST**-RC****V**AR**LR**SL**LR**N**MT**L**TL**RR**V**EG**N**Q**IL**LE**VP**PH**GL**LY**VL**NT**GP**VT**WE**K**GL**AL**CV**LP**PL**PH**GL**SV**LR**HP**VS**GN**WN**VL**PW**IL**EP****PL**A**VE**IN**QR**I  
 TuHV-1 **MA**AD**TT**V**CT**PH**KL**LI**AD**VG**KL**AN**LT**SA**IP**I**RR**HL**L**IK**H**IGH**T**AI**GL**HR**LL**DE**T**-**RA**AR**LR**TL**LR**N**MT**L**TL**RR**V**EG**N**Q**IL**LE**VP**PH**GL**LY**VL**NT**GP**VT**WE**K**GL**AL**CV**LP**PL**PH**GL**PER**AV**RI**GN**WS**VL**PW**IL**EP**PL**A**VE**IN**QR**I  
 MuHV-2 ---**KE**TT**VL**CT**PH**KL**LI**AD**V**SK**L**AK**MI**CA**IV**PI**PR**HH**L**IG**S**AQ**V**GL**NA**I**IN**DE**QR**D**Y**AR**VR**TR**FR**DM**TY**TV**MR**Q**VE**GN**Q**ML**GL**E**Y**GR**VT**Y**K**NT**SH**AL**WE**GD**VL**TAL**AP**V**S**--**TS**AA**LV**TV**GN**W**Q**VL**PW**I**VP****PL**A**TE**IN**QR**I  
 HHV-6 ---**KE**S**-**Y**CT**PH**KL**LI**SD**I**ST**L**CK**LM**NI**V**IP**PA**HH**HL**IG**SN**L**GL**Y**P**VS**SN**-K**D**Y**V**HR**VR**LR**TM**V**VT**IL**Q**K**VE**GN**Q**LV**L**K**PV**PH**Q**Y**AI**K**NT**GP**FP**WE**K**GL**TL**IL**IP**PL**ST**--**H**SE**E**KL**L**K**L**Q**W**EL**TV**PL**VP**PL**A**AE**IN**IR**L  
 HHV-7 ---**MS**-**I**Y**CT**PH**KL**LI**SD**I**ST**L**CK**LM**NI**V**IP**PR**HH**L**IG**NT**N**L**GL**Y**TV**LS**TI**-**TD**Y**I**Q**IR**DL**IK**TM**V**LT**IL**Q**K**VE**GN**Q**LV**L**K**PV**PH**Q**Y**AI**K**NT**GP**FP**LE**K**GL**SL**LP**PF**IK**--**PS**Q**-**L**LV**LP**FW**EL**L**PL**LI**EP**TV**ATE**IN**VR**M**  
 PCMV ---**KE**N**-**I**CT**PH**KL**LI**SD**V**SK**L**C**NA**IG**TV**IP**VS**HH**L**IN**NY**IG**L**S**Q**I**F**ST**I**-K**D**Y**L**Q**IK**EA**L**R**K**M**N**L**T**IL**R**N**VE**GN**Q**LV**L**K**PV**PH**Q**Y**AI**K**NT**GP**FP**LE**K**GL**TL**IL**IP**PS**SA**--**P**CT**PK**IM**SL**Q**W**EL**VL**PL**TV**SE**VATE**VN**LR**L**  
 Consensus **MA**ME**AN**T**FC**T**PH**KL**LI**AD**V**G**KL**T**LV**AV**PI**P**RL**HL**L**IKHY**QL**GL**HQ**F**V**D**HT**-**RC****V**RL**R**GL**L**R**N**MT**L****FL**RR**V**EG**N**Q**IL**LE**VP**PH**GL**LY**VL**NT**GP**VT**WE**K**GL**AL**CV**LP**PL**PH**GL**LARE**NL**TL**Q**W**EL**VL**PW**I**VP****PL**A**LE**IN**QR**I

HCMV **L**IM**GL**PS**L**DR**S**Y**E**Y**K**AA**V**Q**Q**L**Q**T**I**FR**AT**FT**IP**P**V**ID**Q**H**L**L**I**D**M**K**T**AC**L**S**M**S**M**VA**N**L**A**SE**L**MT**Y**VR**K**L**A**LE**D**S**M**L**L**V**K**Q**E**LL**M**K**L**DR**R**ERS**-**V**G**E**PR**TP**----**AR**PH**VS**P**DE**I**AR**L**S**A**L**P**V**M**L**R**Q**L**DD**L**IR**Q**W**EL**VL**PW**I**VP****PL**A**VE**IN**QR**I  
 PaHV-2 **L**IM**GL**PS**L**DR**S**Y**E**EV**K**AA**V**H**Q**L**K**T**Y**R**AT**FT**IP**P**V**ID**NY**L**L**DM**K**T**AC**L**S**M**S**M**V**AN**L**AS**L**IV**TV**VR**K**L**A**LE**D**SS**M**L**L**V**K**Q**E**LL**M**K**L**DR**R**ERS**-**V**G**ER**VP**----**PR**PH**VS**P**DE**I**S**R**L**S**A**L**P**V**M**L**R**Q**L**DD**L**IR**Q**W**EL**VL**PW**I**VP****PL**A**VE**IN**QR**I  
 CeHV-8 **L**IM**A**Y**SL**EE**VE**V**K**AA**V**HL**Q**HT**FR**D**AT**I**PD**V**EN**HL**LD**AK**T**AC**L**S**M**S**M**VA**N**L**A**S**L**IV**TV**VR**K**L**A**LE**D**SS**M**L**L**V**K**Q**E**LL**M**K**L**DR**R**ER**D**AG**T**GG**MS**NR**LL**R**-**GR**PS**VS**P**DE**I**S**R**L**S**A**L**P**V**M**L**R**Q**L**DD**I**HE**VM**PT**VC**D**V**SP**  
 TuHV-1 **L**IV**AL**LS**LE**Y**EQ**V**RA**AT**A**Q**L**R**T**HY**R**AT**FT**SL**PD**V**Q**LE**PT**LE**DEL**K**T**VC**L**S**M**S**M**V**T**L**A**SE**L**IV**TV**VR**K**L**A**LE**D**HS**M**L**L**L**K**Q**E**LL**A**RR**T**AA**AG**RR**E**--**DD**AAA**AA**AV**EH**L**IP**DE**V**AK**L**T**AP**FM**VR**Q**L**DD**V**VT**EP**VP**Q**VD**V**AS  
 MuHV-2 **L**L**I**ALL**SL**RS**R**SE**EV**RA**AS**A**Q**L**R**T**Y**VR**AT**L**LE**PE**L**SL**ND**TT**LD**V**LR**NV**CT**LS**M**IA**N**L**AS**LV**TV**VR**K**L**A**LE**D**SS**M**L**L**M**K**Q**E**LL**TR**Q**DL**PR**CG**N**AG**H**GD**SL**AA**GR**PN**MT**PE**EL**N**KL**T**AL**LV**MI**R**Q**L**SD**IV**A**Q**HL**F**V**CD**V**IS**  
 HHV-6 **L**C**I**GL**I**AV**HR**Y**EM**Q**T**I**D**EL**CS**Q**Y**R**D**VL**KL**FP**AV**ND**K**SM**Y**SM**K**T**AC**LS**M**IT**AM**AP**D**IV**RT**Y**ID**RL**LE**D**HS**M**L**L**I**K**Q**E**L**L**S**ERT**IL**MT**QR**SN**-----**Q**L**HA**FE**I**K**DEL**KK**V**K**S**V**LT**M**ID**Q**IN**SL**T**N**E**K**Y**FP**V**CD**V**S**  
 HHV-7 **L**C**I**S**LL**S**I**HR**RY**EV**Q**I**I**D**EL**R**T**Q**Y**R**VT**IK**L**EV**I**D**AK**ST**FM**K**T**AC**I**SP**EM**I**AT**MA**P**D**IV**Q**TY**IER**L**S**LE**CS**M**L**L**I**K**Q**E**LL**L**AK**N**-**PS**Q**PS**-----**EP**TA**E**IK**TE**L**Q**K**I**K**T**V**T**M**IN**Q**I**NS**L**T**Q**KT**FF**I**V**AD**V**S  
 PCMV **L**C**M**SL**L**SA**H**KT**L**ID**IR**AE**L**EL**R**I**RY**R**D**VT**VL**LD**V**T**ND**EL**MP**D**I**KN**AC**I**AF**SM**IG**L**AP**IV**TY**Y**IN**Q**L**A**LE**D**S**M**L**I**I**K**Q**E**L**L**S**K**R**LN**-**ID**V**TP**-----**RP**NI**EN**V**ND**R**I**K**KL**K**T**I**PT**M**IN**Q**I**HT**I**E**KA**SP**I**I**CD**V**SP  
 Consensus **L**IM**GL**PS**L**DR**S**Y**E**EV**K**AA**V**Q**Q**L**Q**T**I**FR**AT**FT**IP**P**V**ID**NY**L**L**DM**K**T**AC**L**S**M**S**M**V**AN**L**AS**L**IV**TV**VR**K**L**A**LE**D**SS**M**L**L**V**K**Q**E**LL**M**K**L**DR**R**ERS**-**V**G**ER**PT**----**AR**PH**VS**P**DE**I**AR**L**S**A**L**P**V**M**L**R**Q**L**DD**L**IR**Q**W**EL**VL**PW**I**VP**PL**A**VE**IN**QR**I

HCMV **DK**SAT**C**I**FK**  
 PaHV-2 **DK**SAT**C**I**FK**  
 CeHV-8 **DK**SAT**C**I**FK**  
 TuHV-1 **DK**SAT**C**L**FK**  
 MuHV-2 **DK**SA**V**C**I**FK**S**  
 HHV-6 **DK**RMAT**C**I**Y**K**S**  
 HHV-7 **DK**RLA**T**C**I**FK**S**  
 PCMV **DK**KQ**I**T**C**L**FK**  
 Consensus **DK**SAT**C**I**FK**

**K** = Alpha helix  
**S** = Sheet  
**L** = Loop

Figure 5.2E: Secondary structure predictions for betaherpesvirus triplex  $\beta$ -subunits .

Sequences were obtained from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) and aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. Structural elements present in >50% of examples were calculated and plotted below the CLUSTAL W alignment. The sequence of the triplex  $\beta$ -subunit for HCMV was used to show the presence of conserved structural elements in the 'consensus' for secondary structure.

HHV-4 --MLKIVVSLSSALYTDRIAMQQRICLLPLASTHGTQNVQGLGLGQVYSLETPDVFVEMNYLSDCTLAVLDEVVDSLMLTKIIPGQTYAIKNIYYSLETPDVFVSMYNYLSDCTLAVLDEVVDSLMLTKIIPGQTYAIKNIYYP  
 CeHV-15 --MLKIVVSLSSALYTDRIAMQQRICLLPLASTHGTQNVQGLGLGQVYSLETPDVFVEMNYLSDCTLAVLDEVVDSLMLTKIIPGQTYAIKNIYYSLETPDVFVSMYNYLSDCTLAVLDEVVDSLMLTKIIPGQTYAIKNIYYP  
 CaHV-3 --MLKIVVSLSSALYTDRIAMQQRICLLPLASTHGTQNVQGLGLGQVYSLETPDVFVEMNYLSDCTLAVLDEVVDSLMLTKIIPGQTYAIKNIYYSLETPDVFVSMYNYLSDCTLAVLDEVVDSLMLTKIIPGQTYAIKNIYYP  
 PLHV-1 KFNKCTIVVLLSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 ALHV-1 MAVSKIVVSCSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 AthV-3 ELTKTITVSLTSLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 SaHV-2 ELTKTITVSLTSLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 KSHV KALEKSVVNPVTSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 McHV-5 KALDKSVVNPVTSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 BoHV-4 KIQDRSIVVNPVTSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 EHV-2 KIQDRSIVVNPVTSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 MuHV-4 KASRNIVVNPVTSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD  
 Consensus MALKKSVVNPVTSRLYADRIACKLQMIKVIPIEISHGFGNQIQSLGKPAVAGTDAILDVILMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYDAGDASLDYVLMPSYLSKCTLAITLDEVNADSLVLTKIINDNAYDIKNVYD

HHV-4 FFQWHTGSLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 CeHV-15 FFQWHTGSLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 CaHV-3 FFQWHTGSLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 PLHV-1 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 ALHV-1 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 AthV-3 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 SaHV-2 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 KSHV FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 McHV-5 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 BoHV-4 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 EHV-2 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 MuHV-4 FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI  
 Consensus FFIWCSNIQLVMPVPGRRHATIKLESNVDIVFPMVLTPIAEVFLQKILLFNVYSRVVMAAGADMDDVIRMHLSGSVSYLGHHELALEFVFGPLDIALLDNLSLYFCIMVTLPRASMRIVRGLRHRHHDLNLPQDMVPEIAI

HHV-4 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 CeHV-15 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 CaHV-3 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 PLHV-1 VENLDRHSYEDLTKLQMLMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 ALHV-1 AMNLNADTVTPMFKIGLLITYLQTLGSPNIGRLATAAYSQETLTATCWLR--  
 AthV-3 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 SaHV-2 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 KSHV -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 2McHV-5 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 BoHV-4 -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--  
 EHV-2 ----DFIEDDQKMRSMFAYMOSVSLFPNIGRLATAAYSQETLTATCWLR--  
 MuHV-4 ----QENYVQDILMGLPMAYTHKLSGIPNIGRLATAAYSQETLTATCWLR--  
 Consensus -LDLNLGVSADDLRRMRVMTYLSLSLFPNIGRLATAAYSQETLTATCWLR--

█ = Alpha helix  
 ▤ = Sheet  
 █ = Loop

**Figure 5.2F:** Secondary structure predictions for gammaherpesvirus triplex  $\beta$ -subunits .

Sequences were obtained from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) and aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). Secondary structure predictions were carried out on each individual sequence using PredictProtein (Rost *et al.*, 2004) (<http://cubic.bioc.columbia.edu/predictprotein/>). The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. Structural elements present in >50% of examples were calculated and plotted below the CLUSTAL W alignment. The sequence of the triplex  $\beta$ -subunit for KSHV was used to show the presence of conserved structural elements in the 'consensus' for secondary structure.

## TRIPLEX $\alpha$ -SUBUNIT

Alpha **K**T**I**PLPATPSVWGGSTVELPPTTRD**L**AGQGLLRVLRPPISRRDGPVLP**R**SS**G**PRRAASTLWLLGLDGD**T**APPGALTPN**D**TEQALDKILRGTMRGGAAL**L**G**S**  
 Beta MD**R**AVAKRPRD  
 Gamma **K**VQAE**N**AAR**L**GR**V**LGL**L**PP**P**THR**V**SL**R**GP**E**FARG**V**R**D**LL**S**KYA**A**ST**R**FT**V**BS**L**HH**A**L**R**

Alpha **S**HH**T**R**Q**W**I**L**T**DL**C**Q**P**NA**D**EA**T**L**L**L**R**HPAD**L**G**C**T**R**AG**L**VS**P**N**F**LV**A**ACAAS**Y**D**A**R**D**A**D**AV**R**A**H**V**T**A**N**Y**N**G**T**R**V**G**A**R**L**D**R**F**S**E**C**L**R**A**M**V**H**T**H**V**P**HE**V**M**R**F**F**G**L**V**S**W**V**T**Q**D**E**L**A**S**V**T**A**V**C**A**S**Q**E**A**H**T**G**H**P**G**H**  
 Beta PAD**T**D**N**E**L****V**T**A**L**K**A**R**E**V**N**T**I**S**V**R**Y**L**Y**H**A**H**Q**A**L**T**A**R**F**F**PE**G**L**V**E**F**E**A**Q**P**G**A**L**L**I**R**M**E**T**G**C**D**S**P**R**H**L**Y**I**S****L**Y**L**L**G**I**R**A**S****V**S**A**S**T**R**C**L**L**E**S**V**Y**T**A**S**A**A**R**A**A**L**Q**W**I**D**L**G**P**H**L**L**H**R**L**E**T**L**G**C**V**K**T**V**S**L**I**T**S**L**L**T**C**V**M**R  
 Gamma Q**A**P**F**F**R**Q**E**T**Y**G**D**F**L**V**S**Y**Q**T**F**S**F**C**E**H**L**G**T**F**L**F**S**F**K**Q**E**D**N**G**S**S**M**D**M**L**L**T**P**T**S****L**P**N**L**S**G**E**A**K**A**P**Q**T**H**K**V**A**G**V**W**Y**G**S**G**S**L**A**D**F**I**P**N**L**S**E**L**M**D**T**C**E**F**H**T**L**L**T****V**V**C**P**M**V**Q**S**V**H**S**T**F**V**T**K**V**T**S****A**M**K**G**V**G**L**A**R**D**E**P**A**H**V****L**T**L**P

Alpha **R**SA**V**I**L**P**A**C**A**F**V**D**L**A**E**L**G**L**G****S**P**S**A**F**L**Y**L**V**L**T**Y**R**R**R****V**E**L**C**C**V**Y****V**I**K**S**C**L**P**R**G**L**E**P**A**L**E**R**L**F**G**R**L**R**I**T**N**T**I**H**G**T**E**D**M**T**P**P**A**R**T**C**T**Y**A**A**F**A**E**L**S**M**P**P**E**D**S**P**R**C**L**H**R**T**E**R**F**A**V**T**V**P**V**I**L**E**G**V**V**W**C**P**G**E**W**R**A**C**H**  
 Beta G**Y**L**Y**N**T**L**K**T**E**V**F**A**L**M**I**P**K**D**M**Y**L**T**W**E**T**R**G**R**L**Q**V**Y**V**L**I**I**V**Y**D****S**P**E**T**E**P**G**L**Y**V**L**T**S**S**I**A**H**W**O**T**L**V**D**V**A**R**G**K**F**A**R**E**R**C**S**F**V**N**R**R**I**T**R**P**R**O**I**L**L**C**T**G**V**I**Q**L**G**W**C**L**A**D**D**I**H**T**S**F**L**V**H**K**E**L**K**L**S****V**V**R**L**D**N**F**S**V**E**L**A**D**F**R**E**F**  
 Gamma C**D**M**L**V**D**L**D**E**S**C**P**M**V**L**D**E**S**C**P**M**V**Q**R**R**E**P**A**G**I**N**V**T**I**Y**A**S**L**V**Y**I**R**V**N**O**E**P**S**M**A**L**T**F**P**Q**S****G**F**A**E**V**V**A**M**I**K**D**H**F**T**D**V**I**R**T**K**Y**I**Q**L**R**H**E**L**Y**I**N**R**L**V**F**A**G**A**V**C**T**L**G**T**V**F**D**S**H**P**V**H**Q**S**L**N**V**K**G**T**S**L**P**V**L**V**F**A**N**F**E**A**L**C**G**P**W**T**V**L

## TRIPLEX $\beta$ -SUBUNIT

Alpha **M**L**A**D**F**E**T**D**I**A**I**P**S**G**I**S**R**P**D**A**A**L**Q**R**C**E**R**V**E**L**P**T**I**R**R**Q**L**T**L**A**D**V**A**H**E**S**F**V**S**G**G**V**S**P**D**T**L**G**L**L**A**Y**R**R**R**F**P**A**V**I**T**I**R**V**L**E**T**R**I**V**A**C**E**L**D**V**S**L**T**H**A**G**T**V**N**L**R**N**T**S**P**V**D**L**C**N**G**I**P**I**S**L**V**P**P**V**F**E**G**Q**A**T**D**V**R**L**D**S**L**L**T**L**R**F  
 Beta **M**A**A**N**E**A**N**I**F**C**T**F**H**K**L**S**I**A**D**V**G**K**L**T**L**V**L**A**V**P**I**P**R**L**H**L**I**K**H**Y**Q**L**G**L**H**Q**F**V**H**T**R**G**V**R**L**G**L**L**R**N**M**T**L**L**M**R**R**V**E**G**N**Q**L**L**H**V**P**E**H**L**L**Y**T**V**N**T**G**P**V**T**W**E**K**C**A**L**C**V**L**P**L**E**H**G**P**L**R**E**N**L**L**T**L**S**O**W**E**L**V**L**P**W**I**L**E**  
 Gamma **M**L**A**L**K**S**I**V**V**N**F**F**S**R**L**F**A**D**E**L**A**A**L**Q**S**K**I**S**V**L**P**L**G**D**C**H**R**L**Q**N**I**Q**A**L**G**L**G**C**V**C**S**R**E**T**S**P**D**V**I**Q**I**M**Q**V**L**S**K**T**L**A**V**L**E**V**R**P**P**S**L**R**L**T**R**M**P**S**E**N**L**Q**I**K**N**I**Y**E**F**Q**W**D**S**N**I**Q**L**A**V**L**P**P**F**S**R**K**D**S**T**V**L**E**S**N**G**F**D**L**V**F**P**M**V**

Alpha **V**P**L**F**S**P**L**A**R**E**I**V**A**R**L**V**A**R**G**I**R**D**L**N**P**S**P**R**N**P**G**G**L**P**D**L**N**V**L**Y**N**S**R**L**S**L**L**A**D**V**Q**L**G**P**V**N**A**E**L**R**S**L**V**I**N**M**V**S**I**T**E**G**T**T**I**L**T**L**I**P**R**L**F**A**L**S**A**D**G**Y**V**N**A**L**L**Q**M**Q**S**V**T**R**E**A**Q**L**I**H**P**E**A**P**A**L**M**D**G**E**R**R**I**P**L**Y**E**A**L**V**A**W  
 Beta **M**P**L**A**L**R**I**N**Q**R**L**L**I**M**G**L**F**S**L**D**R**E**Y**E**V**K**A**A**V**Q**L**O**T**T**F**R**A**T**F**T**I****S**D**E**V**I**L**Q**H**L**L**I**D**M**K**T**A**C**L**S**M**S**M**V**A**N**L**A**S**E**L**I**M**T**Y**V**R**K**L**A**E**D**S**S**M**L**L**V**K**C**O**E**L**L**M**R**L**D**R**E**R**S**V**G**E**P**R**T**P**A**R**P**Q**H**V**S**P**D**E**I**A**R**L**S**A**L**F**V**M**L**R**Q**L  
 Gamma **V**P**Q**L**G**H**A**I**L**Q**L**L**V**Y**H**I**Y**S**K**I**S**A**G**A**P**D**D**V**N**M**A**E**L**D**L**Y**T**T**N**V**S**F**M**G**R**T**Y**R**L**D**V**N**T**I**P**R**T**A**L**R**V**L**D**D**L**S**M**Y**L**C**L**S**A**L**V**P**R**G**C**L**R**L**L**T**A**L**V**R**H**R**H**P**L**T**E**V**F**E**G**V**P**D**E**V**T**R**I**L**D**Q**L**S**V**P**D**I**T**M**R**M**V**M**F**S**Y**L**Q**S**L**S**S**

Alpha **L**T**H**A**G**Q**L**G**D**T**L**A**A**E**V**V**R**V**C**T**F**D**G**A**V**V**R**S**G**D**M**A**P**V**I**R**Y**P  
 Beta **D**D**L**I**R**E**G**V**V**F**T**V**C**D**V**E**P**N**K**S**A**T**C**I**E**K**S**  
 Gamma **F**N**L**G**R**L**V**Y**A**Y**S**A**E**T**L**A**A**S**C**W**Y**S**P**R

**X** = Alpha helix  
**X** = Sheet  
**X** = Loop

Figure 5.2G: Secondary structure predictions for triplex  $\alpha$  and  $\beta$  subunits of alpha, beta, and gammaherpesviruses.

Consensus sequences from Figures 5.2A-F are shown. The positions of  $\alpha$ -helices are shown in red,  $\beta$  sheet in turquoise, and coil in green. In the  $\alpha$ -subunit alignment, the amino acids in insert 1 and insert 2 (as described in Figure 5.1) have been removed from the alphaherpesvirus consensus, and the remaining amino acids have been aligned from the C-terminal. In the  $\beta$ -subunit alignment, amino acids have been aligned from the N-terminal. (Greater levels of consensus between alpha, beta, and gammaherpes were observed when triplex  $\alpha$  subunits were aligned from the C-terminal, while in the case of triplex  $\beta$  subunits, greater levels of consensus were observed when alignment was carried out from the N-terminal.) Although secondary structure predictions confirm similarities between viruses for each type of protein, they do not indicate an evolutionary relationship between the two proteins.

VP19C MKTNPLPATPSVWGGSTVELPPTTRDTAGQGLLRVLRPPISRRDGPVLPRGSGPRRAAS 60  
 VP23 MLADGFETDIAIPSG-----ISRPDAAALQRCEG----- 29  
 \* \* \* \* \*

VP19C TLWLLGLDGTDAPPGALTPNDDEQALDKILRGTMRGGAALIGSPRHHLTRQVILTDLCQ 120  
 VP23 -----RVVFLPTIRRQLTLADVAHESFV 52  
 \* \*

VP19C PNADRAGTLLLALRHPADLPHLAHQRAPPGRQTERLGEAWGQLMEATALGSGRAESGCTR 180  
 VP23 SGGVSPDTLGLLLAYRRRFPVITRVLP-----TRIVACPLDVGLTH 94  
 \* \* \* \* \*

VP19C AGLVSFNFLVAACAASYDARDAADAVRAHVVTANYRGTRVGARLDRFSECLRAMVHTHVFP 240  
 VP23 AGTVNLRNTSPVDLCNGDP-----ISLVPPVFEGQATDVRDLSLDLTLRFPVP----- 142  
 \* \* \* \* \*

VP19C HEVMRFFGGLVSWVTQDELASVTAVCAGPQEAHAHTGHPRPRSAVILPACAFVDLDAELG 300  
 VP23 -----LPSPLAREIVARLVARGIRDLNPSPRNPGGLPDLNVLYYNGSRLSLLADVQ 193  
 \* \* \* \* \*

VP19C LGGPGAFLYLVLTYRQRDQELCCVYVIKSQLPPRGLEPALERLFGRLRITNTIHGTED 360  
 VP23 QLGP-----VNAELRSLVLMVYSITEGTT--IILTLPRLFALSAQDGYVNALLQ 242  
 \* \* \* \* \*

VP19C MTPPAPNRNPDFLAGLANPQTPRCSAGQVTNPQFADRLYRWQPDLRGRPTARTCTYAA 420  
 VP23 MQSVTR-----EAAQLIHPEAPALMQDGERRLPYEAALVAVLTHAG----- 283  
 \* \* \* \* \*

VP19C FAELGMMPEDSPRCLHRTERFGAVTVPVVILEGVVWCPGEWRACA 465  
 VP23 --QLGDTLALAP--VVRVCTFDGAAVVRSGDMAPVIRYP----- 318  
 \* \* \* \* \*

**Figure 5.2Hi**



VP19C MKNPLPATPSVWGGSTVELPPTTRDTAGQGLLRRVLRPPISSRRDGPVLPKSGPRAAS 60  
 random -----

VP19C TLWLLGLDGTDAAPPALTPNDDTEQALDKILRGTMRGGAALIGSPRHHLTRQVILTDLCQ 120  
 random -----YDKRPQFSYESWKRYNRDQDNYHGHRMIAWMRLWPQHSFCFSDTRVILENLYS 53  
                   \*                  \*                  \*                  \*                  \*\*\*

VP19C PNADRACTLLLALRHPADLPHLAHQRAPPGRQTERLGEAWGQLMEATALGSGRAESGCTR 180  
 random -----YEHMCTEPCHGIDHYTTENAWRTYLMKEIWSRHLTACIKWG---SFTAMR 100  
                                   \*                  \*                  \*                  \*                  \*

VP19C AGLVSFNFLVAACAASYDARDAADAVRAHVVTANYRGTRVGARLDRFSECLRAMVHTHVFP 240  
 random HNCYCMRHHMIMDGRWYSPFRVAIAIAKHSKG-----YSWKMANKEGIP 144  
                                   \*                  \*                  \*                  \*                  \*

VP19C HEVMRFEGGLVSWVTQDELASVTAVCAGPQEAHTGHPGRPRSAVILPACAFVDLDAELG 300  
 random HRWIIETAKRRRWYDWQFYHP---IEQPHNEGFTDINN--TRILMIRSCSKNDLDEE- 198  
                   \*                  \*                  \*                  \*                  \*                  \*\*\*

VP19C LGGPGAFLYLVLTYRQRDQELCCVYVIKSQLPPRGLEPALERLFGRLRITNTIHGTED 360  
 random -FSSQAFQCQRFDFTGERKDDFTDFQTIQDWTAN--IAAFQRQCCKS-CTVALHGHPT 254  
                   \*                  \*                  \*                  \*                  \*                  \*\*

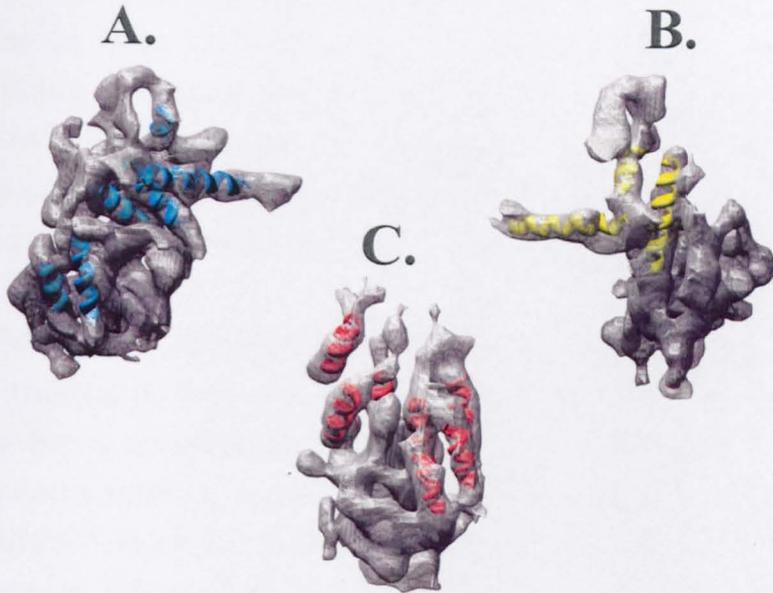
VP19C MTPPAPNRNPDFPLAGLANPQTPRCSAGQVTNPQFADRLYRWQPDLRGRPTARTCTYAA 420  
 random IGPEACSMN-----NPTR---TGQRQ-----WDRDQ--EYFFWSIEYQF 288  
                   \*                  \*                  \*                  \*                  \*                  \*

VP19C FAELGMPEDSPRCLHRTERFGAVTVPVVILEGVVWCPGEWRACA 465  
 random DMQIPFWEGDSPEGLD-----VATHCWFQCKKGW---- 318  
                   \*\*\*                  \*                  \*                  \*

**Figure 5.2Hiii**

**Figure 2H:** VP19C/VP23 alignments do not give an indication of an evolutionary relationship between the two proteins.

**(i).** HSV-1 VP19C and VP23 were aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw>) multiple sequence alignment program (Thompson *et al.*, 1994). **(ii).** VP23 was shuffled using the Sequence Manipulation Suite Shuffle Protein program ([http://bioinformatics.org/sms2/shuffle\\_protein.html](http://bioinformatics.org/sms2/shuffle_protein.html)) (Stothard, 2000) and aligned to VP19C using CLUSTAL W. **(iii).** A random 318 amino acid sequence was generated using the Sequence Manipulation Suite Random Protein Sequence program ([http://bioinformatics.org/sms2/random\\_protein.html](http://bioinformatics.org/sms2/random_protein.html)) (Stothard, 2000), and aligned to VP19C using CLUSTAL W. Conserved residues are denoted by asterisks. Similar numbers of conserved residues were obtained from the three alignments (66, 65, and 54 respectively).



**Figure 5.3:** Three-dimensional reconstructions of the HSV-1 triplex proteins.

**A. and B.** Three-dimensional reconstructions of the two copies of VP23. **C.** Three-dimensional reconstruction of VP19C. The three proteins of the triplex are shown in an 'exploded' view. Although the similar structural elements in the two VP23 molecules are evident, there appears to be little obvious similarity between VP23 and VP19C. The predicted positions of  $\alpha$ -helices are shown in each reconstruction. Four  $\alpha$ -helices are visible in VP19C, and they are not predicted to be in similar positions to  $\alpha$ -helices in VP23.

Images were supplied by Matthew Baker.

and four insertions between amino acids 327 and 330, while no insertions occur between amino acids 144 and 181 (Table 3.2). The GC content of HSV-1 is reported as being 68.3% (McGeoch *et al.*, 1988). The average GC content of amino acids 144-181 in VP19C is 80.7%, while the GC contents of 234-248 and 327-330 are 60% and 41.7% respectively, indicating that the mutagenesis system might have a preference for less-GC-rich DNA. The three cycles of mutagenesis resulted in the production of 11, 19, and 20 unique insertional mutants of the UL38 ORF respectively (Section 3.2). As well as producing the most mutants, the method used for the third cycle produced mutants in the shortest time. Although the plasmid used in this cycle was larger, and therefore contained a smaller proportion of insertions in the UL38 ORF, the number of cloning steps required in the procedure was reduced, accounting for the faster production.

The results of mutational analysis must be interpreted with caution, particularly when considering the behaviour of individual mutants. In some cases, pairs of mutants with contrasting properties were very close together. For example, the plasmids containing insertions in286 and in289 differ markedly in their ability to complement the growth of  $\nu\Delta 38$ YFP (Figure 3.6). In addition, not all functions are equally sensitive to insertion. For example, although in38 maps within the sequences identified as containing a nuclear localisation signal, the ability of this mutant to enter the nucleus was not affected (Figure 5.4). Therefore, without detailed information on the structure of the target protein or specific assays to measure particular properties it is usually not possible to make other than a general interpretation of results. Three-dimensional structural analysis of the capsid has shown that the component proteins interact in different ways at distinct locations in the capsid (Zhou *et al.*, 1998a). Furthermore, the reconfiguration that occurs during capsid maturation involves domain movements leading to changes in the patterns of interaction (Trus *et al.*, 1996; Heymann *et al.*, 2003). This type of behaviour means that the importance of particular regions of a protein may vary with maturation status and location.

This type of analysis may not be optimal for structural proteins lacking separable domains. Nevertheless, the distribution of severely disabling mutations across VP19C was non-uniform, with 16 in the central 251 residues of VP19C but none in the N-terminal 107 residues and only two in the C-terminal 107 residues. The most N-terminal severely disabled mutant, in113, occurs almost immediately after the first evidence of conservation between the sequences of the alphaherpesvirus  $\alpha$ -subunits (Figure 3.6). Originally, it had appeared that an equivalent distinction marked that boundary of the C-terminal 107 residues. Preliminary alignments of the sequences of VP19C and its homologues in

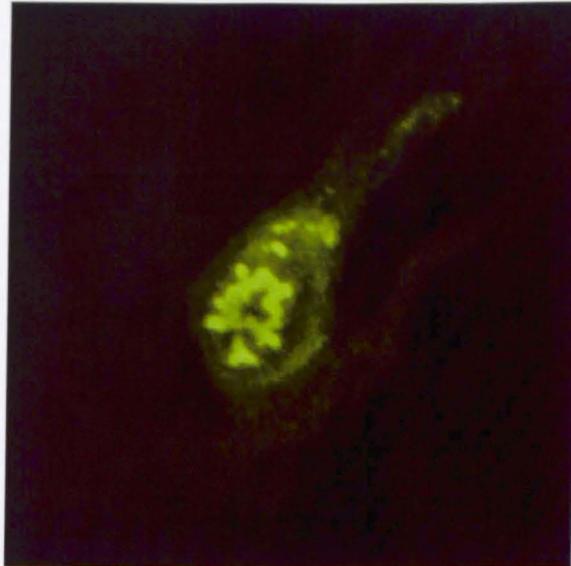


Figure 5.4: Nuclear localisation of UL38in38.

BHK-21 cells were cotransfected with VP19C insertional mutant in38. VP19C was detected with the monoclonal antibody mAb02040, and visualised using FITC GAM (green).

*Alphaherpesvirinae* appeared to show a lack of conservation in amino acid sequence that began approximately at residue 360 in HSV-1 and extended to the C-terminus, however closer examination of the DNA sequences revealed that the apparent lack of conservation had been caused by an error in the sequence of the VP19C homologue in gallid herpesvirus-1 provided by University College London's Virus Database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA3/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA3/VIDA.html); Mar Albà *et al.*, 2001). The introduction of a single extra thymine at position 1237 in the DNA sequence of this gene caused a frameshift which resulted in the next amino acid (aspartic acid (GAT)) being substituted with a STOP codon (TGA). This error resulted in the C-terminal 68 residues of the gallid herpesvirus-1 VP19C homologue being omitted from the protein sequence. When the error in the protein sequence was corrected and it was aligned with the VP19C homologues of other *Alphaherpesvirinae*, there was no decrease in levels of sequence conservation in the C-terminal region of VP19C compared with the middle region.

Despite there being no evidence of a loss of sequence conservation in the 107 C-terminal residue region, its start at in358 coincides with the beginning of the second large alphaherpesvirus specific insert (insert 2; Figure 5.1). However, unlike the poorly conserved N-terminal region, insert 2 contains a severely disabling mutation, in390 (Figure 3.6). Only one insertion (in143) mapped within the insert 1 locus, and this was also severely disabling (Figure 3.6), indicating that both inserted sequences are functionally important. The distribution of severely disabling mutants suggests that VP19C can be divided into three regions based on their ability to tolerate insertions, and may suggest that the three regions have differing roles. However, this interpretation must be treated with caution as the distribution of insertions is not uniform and clustering of mutations with similar phenotypes (such as in327-330) tends to distort the pattern. Statistical analysis by chi-squared analysis (Table 5.1) indicates that the N- and C-terminal regions are more tolerant of insertions than the middle region (in contrast to what was stated in Adamson *et al.*, 2006).

The sensitivity to insertion of the central region suggests that it may be important for folding of VP19C in a similar manner to that suggested for the central region of the major capsid protein, VP5, which has been proposed to act as a folding nucleus for the more flexible N and C-terminal regions (Bowman *et al.*, 2003). The large number of tolerated insertions in the C-terminal region is rather surprising, as deletional analysis had shown that removal of as few as 15 amino acids from the C-terminus destroyed the ability of

	Rescuing	Non-Rescuing	Total
N-Terminal Region	12	0	12
Middle Region	8	15	23
Total	20	15	35
$\text{Chi-squared} = \frac{35[(12 * 15) - (0 * 8)]^2}{12 * 23 * 20 * 15} = 13.70$			
Two rows, two columns, therefore $(2-1)*(2-1) = 1$ degree of freedom.			
$P < 0.001$			
Difference in ability to tolerate five amino acid insertions between the N-terminal region and middle region of VP19C is very highly significant.			

	Rescuing	Non-Rescuing	Total
Middle Region	8	15	23
C-Terminal Region	11	2	13
Total	19	17	36
$\text{Chi-squared} = \frac{36[(8 * 2) - (15 * 11)]^2}{23 * 13 * 19 * 17} = 8.276$			
Two rows, two columns, therefore $(2-1)*(2-1) = 1$ degree of freedom.			
$0.001 > P > 0.0001$			
Difference in ability to tolerate five amino acid insertions between the C-terminal region and middle region of VP19C is highly significant.			

**Table 5.1:** Chi-squared tests to showing differences in the three regions of VP19C in terms of ability to tolerate insertional mutagenesis.

VP19C to support capsid assembly (Spencer *et al.*, 1998). When considered alongside the behaviour of in463, which completely fails to support virus growth (Figure 3.6), this seems to indicate that sequences very near the C-terminus of VP19C play an essential role in capsid assembly. However, the viability of most of the mutants in the C-terminal region suggests that it has considerable structural flexibility despite the secondary structure predictions, which appear to show that conservation of secondary structure increases towards the C-terminus (Figures 5.2A-5.2C and 5.2G). Previous work has highlighted the structural flexibility of triplex proteins, with it being suggested that in isolation the two triplex proteins (VP19C and VP23) behave like molten globules – proteins in an intermediate state of folding. The plasticity inherent in such a molten globule-like state might provide the flexibility required to attain the non-equivalent interactions observed in the triplex (Kirkitaдзе *et al.*, 1998).

The 18 severely disabled mutants can be divided into four categories based on their ability to interact with VP5 and VP23 (Table 3.3). Five (including in242A and in242B, which are in the same codon) failed to transport either VP23 or VP5 to the nucleus (Figures 3.7 and 3.8), suggesting that these interactions may have altered the overall folding of the protein. These insertions (at residues 113, 242, 289, and 330) are evenly spaced across the central region and do not appear to demarcate any obvious domain although they may lie close together in the folded protein. When insertional mutagenesis was carried out on the region of VP5 for which the crystal structure has been obtained, it was shown that lethal insertions were into regions of the protein that were deeply buried away from the surface, possibly disrupting the overall folding of the protein (Webster, 2004). This may also be the situation with the VP19C insertional mutants described here, although this could not be confirmed without the crystal structure of VP19C.

Four of the mutants interact with VP5 but not VP23, while three interact with VP23 but not VP5 (Figures 3.7 and 3.8). The specificity of these mutants suggests that the insertions do not cause global misfolding of the protein but have a more localised effect. The mutants that failed to take VP23 to the nucleus were located N-terminal (between residues 193 and 310) to those that failed to bind VP5 (between residues 328 and 463), suggesting that the regions involved are particularly important for the binding of VP23 and VP5 respectively. Since VP23 and VP5 are both essential for capsid assembly (Desai *et al.*, 1993; Tatman *et al.*, 1994; Thomsen *et al.*, 1994) it is not surprising that the mutants that fail to interact with one or both of them do not support virus growth. The nature of the defect in the remaining severely disabled mutants is unclear. Further analysis is required to determine

whether they support the assembly of structurally normal capsids. However, examination of mutants in143, in234, in235, in327, in350, and in390 in a plasmid-borne DNA packaging assay (Hodge and Stow, 2001) showed that they were unable to support DNA packaging in  $\nu\Delta 38$ YFP infected cells (Figure 3.9). This suggests either that the assembly of the capsid shell was disrupted, presumably due to some subtle alteration in the interactions with VP23 or VP5, or that an interaction involving one of the other virion components had been affected. An interaction between VP19C and the UL25 protein has been indicated from far-western blotting (Ogasawara, 2001), although it has not been possible to replicate this result using immunofluorescence (Valerie Preston, personal communication). The tegument protein present at the vertices in mature virions is in contact with the peri-pentonal triplexes and may also interact with VP19C (Zhou *et al.*, 1999). Failure to bind tegument would be lethal but would not be expected to prevent DNA packaging. However, UL25 mutants show aberrant DNA packaging (McNab *et al.*, 1998; Stow, 2001) and disruption of this interaction might result in the type of phenotype seen here.

### **5.3. Mutational analysis of the N-terminal region of VP19C**

The N-terminal regions of VP19C and its homologues in all three herpesvirus subfamilies are particularly poorly conserved in both length and sequence, with effectively no homology evident until after amino acid 111 of VP19C (Figure 5.1). This raises questions regarding the function of this highly variable region. Part of that function was revealed by the deletion of 45 residues from the N-terminus of VP19C, which removed a nuclear localisation signal (Section 4.2.1 and 4.2.3). That the failure of this deleted protein to congregate in the nucleus was entirely responsible for the 50-fold reduction in its ability to complement growth of a VP19C minus mutant was demonstrated by the recovery in virus growth when an exogenous (SV40) nuclear localisation signal was present (Section 4.2.3). From this observation it is evident that correct transport of the component proteins to the site of capsid assembly is an important function of VP19C. Although the precise sequence of the VP19C NLS has not been identified, it has been mapped to a region containing 33 amino acids (Section 4.3, Figure 4.6). Database searching has revealed that it does not belong to any known classes of NLS. Interestingly, the pattern of arginine residues seen in the VP19C NLS is not present in the N-terminal sequences of any other alphaherpesvirus VP19C homologue apart from that of the very closely related HSV-2 (Figure 5.5). Furthermore, examination of triplex  $\alpha$ -subunit sequences from all three herpesvirus

```

HSV1 -----MKTNPLPATPSVWGGSTVEL-----PPTTRDTAQQGLLRR 35
HHV2 -----MKTKPLPTAPMAWAESAVET-----TTSPRELAGHAPLLRR 35
gallidHV2 -----MKPLLRSHETQYYSLYPDT-----LNGLHNVFRTIGNSV 34
m'eagridHV1 -----MALPRSRNTKTASRTSEMKYSSAHSNR-----PYELHSLMRSIGKNI 43
cercoHV7 MEANQRNSKLTQLTQIPT-SNVMLMGDNRFIQIGNGLN-----MSYNSHNRF-----S 47
VZV MGSQPTNSHFTLNEQTLCGTNISLLGNRFIQIGNGLH-----MTYAPGFFGWNSRDL 53
equidHV4 -----MNLGGNRFVQIGNGMSN-----IMYTDANGG--VRWE 30
equidHV1 -----MNLGGNRFVQIGNGMSN-----IMYTDANGA--VRWE 30
psedorabies -----MSVQIGNGLL----- 10
bovineHV1 -----MAAPNGSSSYIQIGNHLRMRLPAAAPPFSGVPAAAAAAA 40
gallidHV1 -----MSGLLARNGSNDKIRAPERGLSG-----FANTRTHPYLQSV 36

HSV1 VLRPPISRRDGPVLPRGSGPRRAASTLWLLGLDGTDAPPGALTPNDDTEQALDKILRGTM 95
HHV2 VLRPPIARRDGPVLLGDRAPRRTASTMWLLGIDPAESSPGTRATRDDTEQAVDKILRGAR 95
gallidHV2 QSDTVRRLNLGYLDGDNRRGNLAGGLELLRDATTPNASRMNITRPLDTSTSGATAMIMQS 94
m'agridHV1 RTSVGGRINLGYMNVGDRS--VNDLNVLWTADTLGMYKMDRTPAFKDSTAKTTAIIAQS 101
cercoHV7 SIKPWN--LSTDKFTNTAN-----GTLVRITPDCIVINNAHLQIQQN---TS 90
VZV TIGPRFGGLNKQIHVPPKRTE-----TASIQVTPRSIVINRMNIQINT---SI 101
equidHV4 QVPPFAGFPHQ---RGRSNAAFGLPNTLDWLPGFVHATPNSITISNMGGIQISSAGVVTA 87
equidHV1 QISPPAGFPQQRGRGRGHVAFGLPNTLDWLPGFVQATPNSITISNMGGIQISSAGVITA 90
psedorabies -----MVVAPGTLTVGSAR----- 24
bovineHV1 ASAAEAGVQSVTAPAVGGARGRAPRGYNPWAGGMLHVSDATVTIQNMSGIQIVTPRQIAV 100
gallidHV1 LPRRTHQHNLLSAALRLTTGRNFSHGLTRPGKFIGSGRNPEETNNATLGKEDTRAI 96

HSV1 R-GGALIGS-----PRHLTRQVILTDLCQPNADRAGTLLLALRHPADLPHLAHQRAPP 149
HHV2 RAGGLTVPGA-----PRYHLTRQVTLTDLCQPNAEPAGALLLALRHPTDLPHLARHRAPP 150
gallidHV2 LRTDVAENTLLLTGDSRATISRQVLTDFCFPDAEMPGLILSIRHPLDINSEALYATPA 154
m'agridHV1 IRSN--SEIREWGGD-RAVITRQVTLTDFCFPDAEFFGSILLSMRHPLDINSEALYATPG 158
cercoHV7 D---VFRLPLEQTHITPDLIKQVSLTDLCRPDVDFLGSPVFLIRHCLDIVEDASVCTPA 146
VZV GNPQVTIRLPLNNFKSTQLIQVSLTDFFRPDIEHAGSIVLILRHPSDMIGEANTLTQA 161
equidHV4 AINSDHNSWVLSNSQPSLKLTRQVTLTDFCDPTAEKPGLPIIRLRHLDAIGSSPSTPP 147
equidHV1 AINSEQNSWMLSSNFPSLKLTRQVTLTDFCDPTAERPGLPIIRLRHLDAIGSSPSTPP 150
psedorabies -----ARLIRQVTLADFCEPQAERPGLVVLALRHPADLAGAAYAATPP 67
bovineHV1 DTPAGTAVLSPGG-QPHIRLSRQVLTDFCDPQLERPGAPVLTLKHPADIIGLAAAAAPP 159
gallidHV1 ASKPLLDVESTRTN--DVISLTQITVIDLCQPGVEESGSLMLFLKGVKDLLKILGSRPAR 154

```

Figure 5.5: Arginine residues in the N-terminal regions of triplex  $\alpha$  subunits of alphaherpesviruses.

The sequences of eleven alphaherpesviruses were aligned using the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (Thompson *et al.*, 1994). The N-terminal sequences from this alignment are shown. Arginine residues are highlighted in red. The N-terminal region of HSV-1 VP19C that contains a NLS is shown in yellow.

subfamilies using the predictNLS program failed to identify any recognisable NLSs. Therefore, it is unclear whether a role in nuclear localisation represents the usual function of the poorly conserved alphaherpesvirus N-termini, or where NLSs, if any are present, might be located in the  $\alpha$ -subunits from beta and gammaherpesviruses.

The mitochondrial distribution of GFP that was observed for p33-66AA-UL38-GFP (Figure 4.6) was unexpected. Presumably the NLS, which occupies the same region, is dominant and the mitochondrial signal only becomes apparent after increasingly large deletions have rendered the NLS inactive. Such cryptic mitochondrial localisation signals may be common in HSV-1; iPSORT identified nine sites in the amino acid sequence of HSV-1 VP19C alone that fit its criteria for potential mitochondrial localisation signals. Regardless of how accurate these predictions might be, the nuclear localisation of VP19C demonstrates that activity from any potential mitochondrial localisation signal is masked by the NLS. It has previously been shown that the distribution of mitochondria is altered by HSV-1, with mitochondria migrating to the perinuclear region of the cytoplasm following infection and colocalising with the gene products of UL41 and UL46 (both of which are tegument proteins). The reason for this migration is unclear, but it was suggested that the mitochondria might supply the energy required for the incorporation of tegument proteins into the virion (Murata *et al.*, 2000). Previous studies have also described alterations in behaviour of the mitochondria during HSV-1 infection (examples include Lund and Ziola, 1985; Lund and Ziola, 1986; Latchman, 1988; Tsurumi and Lehman, 1990), however, there has been no previous evidence of a relationship between VP19C and mitochondria. It appears likely that mitochondrial localisation signals are common in HSV-1 but in most cases are non-functional, and that the region of VP19C attached to GFP in p33-66AA-UL38-GFP contained one such signal that is non-functional in wild-type VP19C.

The study of nuclear localisation signals has shown that they exist in various types. The NLSs found in many proteins contain monopartite or bipartite sequences that are rich in basic residues (reviewed by Adam, 1999; reviewed by Görlich and Kutay 1999). These sequences recruit importin  $\alpha$  proteins involved in nuclear translocation. Importin  $\alpha$  in turn binds importin  $\beta$  protein, which mediates nuclear localisation (Chook and Blobel, 2001). A limited but growing list of proteins translocate to the nucleus in the absence of importin  $\alpha$  binding signals, instead interacting directly with importin  $\beta$  proteins via arginine-rich motifs (Palmeri and Malim, 1999). Examples include human ribosomal proteins (Jakel and

Görlich, 1998) and human immunodeficiency virus (HIV) Tat and Rev (Truant and Cullen, 1999).

Schaley *et al.* (2005) identified a NLS in the N-terminal 58 amino acids of the adenovirus E4-6/7 protein. Although there is no evidence of an evolutionary relationship between this protein and VP19C, they contain NLSs with similar properties. Thus, both proteins have a relatively arginine-rich N-terminal region, neither protein contains monopartite or bipartite clusters of basic amino acids that might act as a NLS, and NLS prediction programs PredictNLS and PSORT II were unable to identify a NLS in this region of either protein. The authors proposed that E4-6/7 translocates to the nucleus via a direct interaction with importin  $\beta$  as described above. It would be interesting to test whether VP19C also interacts directly with importin  $\beta$ .

An alternative explanation for the nuclear localisation of VP19C can be derived from work carried out by Sessler and Noy (2005) in which they attempted to find a nuclear localisation signal in CRABP-II. CRABP-II is a cytosolic protein that moves to the nucleus upon binding of retinoic acid. The amino acid sequence of CRABP-II does not contain a recognisable NLS. The authors showed that in the presence of retinoic acid, CRABP-II undergoes a conformational change, with a surface patch that is neutral in the absence of retinoic acid becoming positively charged following binding. This patch was mapped to three basic amino acids: two lysine residues at positions 20 and 30, and an arginine at position 29, which shift their orientation to place their side chains in an alignment similar to that seen with the SV40 NLS. Similarly, although no NLS is evident from the VP19C protein sequence, in the folded protein the arginine residues of VP19C may lie close together, acting in a similar fashion to more established NLSs. Evidence of an interaction between VP19C and importin  $\alpha$  would support this hypothesis.

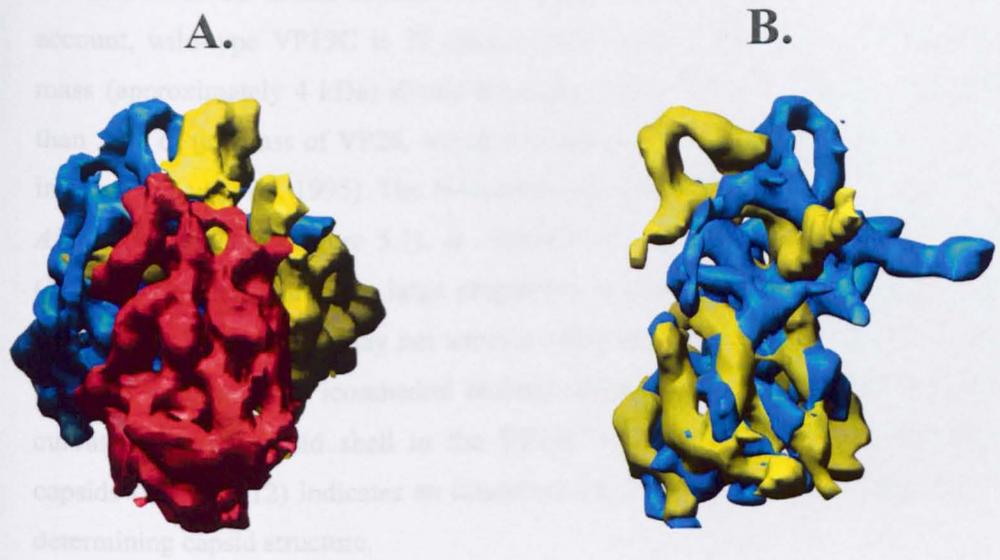
Since the NLS is contained within the first 56 amino acids of VP19C (Section 4.3), the function of the rest of the poorly conserved N-terminal region remains unclear. EM analysis showed that the VP19C-63 truncation was defective for capsid assembly, both in a baculovirus based capsid assembly model, and by complementation of  $v\Delta 38$ YFP (Section 4.4, Figure 4.8). The incomplete capsid shells seen in baculovirus infected cells confirm the fluorescence results showing that the VP19C-63 protein is able to interact with VP5 and VP23, and indicate that the defect in assembly is not a result of a general disruption to the protein's structure. The results presented here contrast with those of Spencer *et al.* (1998), who showed that a mutant of VP19C (nd90), which lacks the N-terminal 90 amino

acids, could support capsid assembly in a baculovirus model system. The reason for this difference is unclear. Although apparently intact capsids were formed with nd90, their assembly was inefficient and the majority of structures seen resembled the incomplete particles shown in Figure 4.4. It is possible that slight differences in the baculovirus model systems used in the two studies account for the contrasting results. Alternatively, when exposed at the N-terminus, amino acids 63-90 may have a disruptive effect on protein folding or interaction that is negated when they are absent. Whatever the reason, it is clear that sequences between amino acids 45 and 63 are necessary for efficient capsid formation and further analysis will be needed to determine precisely what role they play.

#### **5.4. Three-dimensional reconstructions of VP19C**

Previously, VP19C was thought to extend over the top of the triplex to connect with VP5 (Zhou *et al.*, 2000). Recent results have indicated that the density at the top of the triplex should be reassigned to VP23 (Matthew Baker, personal communication). This reassignment was primarily based on interpretation of mass continuity in the new map, and was strengthened by the striking similarity between the tops of the two VP23 subunits (Figure 5.6). The large size difference between VP19C and VP23 seen in HSV-1 is primarily a property of *Alphaherpesvirinae*. For example, in rhesus rhadinovirus (RRV, a gammaherpesvirus), TRI-1 (the VP19C counterpart) and TRI-2 (VP23) are approximately equal in size (O'Connor *et al.*, 2003). It seems unlikely therefore that TRI-1 is large enough to extend over the top of the triplex to make contact with the VP5 subunit on the far side. Despite this, the contacts between the triplexes and capsomers are located in similar positions (Yu *et al.*, 2003). Although the resolution of the RRV map is not sufficient to reveal the dispositions of the triplex proteins accurately, this connection is in the part of the triplex that is believed to be formed by the two TRI-2 molecules.

The size ratio of VP19C to VP23 has been estimated from the cryo-EM reconstruction as 1.2-1.3 (Matthew Baker, personal communication), which compares with a ratio of approximately 1.5 based on their amino acid sequences. Some of this difference may be accounted for by the difficulty of segmenting the triplex from the rest of the capsid. However, the estimated relative size of VP19C is considerably lower than expected, suggesting either that a significant fraction of its size has been assigned to another capsid component, or that some portions of VP19C are not resolved in the icosahedral capsid map. Although the appearance of additional mass in the wild-type triplex when compared to the VP19C-45NLS triplex (Section 4.5, Figure 4.12) would be consistent with the

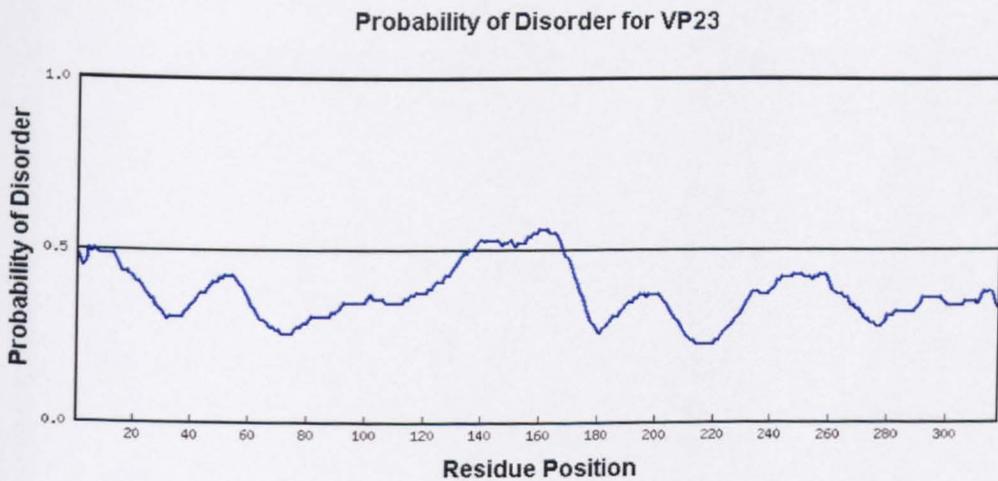
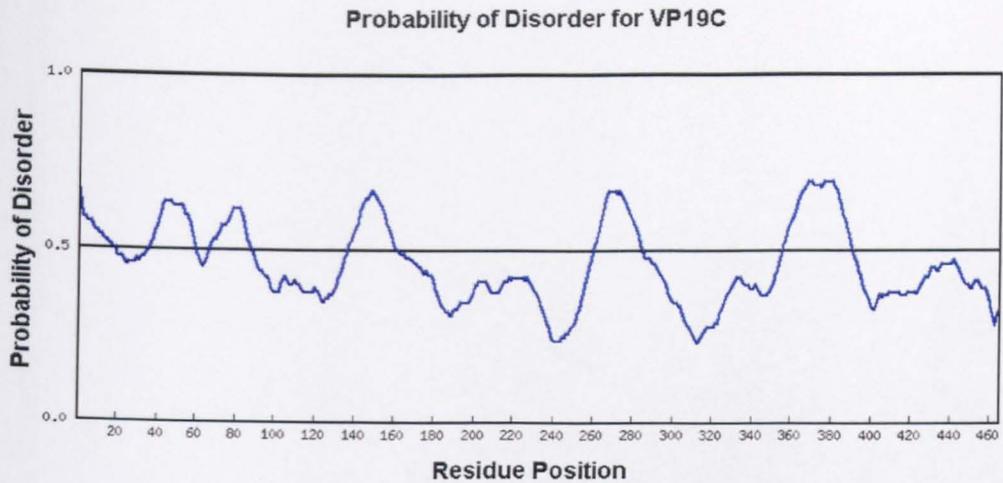


**Figure 5.6:** The similarity between the tops of the two VP23s in a triplex.

**A.** Three-dimensional reconstruction of an HSV-1 triplex. VP19C is shown in red. The two VP23s are shown in blue and yellow. **B.** The two VP23s shown after segmentation and alignment using foldhunter (Jiang *et al.*, 2001). The tops of the two VP23s both contain loop-like and arm-like structures, although they are twisted relative to each other. The loop-like structures were previously assigned to VP19C.

Images supplied by Matthew Baker.

presence of extra residues at the N-terminus of VP19C (Figure 5.1), this putative extra VP19C mass occupies a position that has been assigned to VP23. This, together with the other changes between the two sets of triplexes, suggests that deletion of the N-terminal 45 amino acids has had a global effect on the triplex, consistent with its global effect on capsid structure. Because of this effect, it appears that the location of the VP19C N-terminus cannot be identified by this approach. This was surprising given the size of the deletion involved. When differences relating to the cloning of the NLS are taken into account, wild-type VP19C is 37 amino acids longer than VP19C-45NLS. The addition mass (approximately 4 kDa) should be readily detected. For example, it represents more than 30% of the mass of VP26, which was mapped to the top of the hexons by difference imaging (Zhou *et al.*, 1995). The N-terminal portion of VP19C is poorly conserved among *Alphaherpesvirinae* (Figure 5.1), is insensitive to disruption by insertional mutagenesis (Figure 3.6), and contains a large proportion of disordered residues (Figure 5.7). These properties suggest that it may not adopt a consistent conformation and thus would not be resolved accurately by icosahedral reconstruction techniques. Nevertheless, the reduced curvature of the capsid shell in the VP19C-45NLS capsids compared with wild-type capsids (Figure 4.12) indicates an important role for the N-terminal region of VP19C in determining capsid structure.



**Figure 5.7:** Disorder predictions for HSV-1 VP19C and VP23.

Disorder predictions were carried out using RONN (Yang *et al.*, 2005) (<http://www.strubi.ox.ac.uk/RONN>). Residues showing values above 0.5 are predicted to be disordered. The N-terminal region of VP19C contains a large proportion of disordered residues. Interestingly, insert 1 and insert 2 (Figure 5.1), which comprise amino acids 140-176 and 357-405 in HSV-1 VP19C respectively, also appear to be disordered regions.

## **CHAPTER 6**

### **LABELLING VP5 WITH FLASH-EDT<sub>2</sub>**

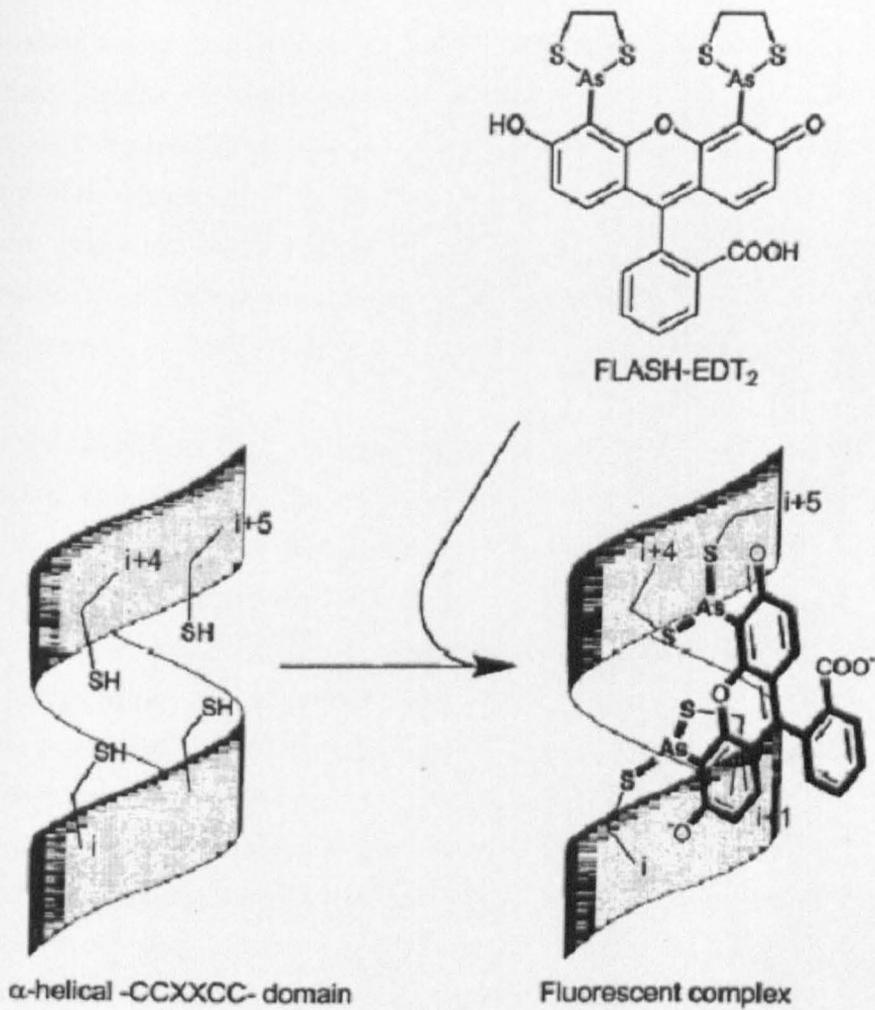
## 6. Labelling VP5 with FIAsh-EDT<sub>2</sub>

### 6.1. Introduction

Tagging proteins by fusing them with fluorescent markers can permit their behaviour within living cells to be examined. The most frequent markers used at present are auto-fluorescent proteins from a variety of marine coelenterates, especially the green fluorescent protein (GFP) from the jellyfish *Aequoria victoria*. These proteins can be fused to the protein of interest to form a chimera that frequently retains the properties of the target protein. However, GFP and its variants are proteins of 25-27 kDa (238 amino acids). The large size of GFP (often larger than the protein of interest) can interfere with the distribution, function and fate of recombinant proteins. Furthermore, in most cases it can only be fused at the N or C terminus of the host protein (Griffin *et al.*, 1998; Gaietta *et al.*, 2002). This means that some proteins may not be amenable to analysis by this method. For example, GFP tagging has been carried out successfully on the smallest capsid protein, VP26 (Desai and Person, 1998), while attempts to tag the triplex protein VP19C with GFP have been unsuccessful (Frazer Rixon, personal communication). The successful tagging of other HSV-1 capsid proteins with GFP has not been reported.

Griffin *et al.* (1998) first described labelling with FIAsh-EDT<sub>2</sub>. In this method (see also Figure 6.1), proteins are modified by genetically inserting a small motif (typically 6-20 residues) containing the sequence CCXXCC (where XX represents any two amino acids). Intact cells expressing the protein are labelled by exposing them to a membrane-permeant non-fluorescent bi-arsenical compound (FIAsh-EDT<sub>2</sub>). FIAsh-EDT<sub>2</sub> binds with high affinity and specificity to the tetracysteine motif and thereby becomes strongly green fluorescent. Toxicity and binding of the trivalent arsenic atoms to endogenous thiols are minimised by the simultaneous administration of micromolar concentrations of antidotes such as 1,2-ethanedithiol (EDT) (Gaietta *et al.*, 2002).

The earliest designs of tetracysteine sequences were intended to encourage  $\alpha$ -helicity, under the assumption that the bi-arsenical dye would ideally fit into the *i*, *i*+1, *i*+4, and *i*+5 positions of an  $\alpha$ -helix (Griffin *et al.*, 1998). When sequences of this type were used, labelled protein had to present in high concentrations to exceed background fluorescence (Stroffekova *et al.*, 2001; Adams *et al.*, 2002). Some reduction of background fluorescence was achieved by increasing the concentrations of the di-thiols EDT and British Anti-Lewisite (BAL) in washes to remove thiol-dependent background, or by including



**Figure 6.1:** FIASH-EDT<sub>2</sub> labelling of proteins.

Fluorescein arsenical helix binder (FIAsH) in complex with two 1,2-ethanedithiol (EDT) molecules (FIAsH-EDT<sub>2</sub>) forms a complex with an  $\alpha$ -helical peptide containing the CCXXCC FIAsH-EDT<sub>2</sub> recognition site. Although the structure is drawn with the  $i$  and  $i+1$  thiols bridged by one arsenic and the  $i+4$  and  $i+5$  thiols bridged by the other, an alternative conformation, in which one arsenic links  $i$  and  $i+4$  while the other links  $i+1$  and  $i+5$ , is also possible.

Figure taken from Griffin *et al.*, 2000.

nonfluorescent dyes (such as Disperse Blue 3 or Patent Blue V) to block hydrophobic binding sites (Griffin *et al.*, 2000). Adams *et al.* (2002) showed that when the helix-breaking amino acids proline and glycine were inserted between the dicysteine pairs to create a hairpin, the resulting sequence substantially increased the affinity and contrast of FIAsh-EDT<sub>2</sub>-labelled tetracysteine fusion proteins in cells, thus allowing the concentration of di-thiol competitors to be increased without a loss of specific fluorescence. To date, most studies in which FIAsh-EDT<sub>2</sub> labelling has been utilised have involved the construction of recombinant proteins containing the sequence CCPGCC (examples include Hoffmann *et al.*, 2005; Martin *et al.*, 2005; Rudner *et al.*, 2005; Ruthel *et al.*, 2005).

Various derivatives of FIAsh-EDT<sub>2</sub> can be designed by the chemical modification of the original flash compound. For example, a red analogue of flash (termed ReAsH) has been synthesised using the red fluorophore resorufin (Gaietta *et al.*, 2002). Using a combination of FIAsh-EDT<sub>2</sub> and ReAsH, Gaietta *et al.* (2002) determined the mechanism by which connexin43 (Cx43), a subunit of gap junction channels, is added to and removed from gap junction plaques. By engineering the FIAsh-EDT<sub>2</sub>/ReAsH binding motif into Cx43 protein, and then alternatively labelling cells with FIAsh-EDT<sub>2</sub> or ReAsH, different pools of protein could be formed over time. This partitioning between red and green fluorescence revealed novel characteristics of Cx43 transport, assembly into channels and turnover. This work demonstrates the benefits of the technique for studying protein ageing over any time frame. It is highly versatile compared to the technique using green-to-red shifting E5 protein, which matures in a fixed time frame (Miyawaki *et al.*, 2003).

Labelling proteins with FIAsh-EDT<sub>2</sub> would be particularly useful for studying the assembly of the HSV-1 capsid. In addition to the problems regarding the tagging of HSV-1 capsid proteins (discussed above), pulse labelling of HSV-1 capsid proteins in a manner similar to that described by Gaietta *et al.* (2002) might provide an insight into the way in which the structures that make up the capsid, or the capsid as whole, form following the initiation of assembly. Therefore, the aim of the work described in this chapter was to examine the feasibility of using FIAsh-EDT<sub>2</sub> to label the HSV-1 major capsid protein, VP5.

## 6.2. Insertion of the CCPGCC motif into the UL19 open reading frame

Stephen Webster produced a panel of 19 insertional mutants for UL19; each of these contained a randomly inserted 12 bp oligonucleotide, which included a recognition sequence for *Bgl*III. The target DNA was the mammalian expression plasmid, pE19 (Nicholson *et al.*, 1994), which contains the UL19 ORF under the control of the HCMV major immediate-early promoter (Stow *et al.*, 1993). The insertion of the oligonucleotide resulted in the incorporation of an extra four amino acids in frame with the VP5 coding sequence. The mutants were tested for their ability to support virus growth, capsid assembly, DNA packaging, and for interactions with other capsid proteins. Eleven of the mutants complemented growth of a VP5 null mutant virus (K5ΔZ) (Webster, 2004). As these mutants indicated points at which the VP5 sequence was able to tolerate the insertion of four extra amino acids, they were considered to be suitable candidates for the insertion of further genetic information, namely the FIAsh-EDT<sub>2</sub> recognition sequence.

Insertion of FIAsh oligonucleotides was initially attempted on two of Stephen Webster's VP5 insertional mutants: in754 and in1235. Subsequently, following the publication of the 3D structure of the VP5 upper domain (Bowman *et al.*, 2003), it was possible to identify the locations of some of the points of insertion in the folded protein (Figure 6.2). The 3D structure also showed that in two more of Stephen Webster's mutants, in844 and in848, the insertion of extra amino acids took place at the top of the VP5 upper domain, away from alpha helices. These two mutants were considered to be particularly good candidates for the insertion of the FIAsh sequence – both were able to complement the growth of HSV-1 deletional mutant K5ΔZ, and as the amino acids inserted would be likely to lie on the surface of the folded protein, the FIAsh recognition sequence should be readily accessible to FIAsh-EDT<sub>2</sub>.

Insertional mutant in1235 was digested with *Bgl*III. Oligonucleotides VP5 41C FLA-S and VP5 41C FLA-A (Table 6.1) were annealed together and ligated into the sequence at the *Bgl*III site as described in Section 2.2.2.5. Each FIAsh oligonucleotide contains a *Sma*I recognition site so its presence in the DNA sequence was confirmed by digestion with *Sma*I. Digesting UL19in1235 with *Sma*I produces bands of 80 bp, 340 bp, 478 bp, 662 bp, 1057 bp, and 1355 bp, while digesting UL19in1235-Flash results in the 340 bp band being replaced with bands of 107 bp and 257 bp (Figure 6.3). This change in band pattern confirmed the insertion of the FIAsh oligonucleotides; however it did not confirm the

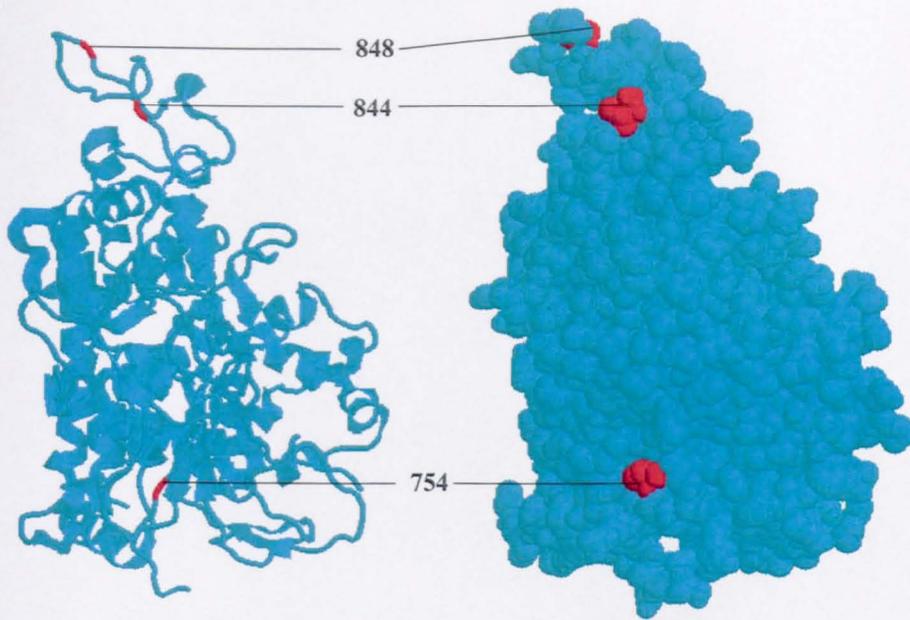


Figure 6.2: Positions in the VP5 upper domain at which extra amino acids were added for the introduction of FIASH-EDT<sub>2</sub> recognition sequence.

Figures were prepared in Protein Explorer 2.45 Beta (<http://www.umass.edu/microbio/chime/pe/protexpl/frntdoor.htm>), using coordinates from the RCSB Protein Databank (<http://www.rcsb.org>) (Berman *et al.*, 2000). Positions at which extra amino acids were added are shown in red. The amino acid 1235 is located outwith the upper domain of VP5, and therefore is not shown here.

UL19 Mutant	FIAsh Oligonucleotides Used	Oligonucleotide Sequence	Amino Acids Inserted	Fragments from <i>Sma</i> I Digest Before Oligonucleotide Insertion	Fragments from <i>Sma</i> I Digest After Oligonucleotide Insertion
in754	VP5 41C FLA-S VP5 41C FLA-A	5' GATCTGTGCTGCCCCGGGTGCTGCCGA 3' 3' ACACGACGGGGCCCACGACGGCTCTAG 5'	CCPGCCEI	80, 90, 238, 478, 662, 1057, <b>1367</b>	80, 90, 238, 478, <b>624</b> , 662, <b>770</b> , 1057
in844	VP5 11A FLA-S VP5 11A FLA-A	5' GATCTGTGCTGCCCCGGGTGCTGCCGA 3' 3' ACGACGGGGCCCACGACGGCTCTAG 5'	CCPGCCRDL	80, 90, 238, 478, 662, 1057, <b>1367</b>	80, 90, 238, 478, <b>497</b> , 662, <b>894</b> , 1057
in848	VP5 56C FLA-S VP5 56C FLA-A	5' GATCTTGCTGCCCCGGGTGCTGCA 3' 3' AACGACGGGGCCCACGACGTCTAG 5'	CCPGCCRS	80, 90, 238, 478, 662, 1057, <b>1367</b>	80, 90, 238, 478, <b>609</b> , 662, <b>782</b> , 1057
in1235	VP5 41C FLA-S VP5 41C FLA-A	5' GATCTGTGCTGCCCCGGGTGCTGCCGA 3' 3' ACACGACGGGGCCCACGACGGCTCTAG 5'	CCPGCCEI	80, <b>340</b> , 478, 662, 1057, 1355	80, <b>107</b> , <b>257</b> , 478, 662, 1057, 1355

Table 6.1: Oligonucleotides used for insertion of the FIAsh sequence into VP5 insertion mutants. Fragments sizes from *Sma*I digests that are altered following insertion of the FIAsh oligonucleotide are shown in red.

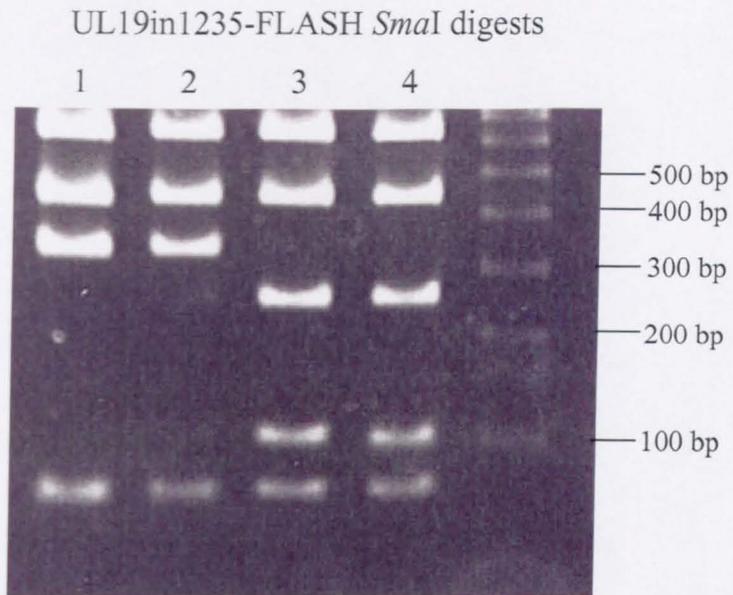


Figure 6.3: *Sma*I digests were used to identify clones with containing the FLASH oligonucleotide.

Digestion of UL19in1235 produces bands of 80 bp, 340 bp, 478 bp, 662 bp, 1057 bp, and 1355 bp. Digestion of UL19in1235-FLASH results in the 340 bp band being replaced with bands of 107 bp and 257 bp. Therefore, the FLASH oligonucleotide has inserted into UL19in1235 in samples 3 and 4, but not in samples 1 and 2.

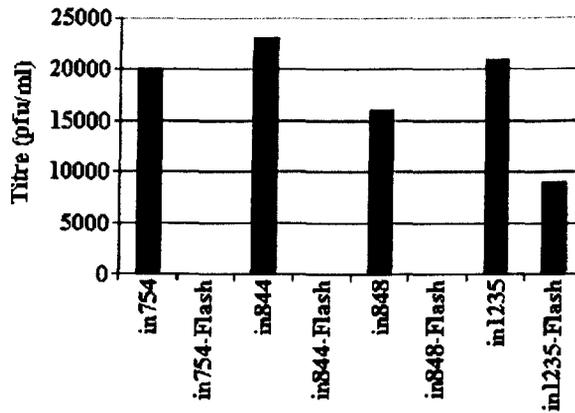
orientation in which the oligonucleotides had been inserted. If the oligonucleotides had inserted in the wrong orientation, the inserted sequence would not code for the desired amino acids. Insertion in the correct orientation was confirmed by sequencing.

A similar process was carried out for the insertion of FIAsh oligonucleotides into UL19in754, UL19in844, and UL19in848, although FIAsh oligonucleotides used and the band pattern expected following the insertion of the oligonucleotides varied in each case (Table 6.1).

Following the insertion of the FIAsh oligonucleotides into the DNA sequence, the resulting plasmids (UL19in754-Flash, UL19in844-Flash, UL19in848Flash, and UL19in1235-Flash) were tested for their ability to complement the growth of the HSV-1 UL19 deletional mutant K5ΔZ. Each mutant was transfected into BHK-21 cells, which were incubated at 37°C for five hours before being infected with 2 pfu/cell K5ΔZ. After one hour, the unabsorbed input virus was harvested and titrated on UL19RSC cells (Section 2.2.4.7). Of the four flash mutants, only UL19in1235-Flash was capable of complementing K5ΔZ (albeit at a reduced efficiency compared with UL19in1235) (Figure 6.4). Therefore, of the four flash mutants, only UL19in1235-Flash is a suitable candidate for FIAsh-EDT<sub>2</sub> labelling experiments.

### 6.3. Labelling with FIAsh-EDT<sub>2</sub>

The ability of FIAsh-EDT<sub>2</sub> to bind to UL19in1235-Flash and produce fluorescence under a confocal microscope was examined. BHK-21 cells on coverslips were co-transfected with UL19in1235-Flash, either alone or along with pJK2, which expresses UL26.5 (Nicholson *et al.*, 1994). 24 hours after transfection, FIAsh-EDT<sub>2</sub> labelling was used to determine the cellular distribution VP5. FIAsh-EDT<sub>2</sub> labelling was carried out in accordance with instructions provided by Guido Gaietta (personal communication). The cells were washed three times with PBS supplemented with 1 g/litre D+ glucose. For each coverslip, 1 μl of 25 mM EDT was added to 2 μl of 1 μM FIAsh. This mixture was incubated at room temperature for ten minutes – a step that ensures that all FIAsh is in FIAsh-EDT<sub>2</sub> form. 200 μl of PBS/glucose was then added to the tube containing FIAsh-EDT<sub>2</sub>. Cells were then overlaid with the resulting mixture and incubated for one hour at 37°C. Following incubation, cells were washed five times in PBS/glucose containing 0.5 mM British Anti-Lewisite (BAL) and 0.1% DMSO, before fixation with 4% formaldehyde. As a control, the



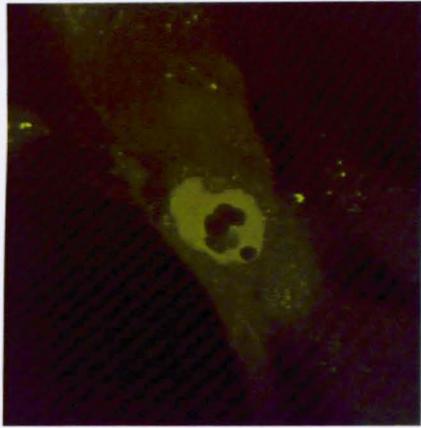
**Figure 6.4:** Abilities of VP5 mutants to complement VP5 null mutant virus with and without the insertion of the FIAsh oligonucleotide..

Complementation of growth of the VP5 null mutant K5ΔZ by transfected plasmids expressing the VP5 site-directed mutants was carried out in BHK-21 cells. The progeny virus was titrated on UL38RSC cells. The titres of the viruses tested are shown.

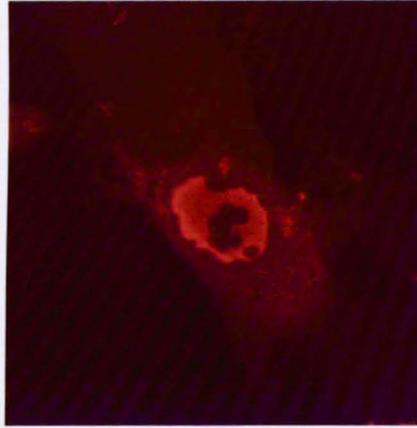
distribution of VP5 in cells was also determined by immunofluorescence using anti-VP5 antiserum (Section 2.2.4.5).

In cells transfected with UL19in1235-Flash and UL26.5, fluorescence apparently produced from the binding of FIAsh-EDT<sub>2</sub> was observed primarily in the nucleus (Figure 6.5). In the absence of UL26.5, fluorescence was confined to the cytoplasm (Figure 6.6). These observations are in accordance with the expected distributions of VP5 in the presence and absence of preVP22a. Furthermore, fluorescence from the anti-VP5 antibody essentially co-localised with the FIAsh-EDT<sub>2</sub> fluorescence (Figures 6.5 and 6.6). Thus, it appeared that the FIAsh-EDT<sub>2</sub> was selectively binding to its recognition sequence. In general, FIAsh-EDT<sub>2</sub> fluorescence observed in the cytoplasm appeared brighter than that observed in the nucleus. This may have been due to FIAsh-EDT<sub>2</sub> being unable to efficiently penetrate the nuclear membrane, or to a reduced ability of FIAsh-EDT<sub>2</sub> to bind to UL19in1235-Flash when it is complexed with preVP22a. However, in cells where VP5 was expected to be localised to the nucleus, a background level of fluorescence was also visible in the cytoplasm. It has been suggested (Gaietta, personal communication) that the cause of this was insufficient levels of BAL during the washing stages (Section 2.2.8). BAL removes unbound FIAsh-EDT<sub>2</sub> from cells, and it may be that a reduced concentration of BAL in the washing solution could result in this being done inadequately. The experiment was repeated, varying the concentrations of BAL in the washing solution (ranging from 0.3 mM to 1.0 mM). However, no tested concentration of BAL was able to alter levels of background fluorescence. Trafficking experiments carried out on the HSV-1 capsid proteins require low levels of background fluorescence in order to identify the locations of the components of the capsid during assembly, therefore in the absence of further improvements to the levels of background fluorescence it was not felt that labelling with FIAsh-EDT<sub>2</sub> was a suitable method for studying the HSV-1 capsid, as discussed in Section 8.1.

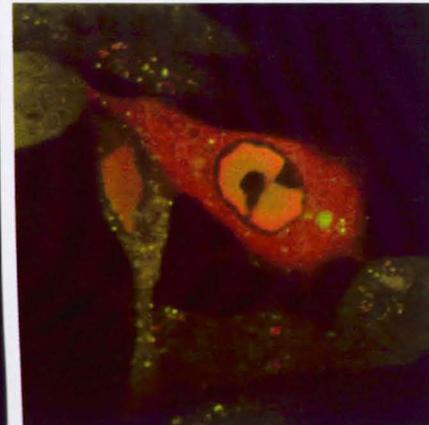
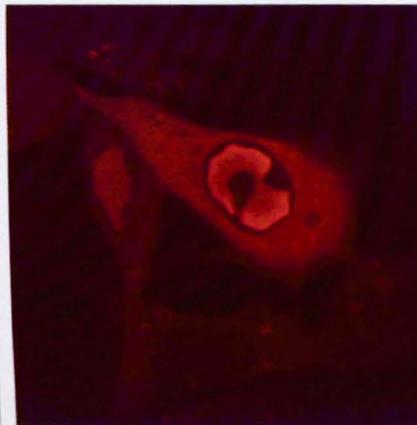
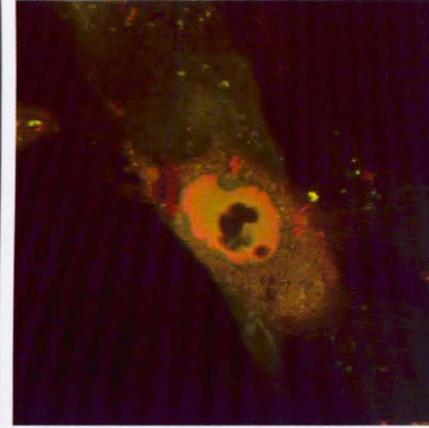
**FLAsH-EDT<sub>2</sub>  
labelled**



**VP5 antibody  
labelled**



**Co-localisation of  
VP5 antibody and  
FLAsH-EDT<sub>2</sub>**



**Figure 6.5:** FIAsh-EDT<sub>2</sub> labelling of cells expressing UL19in1235-Flash and preVP22a. BHK-21 cells were cotransfected with UL19in1235-Flash and pJK2 (expressing VP22a). The distribution of the CCPGCC-tagged VP5 was detected by FIAsh-EDT<sub>2</sub> labelling. The fluorescence attributed to FIAsh-EDT<sub>2</sub> labelling (green) was compared to that obtained by immunofluorescence by also labelling CCPGCC-tagged VP5 with the antiserum rAb184 and visualising using TRITC-conjugated GAR (red).

**FIAsH-EDT<sub>2</sub>  
labelled**



**VP5 antibody  
labelled**



**Co-localisation  
of VP5 antibody  
and FIAsH-EDT<sub>2</sub>**

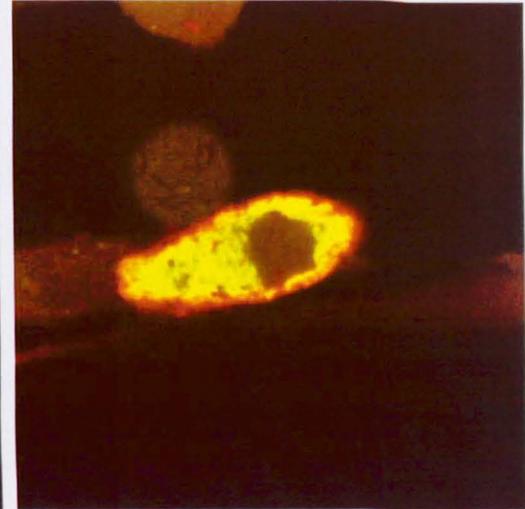


Figure 6.6: FIAsH-EDT<sub>2</sub> labelling of cells transfected with UL19in1235-Flash.

BHK-21 cells were transfected with UL19in1235-Flash. CCPGCC-tagged VP5 was detected by FIAsH-EDT<sub>2</sub> labelling. The fluorescence attributed to FIAsH-EDT<sub>2</sub> labelling (green) was compared to that obtained by immunofluorescence by also labelling CCPGCC-tagged VP5 with the antiserum rAb184 and visualising using TRITC-conjugated GAR (red).

## **CHAPTER 7**

### **SITE-DIRECTED MUTAGENESIS OF VP5**

## 7. Site-directed mutagenesis of VP5

### 7.1. Introduction

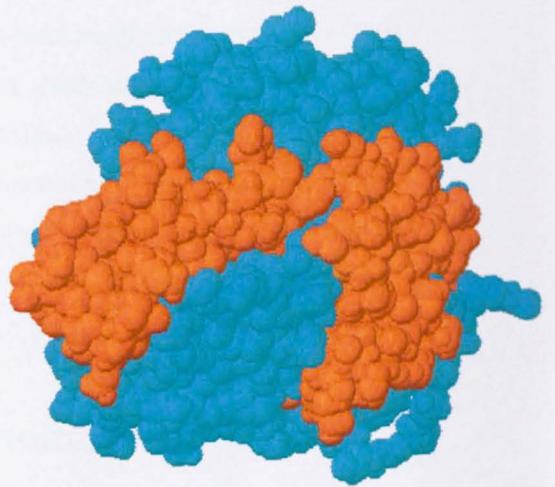
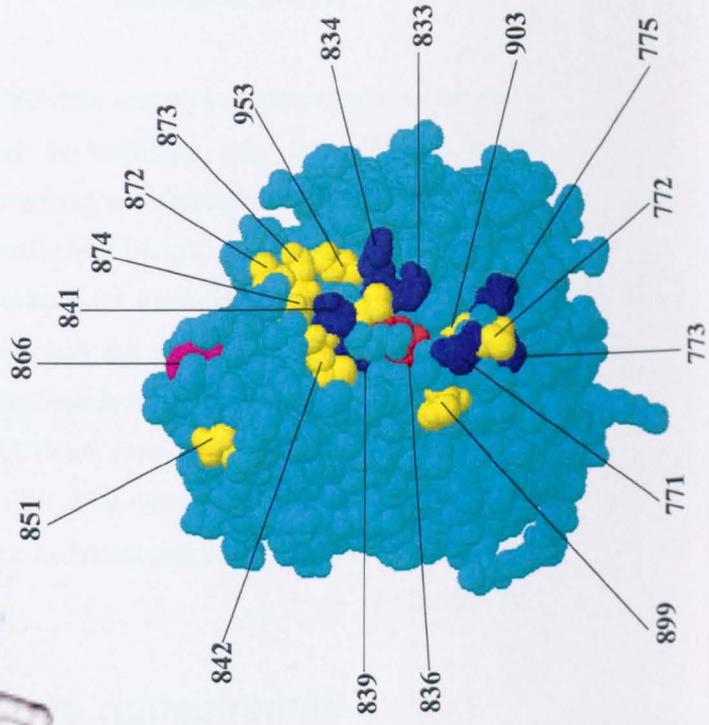
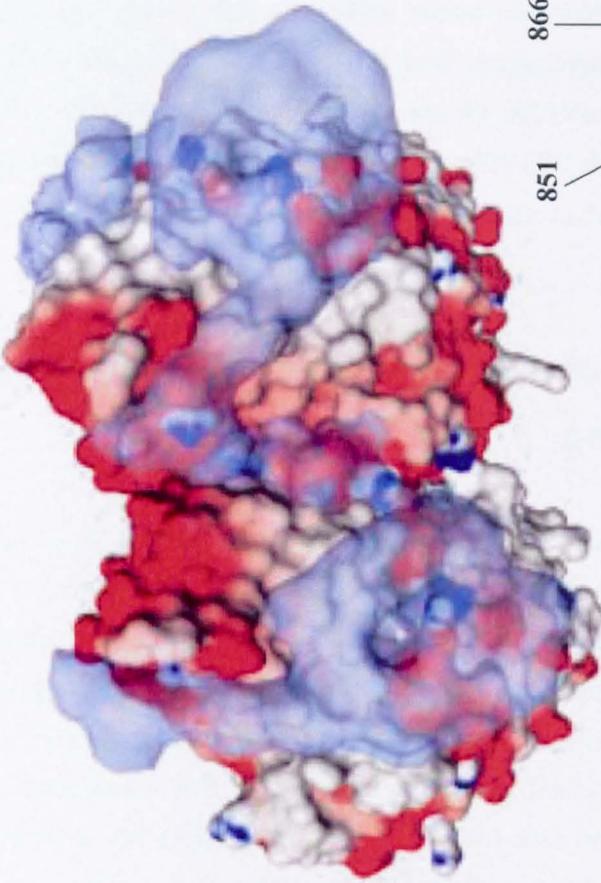
The major capsid protein interacts with VP26 at the tips of hexons (Trus *et al.*, 1995; Zhou *et al.*, 1995), and with tegument at the tips of pentons (Zhou *et al.*, 1999). Our understanding of these interactions has been aided by the resolution of the crystal structure of the VP5 upper domain (Figure 1.16) (Bowman *et al.*, 2003). Since the publication of the crystal structure, a combination of computational and experimental approaches has been used in an attempt to identify the particular amino acids involved in these interactions. Previously, single amino acid substitutions have been used to identify residues in VP5 that are important for interaction with preVP22a (Walters *et al.*, 2003) and residues in VP26 that are important for interaction with VP5 (Desai *et al.*, 2003). Using a similar approach, residues of VP5 that were suggested as being important for interaction with VP26 were examined.

### 7.2. Identification of VP5 residues that might have a role in VP26 binding

Superimposition of the VP5 upper domain (VP5ud) crystal structure onto the 8.5Å capsid reconstruction identified positions in the VP5ud that appeared to be in close contact with VP26. This identified six amino acids near the top of VP5 that represented good candidates for involvement in the interaction (Figure 7.1): Residues 833 (aspartic acid), 834 (arginine), 836 (tyrosine), 837 (alanine), 839 (leucine), 841 (asparagine) and 851 (glutamic acid) (Matthew Baker, personal communication). In order to test this hypothesis, mutagenesis of these amino acids was carried out, and each resulting mutant was assayed for its ability to interact with VP26. Each mutant produced contained a single amino acid change, with the candidate amino acid substituted by an alanine. Since amino acid 837 is already an alanine it was not mutated in this study.

### 7.3. Introduction of an *Mlu*I site into the UL19 open reading frame

PCR was used to generate portions of the UL19 ORF containing the desired mutations (Figure 7.2), and the resulting sequences were cloned into the full length UL19 ORF. The



**Figure 7.1:** The VP5 upper domain.

**Top:** View from above of the upper domains of two hexon VP5s bound to two VP26 molecules (blue-grey semi-transparent). In the VP5 upper domain, acidic residues are coloured red, basic residues are coloured blue and uncharged residues are coloured grey. (Figure reproduced from Bowman *et al.*, 2003).

**Bottom left:** The right-hand VP5 upper domain from the top figure, with residues that can be altered by PCR mutagenesis following the construction of pWA1 shown in orange.

**Bottom right:** Residues suggested by Matthew Baker (personal communication) as having a role in VP26 or tegument binding. Residues that were substituted for alanine in the work described here are shown in blue, red and magenta. Amino acids shown in blue were suggested as having a role in VP26 binding and substitution of alanine did not affect viability. When the tyrosine at position 836 was substituted for alanine (red), no infectious virus was produced. The histidine at position 866 (magenta) was suggested as having a role in tegument binding. Alanine substitutions were not successfully attempted for residues shown in yellow. (Bottom left and bottom right figures were prepared in Protein Explorer 2.45 Beta (<http://www.umass.edu/microbio/chime/pe/protexpl/frntdoor.htm>), using coordinates from the RCSB Protein Databank (<http://www.rcsb.org/pdb/>) (Berman *et al.*, 2000).

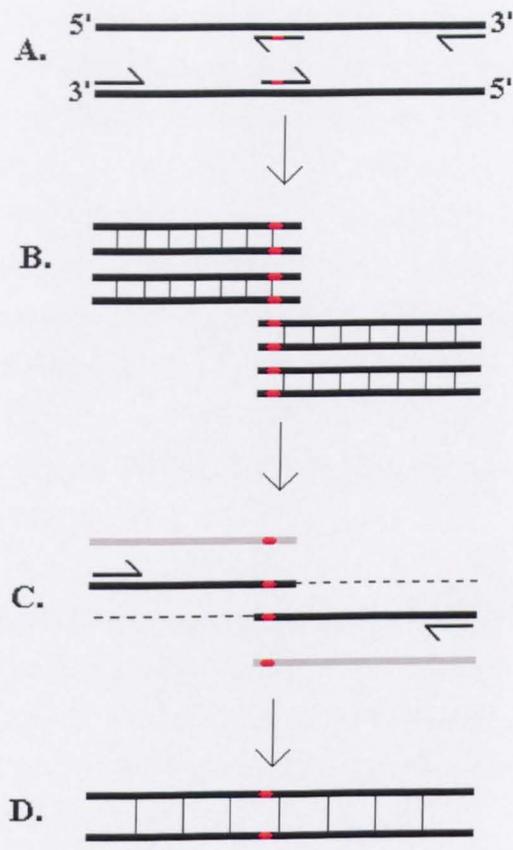


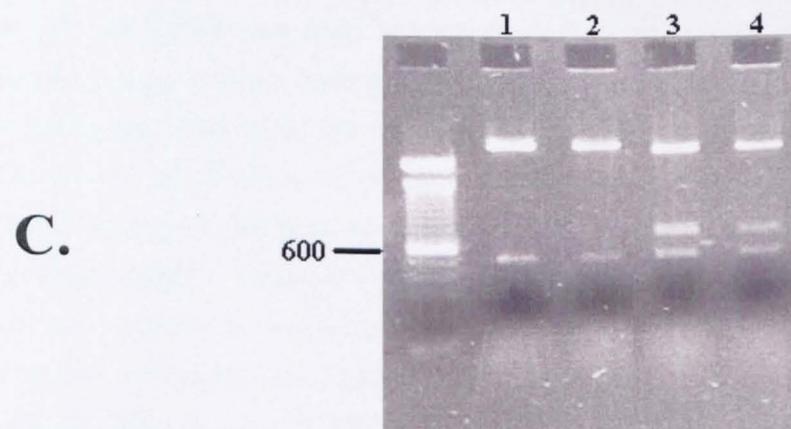
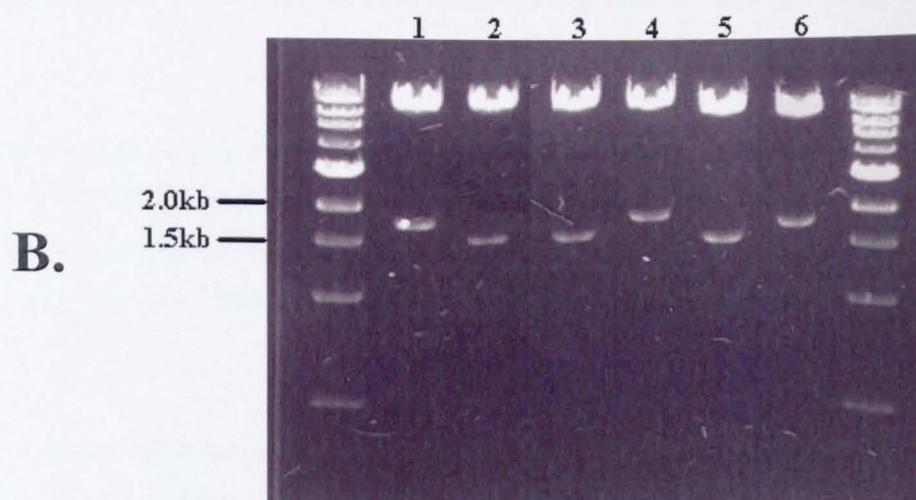
Figure 7.2: PCR method for site-directed mutagenesis

A. PCR was carried out on DNA to be mutagenised. Two reactions were carried out. Each PCR product contained an overlapping region, in which the desired mutation was located. 'Inner' primers contained the desired mutation (denoted by red lines). B. PCR products from the two reactions contained overlapping region and desired mutation. C. The second round of PCR was carried out, using the two products from the first PCRs as template DNA, and the 'outer' primers shown in A. D. The product of the second round of PCR contained the desired mutation. As suitable cloning sites flanked the mutation, the PCR product could then be introduced into the ORF targeted for mutagenesis.

region of the UL19 ORF containing all the amino acids under investigation did not contain many unique recognition sites for restriction enzymes. There is an *NruI* site 159 bp upstream of amino acid 833, while an *FspAI* site is 263 bp downstream of amino acid 851. While *NruI* can be obtained from most suppliers at relatively little cost, *FspAI* proved more difficult to obtain and was considerably more expensive than most currently available restriction enzymes. For this reason, PCR mutagenesis was carried out to introduce a unique restriction enzyme recognition site – an *MluI* site – downstream of the amino acids under investigation. This alteration results in the substitution of a guanine for an adenine in the nucleotide sequence; however, it does not alter the protein sequence, with amino acid 882 of VP5 remaining a glycine. The introduction of this *MluI* site in this region allowed mutagenesis of the UL19 ORF between the *NruI* site (2255 bp from the start of UL19 – amino acid 752) and the *MluI* site (2646 bp from the start of UL19 – amino acid 882) as described below (Figure 7.1).

Analysis carried out by Matthew Baker on the three-dimensional structure of VP5 had also suggested that amino acid 866 (histidine) lies in very close proximity to the site of tegument attachment in pentons (personal communication). Since amino acid 866 lies in the region that can be examined by this method, the study was extended to include mutagenesis of this amino acid.

Attempts to carry out mutational PCR using the Quickchange mutagenesis kit (Stratagene) were unsuccessful due to failure of the PCR reaction, possibly caused by the large difference in GC content between the vector and insert sequences (Frazer Rixon, personal communication). Therefore, an alternative mutagenesis method was used as illustrated in Figure 7.2. Frazer Rixon had previously cloned the UL19 ORF into pFASTBAC pCI. pFASTBAC pCI (provided by Roger Everett) is a derivative of the pFASTBAC baculovirus transfer vector (Invitrogen) that was modified to include the HCMV IE promoter for expression in mammalian cells. To introduce a unique *MluI* site into the UL19 ORF, two separate PCR reactions were carried out. In reaction 1 the primers used were UL19 UO SEN2 and UL19 MLU ANTI, while in reaction 2 the primers used were UL19 MLU SENS and UL19 DOWN ANT (Table 7.1). UL19 MLU SENS and UL19 MLU ANTI are complementary oligonucleotides which when annealed result in a base change (G to A at position 2646 in the UL19 sequence). The expected PCR product from reaction 1 was 497 bp long, while reaction 2 was expected to produce a 218 bp product (Figure 7.3A). Bands of appropriate sizes were identified, extracted from 1xTAE gel and purified by Sephaglas purification. A PCR reaction was then set up using a mixture of the



**Figure 7.3:** Gel photographs from cloning steps during construction of VP5 site-directed mutants.

**A.** (Lane 1) PCR product obtained from reaction 1 (wild-type UL19 ORF was used as a template, UL19 UO SEN2 and UL19 MLU ANTI were used as primers. The expected PCR product from reaction 1 was 497 bp long. (Lanes 2, 3, and 4) PCR product obtained from reaction 2 (wild-type UL19 ORF was used as a template, UL19 MLU SENS and UL19 DOWN ANT were used as primers). The expected PCR product from reaction 2 was 218 bp long. Details of primers are shown in Table 7.1. Band sizes were estimated using a 100 bp ladder (New England Biolabs). Selected band sizes (in numbers of base pairs) in the 100 bp ladder are indicated.

**B.** Digestion with both *NruI* and *FspAI* produces blunt ends, therefore the mutated section of the UL19 ORF could potentially insert in either orientation. Insertion in the correct orientation could be confirmed by the production of bands of 1537 bp and 8580 bp from an *MluI/XbaI* digest (as seen in lanes 2, 3, and 5). Insertion in the incorrect orientation produces bands of 1795 bp and 8322 bp (as seen in lanes 1, 4, and 6). Band sizes were estimated using a 1 kb ladder (New England Biolabs). Selected band sizes in the 1 kb ladder are indicated.

**C.** (Lanes 1 and 2) Digestion of pGEMT UL19MluI with *HincII* should produce bands of either 499 bp and 3216 bp, or 265 bp and 3450 bp, depending on which orientation the PCR fragment inserted into pGEMT Easy. Both lanes 1 and 2 show clones with inserts in the first orientation. (Lanes 3 and 4) In pGEMT UL19MluI clones which produced the 499bp/3126bp digest pattern from *HincII* digestion, digestion with *PvuII* should produced fragments of 469 bp, 682 bp, and 2564 bp. Band sizes were estimated using a 100 bp ladder (New England Biolabs). The position of the 600 bp band is indicated.

products of reactions 1 and 2 as template DNA, while the primers were UL19UO SEN2 and UL19 DOWN ANT. The expected PCR product from this reaction was 693 bp long. A

Primer ID	Primer Sequence
UL19 UO SEN2	5' GAGCTGAATCACCTAATGCGAGACC 3'
UL19 SDM UPSTREAM 3	5' CATAACCGGCAGCGAGCACGTC 3'
UL19 DOWN ANT	5' GCCATTTTGGATGGTATGGTCCAG 3'
UL19 MLU SENS	5' CACAACGGACGCGTGGTAGTG 3'
UL19 MLU ANTI	5' CACTACCACGCGTCCGTTGTG 3'
V771 SENSE	5' GCATGTAACGGCGGACCGCGGAC 3'
V771 ANTI	5' GTCCGCGGTTCGCGCGTTACATGC 3'
T773 SENSE	5' GTAACGTGGCGGCCGCGGACTTC 3'
T773 ANTI	5' GAAGTCCGCGGCCCGCCACGTTAC 3'
D775 SENSE	5' GTGGCGACCGCGGCCCTTCAACC 3'
D775 ANTI	5' GGTGAAAGGCGCGGTTCGCCAC 3'
D833 SENSE	5' GTTCGCTTCGCCCGGTATACGC 3'
D833 ANTI	5' GCGTATACGCGGGCGAAGCGAAC 3'
R834 SENSE	5' GTTCGCTTCGACGCCGTATACGCC 3'
R834 ANTI	5' GGCATATACGGCGTGAAGCGAAC 3'
Y836 SENSE	5' CGACCGCGTAGCCGCCACCCTG 3'
Y836 ANTI	5' CAGGGTGGCGGCTACGCGGTTCG 3'
L839 SENSE	5' GTATACGCCACCCTCAGAACATGGTG 3'
L839 ANTI	5' CACCATGTTCTGAGCGGTGGCGTATAC 3'
N841 SENSE	5' GCCACCCTGCAGGCCATGGTGGTC 3'
N841 ANTI	5' GACCACCATGGCCTGCAGGGTGGC 3'
H866 SENSE	5' CACCCCTGCCCCCGCCAATCTG 3'
H866 ANTI	5' CAGATTGGCCGGGCCAGGGGGTG 3'

Table 7.1: Primers used in cloning UL19 site-directed mutants.

Bases in which the DNA sequence differs from the wild-type ORF are shown in red.

### 7.4. Construction of site-directed mutants

To mutate the required codons in the UL19 ORF, site-directed PCR reactions, similar to that described above, were carried out using the primers described in Table 7.1. The resulting sequences were ligated into pGEMT100 to produce pGEMT834, pGEMT836, pGEMT839, pGEMT841, pGEMT843, and pGEMT845. The presence of the mutations in the resulting plasmids was confirmed by DNA sequencing. The pGEMT834, pGEMT836, pGEMT839, pGEMT841, and pGEMT845 all contained the expected sequences. However, pGEMT843 contained a single base change, which resulted in a frameshift. Unfortunately, the loss of a base in the primer did not allow the construction of pGEMT831 to be repeated.

Each of the pGEMT plasmids described above was digested with *Xba*I and *Kpn*I. In each case this produces two fragments of 391 bp and 906 bp. The 391 bp fragment was

products of reactions 1 and 2 as template DNA, while the primers were UL19UO SEN2 and UL19 DOWN ANT. The expected PCR product from this reaction was 695 bp long. A band of the appropriate size was identified, extracted from 1xTAE gel and purified by Sephaglas purification. The purified band was ligated into pGEMT Easy (Promega) to obtain pGEMT UL19MluI. The presence of the mutated section of the UL19 ORF in pGEMT Easy was confirmed by digestion with *EcoRI* (which was expected to produce a band of 718 bp), *HincII* (either 265 bp and 3450 bp, or 499 bp and 3216 bp, depending on the orientation that the PCR product inserted into pGEMT Easy), and *PvuII* (either 358 bp, 793 bp, and 2564 bp or 469 bp, 682 bp, and 2564 bp) (Figure 7.3C). The presence of a mutated section of the UL19 ORF in the plasmid was confirmed by sequencing.

Following sequencing, pGEMT UL19MluI and 1C7/pC1 were digested with *NruI* and *FspAI*, and bands of respectively 524 bp and 9593 bp were extracted from the two digests. These bands were ligated to produce pWA1, a pFASTBAC clone of UL19 with a single base change that introduced a unique *MluI* site. As digestion with both *NruI* and *FspAI* produces blunt ends, the mutated section of the UL19 ORF could potentially insert in either orientation. To check this, *MluI/XbaI* digests were carried out. Insertion in the correct orientation could be confirmed by the production of bands of 1537 bp and 8580 bp from this digest (Figure 7.3B). A clone showing the correct digestion pattern was selected for further use and designated pWA1. The sequence of pWA1 was confirmed by DNA sequencing.

#### 7.4. Construction of site-directed mutants

To mutate the required codons in the UL19 ORF, two-stage PCR reactions, similar to that described above, were carried out using the primers described in Table 7.1. The resulting sequences were ligated into pGEMT Easy to produce pGEMT833, pGEMT834, pGEMT836, pGEMT839, pGEMT841, pGEMT851, and pGEMT866. The presence of the mutations in the resulting plasmids was analysed by DNA sequencing. pGEMT833, pGEMT834, pGEMT836, pGEMT839, pGEMT841, and pGEMT866 contained the expected sequences. However, pGEMT851 contained a single base insertion, which resulted in a frameshift. Unfortunately the time constraints of this project did not allow the construction of pGEMT851 to be repeated.

Each of the pGEMT plasmids described above was digested with *NruI* and *MluI*. In each case this produces two fragments of 391 bp and 3324 bp. The 391 bp fragment was

isolated and purified. Digestion of pWA1 with *NruI* and *MluI* generated two bands of 391 bp and 9726 bp. The 9726 bp band was isolated and purified. The purified 391 bp and 9726 bp bands were ligated together to produce pWA833, pWA834, pWA836, pWA839, pWA841, and pWA866.

## **7.5. Identification of further VP5 residues that might have a role in VP26 binding**

Shortly after construction of these mutants, a refined analysis of amino acids of VP5 likely to be involved in the interaction with VP26 became available (Matthew Baker, personal communication). Bioinformatic analysis of VP26 and the VP5ud suggested that there are three primary regions of contact between VP5 and VP26. In VP26 the N- and C-termini contain high concentrations of charged amino acids, and estimates obtained from bioinformatic approaches suggest that the areas of VP5 in close contact with these regions of VP26 also contain charged residues. It appears likely that residues 834 (arginine) and 953 (aspartic acid) are both in close proximity to the N-terminal arm of VP26, while 775 (aspartic acid), 833 (aspartic acid), 851 (glutamic acid), 899 (histidine), and 903 (glutamic acid) are in close proximity to the C-terminal arm. VP5 also appears to have two loops in close proximity to VP26, with the positions of amino acids 771 (valine), 772 (alanine), 773 (threonine), 841 (asparagine), 842 (methionine), 872 (alanine), 873 (asparagine), and 874 (threonine) making them possible candidates for VP26 binding. Residues suggested for site-directed mutagenesis are summarised in Figure 7.1.

As described above, construction of pWA833, pWA834, and pWA841 had already been carried out. Amino acids 772 and 872 were alanine, and therefore not suitable for this method of analysis, while amino acids 899, 903, and 953 were upstream of the *MluI* site and would therefore require an alternative cloning strategy for mutagenesis. Therefore, for this analysis of the interaction between VP5 and VP26, mutagenesis resulting in amino acid substitutions at positions 771, 773, 775, 842, 873, and 874 was attempted. Cloning was carried out as described above, with primers used for PCR mutagenesis as listed in Table 7.1. When the resulting DNAs were sequenced, many were shown to contain unexpected single base substitutions and/or insertions, suggesting that the extension temperature during the PCR might have been too low; this could have resulted in the incorporation of erroneous bases in the sequence. Unfortunately, there was not enough time to repeat the cloning; nevertheless, full-length UL19 ORFs containing bases coding

for amino acid substitutions at positions 771, 773, and 775 (pWA771, pWA773, and pWA775 respectively) were produced.

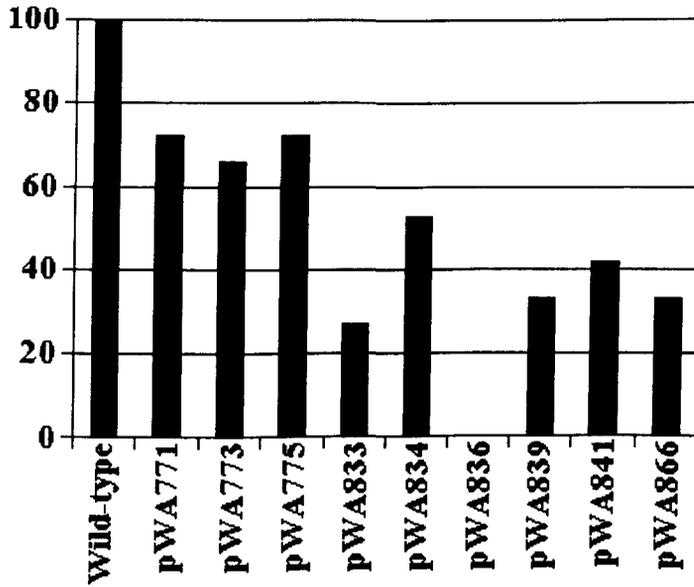
## **7.6. Analysis of site-directed mutants**

### **7.6.1. Complementation**

To analyse the effects of the mutations on the properties of VP5, their ability to complement the growth of the HSV-1 deletional mutant K5ΔZ (Desai *et al.*, 1993) was examined. DNA from each site-directed mutant was transfected into BHK-21 cells. These were incubated at 37°C for five hours before being infected with 2 pfu/cell of K5ΔZ. After one hour, the unabsorbed input virus was neutralised by an acid wash. Incubation was continued at 37°C for 40 hours, at which time the progeny virus was harvested and titrated on UL19RSC (Section 2.1.4). Figure 7.4 shows the relative yield obtained with each of the mutant plasmids compared to that obtained with the wild-type control (pE19). With the exception of pWA836, all the mutants retained the ability to support virus growth to some degree, with the mutants in amino acids 771-775 producing slightly higher average yields (66-72% of the VP5 wild-type control level) than mutants in amino acids 833-866 (27-53%).

### **7.6.2. Interaction with VP19C and preVP22a**

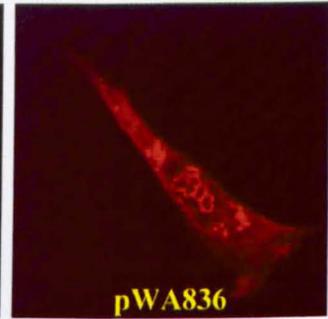
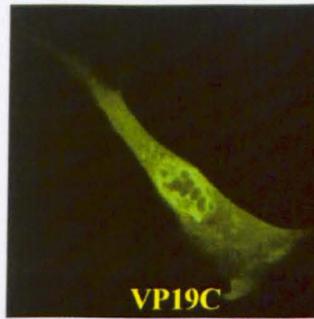
In order to ensure the formation of functional capsids, and hence infectious virions, VP5 must interact correctly with VP19C and preVP22a. The presence of either of these proteins is sufficient to transport VP5 to the nucleus. In the absence of both proteins, VP5 is distributed throughout the cell (Rixon *et al.*, 1996). Since these interactions are essential, they can be assumed to have formed correctly in the VP5 mutants that complemented K5ΔZ. To examine these interactions in pWA836, it was co-transfected into BHK-21 cells along with either pE38 (Rixon *et al.*, 1996), which expresses VP19C, or pJK2, which expresses preVP22a (Nicholson *et al.*, 1994). The distributions of the proteins were determined by immunofluorescence (Section 2.2.4.5). These co-transfections demonstrated that the mutation in pWA836 did not affect interactions with VP19C and preVP22a (Figure 7.5).



**Figure 7.4:** Abilities of VP5 site-directed mutants to complement VP5 null mutant virus.

Complementation of growth of the VP5 null mutant K5ΔZ by transfected plasmids expressing the VP5 site-directed mutants was carried out in BHK-21 cells. The progeny virus was titrated on UL38RSC cells. The results are shown as percentages of the titre obtained with the wild-type VP5 control plasmid.

**pWA836/VP19C  
co-transfection**



**pWA836/preVP22a  
co-transfection**

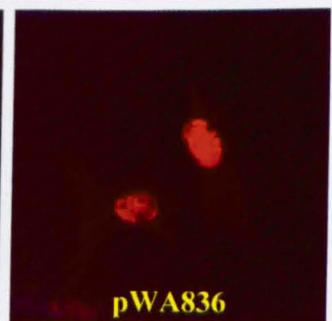
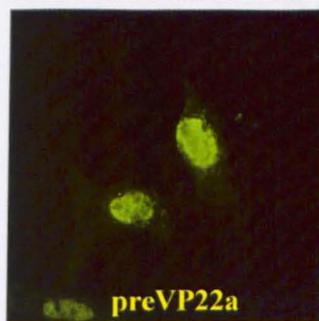


Figure 7.5: Influence of VP19C and preVP22a on the distribution of VP5 insertional mutant pWA836.

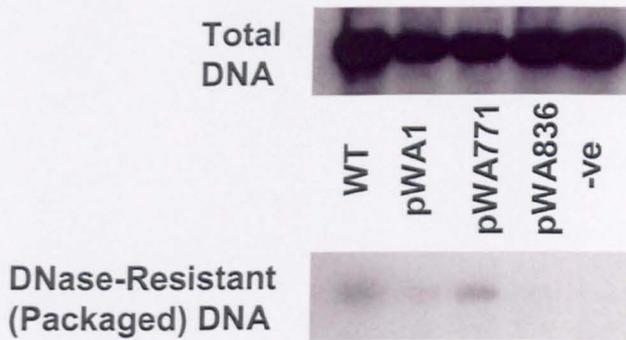
BHK-21 cells were cotransfected with pWA836 and either pUL38FBpCI (which expresses wild-type VP19C), or pJK2 (wild-type preVP22a). VP19C was detected with the monoclonal antibody mAb02040 and visualised using FITC GAM (green). preVP22a was detected with the monoclonal antibody mAbMCA406 and visualised using FITC GAM (green). pWA836 was detected with the antiserum rAb186, and visualised using TRITC-conjugated GAR (red).

### **7.6.3. DNA packaging**

As shown in Figure 7.5, both VP19C and VP22a were able to transport the mutant form of VP5 expressed from pWA836 to the nucleus, indicating that the mutation had not disrupted VP5's binding to either protein. Since successful DNA packaging depends on the presence of functional capsids, the ability of pWA836 to support the packaging of DNA was also examined in a plasmid-based packaging assay. As described in Section 2.2.9, the presence of DNase-resistant DNA in the WT, pWA1, and pWA771 transfected samples (Figure 7.6) demonstrates that DNA packaging had taken place, while in pWA836 the lack of packaged DNA suggests that either the capsid assembly or packaging mechanism were disrupted.

### **7.6.4. Interaction with VP26**

The effects of the VP5 mutations on the interaction between VP5 and VP26 was analysed by immunofluorescence. VP26 does not localise specifically to the nucleus when expressed in the the absence of other viral proteins, but is predominantly nuclear when co-expressed with VP5 and preVP22a (Rixon *et al.*, 1996). To examine the ability of the VP5 mutants to interact with VP26, each mutated plasmid was co-transfected into BHK-21 cells along with pE35T1, which expresses a pp65 epitope tagged form of VP26 (Rixon *et al.*, 1996), and pJK2, which expresses preVP22a (Nicholson *et al.*, 1994). The distributions of the proteins were determined by immunofluorescence (Section 2.2.4.5). VP26 was seen to be located in the nucleus in the presence of preVP22a and each of the VP5 mutants (Figure 7.7), indicating that none of the single amino acid substitutions examined prevented the interaction between VP5 and VP26. This is discussed further in Section 8.2.



**Figure 7.6:** Ability of VP5 site-directed mutants to package DNA.

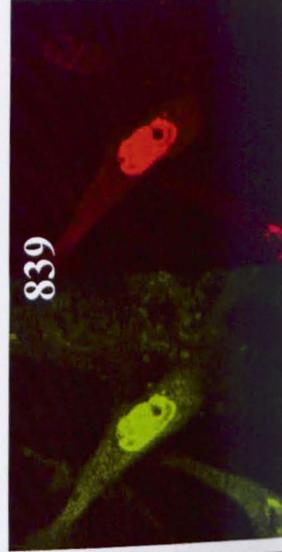
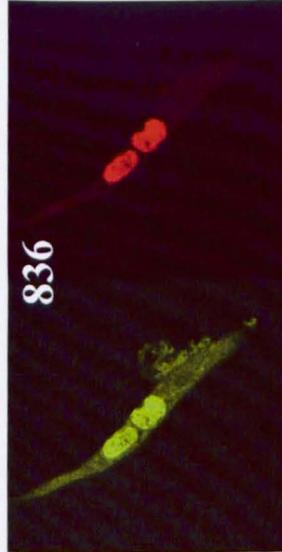
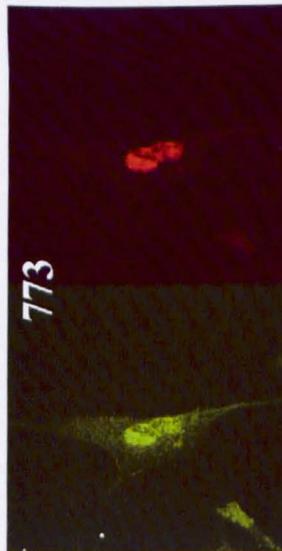
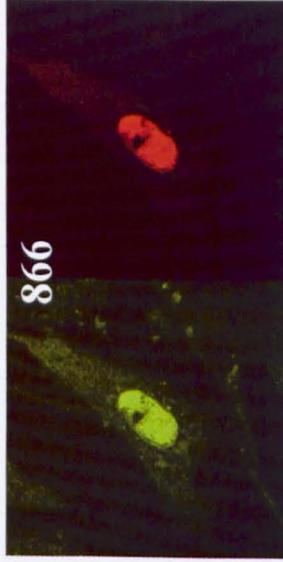
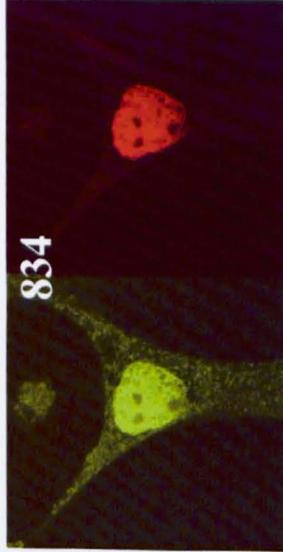
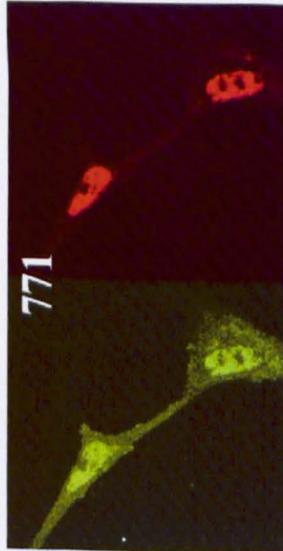
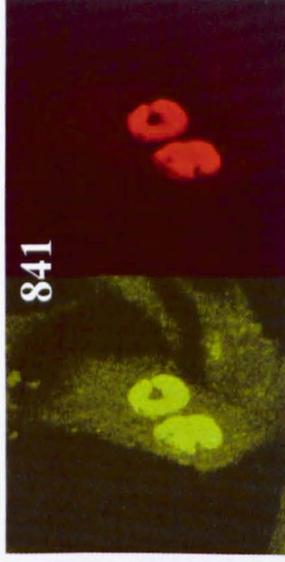
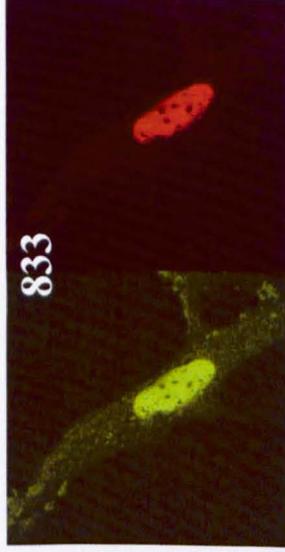
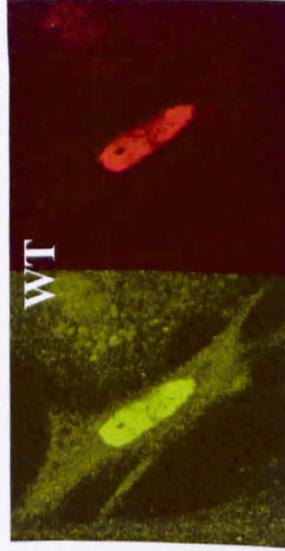
BHK-21 cells were transfected with either pE19 (wild-type (WT) VP5), one of the VP5 site-directed mutants, or a negative (no DNA) control as indicated. Cells were superinfected with K5ΔZ, and total (top) and DNase-resistant (bottom) DNAs were prepared. The DNAs were digested with *Eco*RI and *Dpn*I, and the fragments were separated by electrophoresis through an agarose gel. The gel was blotted, the membrane was hybridised to <sup>32</sup>P-labeled pAT153 DNA and, after being washed, the membrane was exposed to a phosphoimager screen.

VP26

VP26

VP5

VP5



**Figure 7.7:** Influence of VP5 site-directed mutants on the distribution of VP26.

BHK-21 cells were cotransfected with pE35T1 (expressing pp65 epitope tagged form of VP26), pJK2 (preVP22a), and either pE19 (wild-type (WT) VP5), or one of the VP5 site-directed mutants as indicated. VP26 was detected with the monoclonal antibody 9220 (Capricorn Products), and visualised using FITC GAM (green). VP5 was detected with the antiserum rAb184, and visualised using TRITC-conjugated GAR (red).

## **CHAPTER 8**

### **USING MUTAGENESIS TO STUDY VP5 – DISCUSSION**

## 8. Using Mutagenesis to Study VP5 – Discussion

Alterations and additions to the sequence of the HSV-1 major capsid protein have provided an insight into the functional organisation of the protein. Warner *et al.* (2001) randomly mutagenised the N-terminal 255 nucleotides of VP5 using error-prone PCR in order to identify residues that were important for interactions with the scaffold proteins. (The N-terminal region of VP5 had previously been identified as important for interaction with the scaffold proteins (Desai and Person, 1999; Warner *et al.*, 2000)). Webster (2004) generated a panel of insertional mutants of VP5, each of which had an extra four amino acids inserted at a random position in the protein sequence, and examined each mutant's ability to support virus growth, capsid assembly, DNA packaging and to interact with other capsid proteins.

### 8.1. Labelling VP5 with FIAsH-EDT<sub>2</sub>

Tetracysteine motifs were inserted into four positions in the VP5 amino acid sequence. It had previously been shown that at each of these positions it was possible to insert four extra amino acids without disrupting protein function; however, with the exception of the insertion at amino acid 1235, insertions of a further eight or nine amino acids (Table 6.1) resulted in the protein losing its ability to function normally and complement growth of a VP5 deleted virus (K5ΔZ). Examination of the 3D structure of the VP5 upper domain shows that amino acids 754, 844, and 848 are all in loops that are exposed on the surface of the VP5 upper domain. The introduction of hydrophobic residues such as cysteine to an exposed region of a protein might have an affect on the protein's folding and binding specificities.

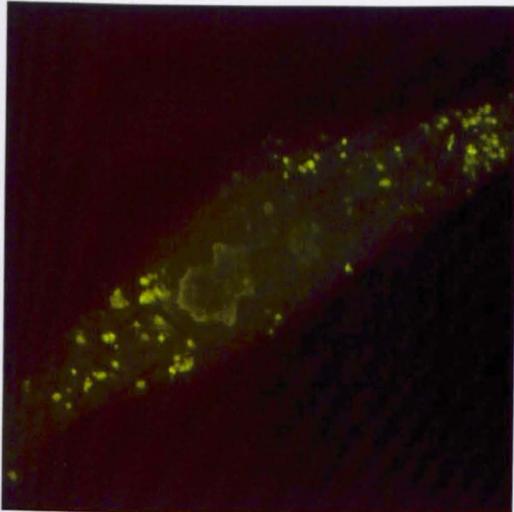
When the sequence CCPGCCEI was inserted into the existing insertion (REIS) in VP5 mutant in1235, it was still able to rescue K5ΔZ (Figure 6.4). In the resulting VP5 mutant, UL19in1235-Flash, a total of twelve extra amino acids had been inserted at amino acid 1235. This observation indicates that this region of the protein is particularly tolerant of insertional mutagenesis; it could suggest that this region is not one that the structure and function of VP5 is particularly dependent on. As this region of the protein is not in the VP5 upper domain, it is not possible to determine the location of amino acid 1235 in the folded protein. Baker *et al.* (2003) used computation analysis to predict the locations of secondary structural elements in the remainder of VP5, and to predict where these secondary structural elements would be located in the folded protein. They predicted that residues

1194-1374 are located in the middle domain of VP5, making up a region known as the VP5 helix bundle with residues from the N-terminal region of the protein. However, the software that they used did not identify any secondary structural elements at or near to amino acid 1235. None of the remaining flash mutants were able to rescue K5ΔZ. It appeared that in each of these cases, although the insertion of four extra amino acids did not affect the mutant protein's ability to complement growth of K5ΔZ, the addition of a further eight or nine amino acids at the same point of insertion rendered VP5 non-functional. This was particularly surprising for in 844 and in848, which are located at the top of the hexon and penton VP5 subunits in a location that had been suggested to be more tolerant of insertions (Webster, 2004), and would suggest that the structure of VP5 is poorly able to tolerate major changes. Some structural proteins are more tolerant to the addition of extra residues than others. For example, in hepatitis B virus (HBV), multiple copies of a 183 amino acid core protein form the nucleocapsid. It has been shown that HBV capsids can still be formed when the entire GFP sequence (238 amino acids) is inserted in a loop between two long  $\alpha$ -helices between amino acids 79 and 80 of the HBV core protein (Kratz *et al.*, 1999).

When BHK-21 cells were transfected with UL19in1235-Flash and VP22a, and then subsequently examined following exposure to FIAsh-EDT<sub>2</sub> and immunofluorescence with a VP5 antibody, fluorescence was observed in similar positions within the nucleus (Figure 6.5). This indicates the insertion of the tetracysteine motif does not disrupt the ability of VP22a to transport VP5 to the nucleus, and that FIAsh-EDT<sub>2</sub> is able to bind to the tetracysteine motif in UL19in1235-Flash with some degree of specificity. When transfections were carried out in the absence of VP22a, the strongest fluorescence appeared in the cytoplasm (Figure 6.6), as would be expected from a transfection of VP5 alone. In general, cytoplasmic fluorescence appeared brighter than nuclear fluorescence. This might be due to a reduced efficiency of FIAsh-EDT<sub>2</sub> staining in the nucleus as compared to the cytoplasm, or to the FIAsh-EDT<sub>2</sub> recognition site being less accessible in VP5 when complexed with preVP22a.

All the cells examined contained considerable levels of background fluorescence (Figure 6.5). Cells with bright punctate fluorescence were often observed (examples are shown in Figure 8.1). These would prove a particular problem for studies of labelled capsids. Although the removal of this was attempted by increasing the concentration of BAL during the washing stages of the procedure (upon advice from Guido Gaietta) (Section 6.3), no effective decrease was observed in the levels of background fluorescence. In fact, the

**Wild-type VP5 +  
preVP22a**



**UL19in1235-Flash  
+ preVP22a**

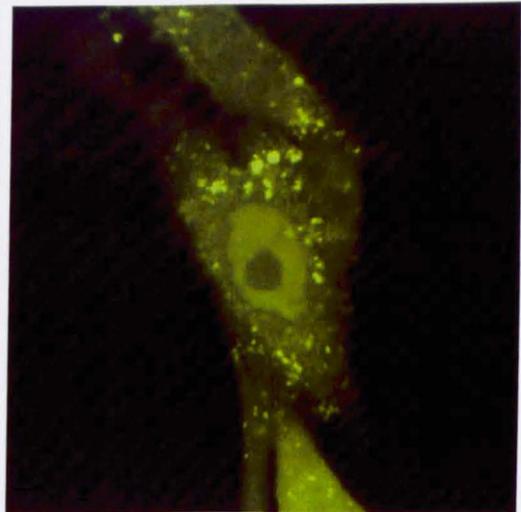


Figure 8.1: Levels of background fluorescence often seen in cells labelled with FIAsh-EDT<sub>2</sub>. BHK-21 cells were cotransfected with UL19in1235-Flash and pJK2 (expressing VP22a) and either pE19 (expressing wild-type VP5) or UL19in1235-Flash. FIAsh-EDT<sub>2</sub> labelling was carried out. Little nuclear fluorescence can be seen in cells transfected with wild-type VP5 (thus containing no CCPGCC motif), and nuclear fluorescence can be observed in cells transfected with UL19in1235-Flash, however background fluorescence can be observed in the cytoplasm of both cells.

levels of background fluorescence are such that the method of FIAsh-EDT<sub>2</sub> labelling is currently inapplicable to studies of the HSV-1 capsid without further optimisation. Several published studies have also reported problems with background fluorescence in FIAsh-EDT<sub>2</sub> labelling. Stroffekova *et al.* (2001) suggested that FIAsh-EDT<sub>2</sub> might bind to endogenous cysteine-containing proteins, making the method not easily applicable to proteins that contain many endogenous cysteine residues. They further suggested that with the levels of background fluorescence observed, FIAsh-EDT<sub>2</sub> labelling might only be applicable for labelling recombinant proteins that are expressed at a very high level. They also reported a variation between cell types in levels of background fluorescence, with higher levels of background fluorescence observed in HEK 293 cells than in HeLa-S3 cells. In this study, BHK-21 cells were the only cell type in which labelling was attempted, therefore no conclusions can be drawn on the relative levels of background fluorescence compared to other cell types.

Recently, attempts have been made to increase the affinity of the tetracysteine motif to FIAsh-EDT<sub>2</sub> (Martin *et al.*, 2005). If this increase could be achieved, more stringent dithiol washes could be used to increase the contrast between the fluorescence obtained from specific and nonspecific binding. These attempts have centred on the amino acids surrounding the tetracysteine motif, and have resulted in the identification of two sequences, HRWCCPGCCKTF and FLNCCPGCCMEP, that showed increased contrast between fluorescence obtained from specific and non-specific binding when used to replace previously characterised tetracysteine motifs (a twenty fold increase in contrast over AEAAARECCRECCARA (Griffin *et al.*, 1998), and a six fold increase over AEAAARECCPGCCARA (Adams *et al.*, 2002)). The molecular explanation of the higher affinity possessed by HRWCCPGCCKTF and FLNCCPGCCMEP for FIAsh-EDT<sub>2</sub> is currently unclear, but based on this recently-published information, it would be interesting to construct recombinant VP5 proteins containing these sequences and to examine whether they reduce the problems of background fluorescence that have up to now been experienced in the application of FIAsh-EDT<sub>2</sub> labelling to the HSV-1 capsid.

## **8.2. VP5 upper domain site-directed mutagenesis**

A panel of nine VP5 mutants was constructed, each of which had an amino acid sequence that differed from wild-type VP5 by one amino acid (Section 7.4). These changes were at positions of apparent close contact between VP5 and VP26, or where bioinformatic analysis suggested a possible role in the binding of VP26 or, in the case of one mutant, of

tegument (Sections 7.2 and 7.5). Eight of the mutants were able to complement virus that did not contain VP5 (Section 7.6.1, Figure 7.4). The failure to complement of the VP5 mutant with an amino acid substitution at position 836 appears to be due to a defect in its ability to package DNA (Section 7.6.3, Figure 7.6). Although it has been demonstrated that this mutant is able to interact with VP19C and VP22a (Figure 7.5), its ability to interact with other HSV-1 capsid proteins such as VP23 has not been examined. The phenotype of this mutant resembles that of some of VP19C insertional mutants described in Section 3.4, in that although the overall conformation of the protein is not disrupted to the extent that it cannot bind its capsid partners, it is sufficiently affected to prevent it from functioning normally. The inability to package DNA suggests either that the assembly of the capsid shell was disrupted due to an alteration in the interactions with other capsid proteins, or that an interaction involving one of the other virion components, such as the scaffold or packaging proteins, had been affected. Due to lack of time, these possibilities were not investigated.

Eight of the mutants contained a substitution at an amino acid that was proposed as having a role in VP26 binding. However, results from co-transfection of each of the mutants with preVP22a and VP26 demonstrated that the mutation had not affected the ability of any of the mutants to bind VP26 (Section 7.6.4, Figure 7.7). Several VP5 amino acids have been suggested as having a role in VP26 binding (Matthew Baker, personal communication). The suggestions were based on computer-based analyses of the crystal structure of the upper domain of VP5; these highlighted charged and hydrophobic residues that appear to be in close contact with VP26. It may be the case that some, or indeed all, of the residues suggested have a role in VP26 binding, but that binding is not abolished by the change made to any one of them. Regarding this point, it should be noted that fluorescent co-localisation is not a quantitative technique, and there is a degree of subjectivity in the assessment of how effectively the two proteins interact. It should also be noted that the VP5 mutants that were constructed and analysed represent a subset of the amino acids that were suggested as having a possible role in VP26 binding. Due to time constraints it was not possible to construct the remaining site-directed mutants, and it is still possible that one of the other amino acids suggested has a central role in VP26 binding. It would be interesting to extend this study to VP5 mutants that contain several amino acid substitutions, i.e. substitutions comprising a combination of the residues suggested by computer-based analysis. Interestingly, Stephen Webster reported that two mutants of VP5 did show reduced interaction with VP26. These mutants contained four amino acid insertions at positions close to many of the mutations described in this section – at amino

## **CHAPTER 9**

## **CONCLUSIONS**

## 9. Conclusions

The work for this thesis built upon earlier work on the HSV-1 capsid structure and was an attempt to analyse the functions and interactions of certain capsid proteins by mutational analysis. The potential identification of amino acids in VP5 that might be important for binding VP26 was made possible by the resolution of the crystal structure of the VP5 upper domain (Bowman *et al.*, 2003). Single amino acid substitutions were previously used to identify residues in VP5 that are important for interaction with preVP22a (Walters *et al.*, 2003) and residues in VP26 that are important for interaction with VP5 (Desai *et al.*, 2003), and using information from the crystal structure of the VP5 upper domain, it was hoped that a similar method of analysis would identify amino acids in VP5 that are important for interaction with VP26.

One-step mutagenesis using the Quickchange procedure was unsuccessful (Frazer Rixon, personal communication). Therefore, mutagenesis was carried out using an overlapping PCR strategy. Unfortunately, this method was prone to unwanted further mutations in the DNA sequence and as the desired mutation involved changes to 1-3 nucleotides, the only way to determine the fidelity of the DNA product was to obtain its sequence. As a result, obtaining site directed mutants was time consuming and site directed mutagenesis was only successful for nine of the eighteen amino acids originally suggested as having a roles in either VP26 or tegument binding. None of these mutations prevented VP26 binding. It would be interesting to see if mutation of any of the remaining amino acids has an affect on the binding of VP26.

The interaction between VP5 and VP26 may involve several amino acids from the two proteins in contact with each other, and if that is the case then it would not be surprising that mutation of just one amino acid did not abolish the interaction. It would be interesting to analyse how the mutation of several candidate amino acids simultaneously would affect VP26 binding, although with eighteen amino acids suggested as candidates the number of different possible combinations would make a complete study impractical. However, the candidate amino acids can be divided into subgroups with respect to regions of VP26 that they might interact with and obtaining VP5 mutants with all of the candidate amino acids in one subgroup mutated simultaneously might reveal more about the nature of the interaction with VP26.

A panel of mutants with five amino acid insertions into the VP5 upper domain is currently under construction (Marion McElwee, personal communication). It will be interesting to see whether any of these mutants have an affect on VP26 binding, and if so, where they map in the VP5 upper domain in relation to the amino acids suggested for mutation in the work described in this thesis.

The main body of results in this thesis describe the mutational analysis of the single copy triplex protein, VP19C. The Mutation Generation System produced a panel of 50 unique insertional mutants of the UL38 ORF. Using this system it was possible to show that VP19C can be divided into three regions with respect to tolerating the insertion of extra amino acids, and highlight regions of the protein that are important for binding VP23 and VP5. The method allowed the rapid production of insertional mutants, and by carrying out the procedure three times it was possible to make alterations to the protocol described in the manufacturers' instructions, which resulted in more efficient generation of mutants. Whether these alterations would be applicable to using the system to study other proteins would depend largely on gene size and the number of insertional mutants desired. A problem encountered using the Mutation Generation System is that insertion of the transposon appears to occur more frequently in areas of the ORF with low GC content. This resulted in some regions of the protein having greater representation among the panel of insertional mutants than others – the region from amino acids 114-181 being particularly poorly represented. If the Mutation Generation System is to be used to produce more VP19C mutants in the future, focusing insertional mutagenesis on regions such as that might prove useful.

The analysis of insertional mutants proved a useful means of making initial assessments in the analysis of a protein that little was known about structurally. Although the system was able to give indications of regions of the protein that are important for its function, it was not possible to identify particular amino acids that might be important for interactions with VP23 and VP5. However, using the information obtained from insertional mutagenesis it might be possible to select regions of VP19C as targets for site-directed and deletional mutagenesis and gain further understanding into these interactions. Identification of amino acids that are important for these interactions will be aided by developing techniques such as site-specific crosslinking, although resolution of the crystal structures of the proteins would provide the most useful tool in understanding these interactions.

The N-terminal region of VP19C turned out to be a particularly interesting region of the protein. It was poorly conserved among *Alphaherpesvirinae* with little evidence of sequence conservation in the N-terminal 110 amino acids, and secondary structure prediction programs were unable to identify much (if any) secondary structure in the N-terminal 101 amino acids. The N-terminal 107 amino acids were particularly tolerant of insertional mutagenesis, and it had previously been reported that the N-terminal 90 amino acids were not absolutely required for capsid assembly (Spencer *et al.*, 1998). Therefore, further analysis of this region was needed to gain a greater understanding of its function.

The analysis of the sub-cellular localisation of VP19C mutants with the first 45 and the first 63 amino acids deleted was an effective means of identifying that this region contained a nuclear localisation signal. Subsequently, fusing amino acid sequences from the VP19C N-terminal region to green fluorescent protein allowed the region containing the nuclear localisation signal to be narrowed down to amino acids 24-55. The nuclear localisation signal in VP19C is unusual, and it would be interesting to characterise it further by identifying the amino acids within this region that are essential for nuclear localisation. This could be achieved by making further GFP-fusion constructs with mutations in individual amino acids. The eight arginines that lie between amino acids 24 and 55 would be the most obvious starting point for such site-directed mutagenesis. It would also be interesting to ascertain whether VP19C's nuclear localisation involves interaction with either importin  $\alpha$  or importin  $\beta$ .

Analysis of N-terminal deletional mutants suggested that amino acids 1-45 are required for nuclear localisation of VP19C, and have an influence on overall capsid shape, while amino acids 46-63 have an as yet unidentified role in capsid formation. Further analysis of this region might be expected to provide a better understanding of the role of VP19C in directing capsid assembly and conformation.

## References

- Abbotts, A. P., Preston, V. G., Hughes, M., Patel, A. H. and Stow, N. D. (2000).** Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. *J. Gen. Virol.* **81**, 2999-3009.
- Adam, S. A. (1999).** Transport pathways of macromolecules between the nucleus and the cytoplasm. *Curr. Opin. Cell Biol.* **11**, 402-406.
- Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J. and Tsien, R. Y. (2002).** New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J. Am. Chem. Soc.* **124**, 6063-6076.
- Adamson, W. E., McNab, D., Preston, V. G. and Rixon, F. J. (2006).** Mutational Analysis of the Herpes Simplex Virus Triplex Protein VP19C. *J. Virol.* **80**, 1537-1548.
- Addison, C., Rixon, F. J. and Preston, V. G. (1990).** Herpes simplex virus type-1 UL28 gene product is important for the formation of mature capsids. *J. Gen. Virol.* **71**, 2377-2384.
- Agirrezabala, X., Martín-Benito, J., Valle, M., González, J. M., Valencia, A., Valpuesta, J. M. and Carrascosa, J. L. (2005).** Structure of the connector of bacteriophage T7 at 8Å resolution: structural homologies of a basic component of a DNA translocating machinery. *J. Mol. Biol.* **347**, 895-902.
- Albrecht, J. C., Nicholas, J., Biller, D., Cameron, K. R., Biesinger, B., Newman, C., Wittmann, S., Craxton, M. A., Coleman, H., Fleckenstein, B. and Honess, R. W. (1992).** Primary structure of the herpesvirus saimiri genome. *J. Virol.* **66**, 5047-5058.

**Alwine, J. C., Steinhart, W. L. and Hill, C. W. (1974).** Transcription of herpes simplex type 1 DNA in nuclei isolated from infected Hep-2 and KB cells. *Virology*. **60**, 302-307.

**Arthur, J. L., Everett, R., Brierly, I. and Efstathiou, S. (1998).** Disruption of the 5' and 3' splice sites flanking the major latency associated transcripts of the herpes simplex virus type 1: evidence for alternative splicing in lytic and latent infections. *J. Gen. Virol.* **79**, 107-116.

**Arzul, I., Renault, T., Lipart, C. and Davison, A. J. (2001).** Evidence for interspecies transmission of oyster herpesvirus in marine bivalves. *J. Gen. Virol.* **82**, 865-870.

**Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, 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.

**Baines, J. D. and Roizman, B. (1991).** The open reading frames UL3, UL4, UL10 and UL16 are dispensible for the replication of herpes simplex virus 1 in cell-culture. *J. Virol.* **65**, 938-944.

**Baines, J. D., Ward, P. L., Campadelli-Fiume, G. and Roizman, B. (1991).** The UL20 gene of herpes simplex virus 1 encodes a function necessary for viral egress. *J. Virol.* **65**, 6414-6424.

**Baker, T. S., Newcomb, W. W., Booy, F. P., Brown, J. C. and Steven, A. C. (1990).** Three-dimensional structures of maturable and abortive capsids of equine herpesvirus 1 from cryoelectron microscopy. *J. Virol.* **64**, 563-573.

**Baker, M. L., Jiang, W., Bowman, B. R., Zhou, Z. H., Quioco, F. A., Rixon, F. J. and Chui, W. (2003).** Architecture of the herpes simplex virus major capsid protein derived from structural bioinformatics. *J. Mol. Biol.* **331**, 447-456.

**Baker M. L., Jiang W, Rixon F. J. and Chiu W. (2005).** Common ancestry of herpesviruses and tailed DNA bacteriophages. *J. Virol.* **79**, 14967-14970.

**Baldick, C. J., Marchini, A., Patterson., C. E. and Shenk, T. (1997).** Human cytomegalovirus tegument protein pp71 (ppUL82) enhances the infectivity of viral DNA and accelerates the infectious cycle. *J. Virol.* **71**, 4400-4408.

**Banfield B. W., Leduc, Y., Esford, L., Schubert, K. and Tufaro, F. (1995).** Sequential isolation of proteoglycan synthesis mutants by using herpes simplex virus as a selective agent: evidence for a proteoglycan-independent virus entry pathway. *J Virol.* **69**, 3290–3298.

**Batterson, W., Furlong, D. and Roizman, B. (1983).** Molecular genetics of herpes simplex virus. VIII. Further characterization 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.

**Bazinet, C. W. and King, J. (1985).** A late gene product of phage P22 affecting virus infectivity. *Virology*, **143**, 368-79.

**Beard, P. M., Taus, N. S. and Baines, J. D. (2002).** DNA cleavage and packaging proteins encoded by genes U<sub>L</sub>28, U<sub>L</sub>15, and U<sub>L</sub>33 of herpes simplex virus type 1 form a complex in infected cells. *J. Virol.* **76**, 4785-4791.

**Beard, P. M. and Baines, J. D. (2004).** The DNA cleavage and packaging protein encoded by the U<sub>L</sub>33 gene of herpes simplex virus 1 associates with capsids. *Virology.* **324**, 475-82.

**Bender, F. C., Whitbeck, J. C., Ponce de Leon, M., Lou, H., Eisenberg, R. J. and Cohen, G. H. (2003).** Specific association of glycoprotein B with lipid rafts during herpes simplex virus entry. *J. Virol.* **77**, 9542-9552.

**Benson, S. D., Bamford, J. K. H., Bamford, D. H. and Burnett, R. M. (1999).** Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell* **98**, 825-833.

**Benson, S. D., Bamford, J. K. H., Bamford, D. H. and Burnett, R. M. (2002).** The X-ray crystal structure of P3, the major coat protein of the lipid-containing bacteriophage PRD1, at 1.65 Å resolution. *Acta Cryst. Sect. D* **58**, 39-59.

**Berg, C. M. and Berg, D. E. (1995).** Transposable elements as tools for molecular analyses in bacteria. *In* *Molecular genetic elements; Frontiers in molecular biology.* Edited by Sherrat, D. J. Oxford University Press, Oxford, UK, 38-68.

**Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000).** The Protein Data Bank. *Nucleic Acids Research*, **28**, 235-242.

**Bernard, J. & Mercier, A. (1993).** Sequence of two *Eco* RI fragments from Salmonis herpesvirus 2 and comparison with Ictalurid herpesvirus 1. *Arch. Virol.* **132**, 437-442.

**Bhella, D., Rixon, F. J. and Dargan, D. J. (2000).** Cryomicroscopy of human cytomegalovirus virions reveals more densely packaged genomic DNA than in herpes simplex virus type 1. *J. Mol. Biol.* **295**, 155-161.

**Bishop, D. H. L. (1992).** Baculovirus expression vectors. *Semin. Virol.* **3**, 253-264.

**Bloom, D. C., Hill, J. M., Devi-Rao, G., Wagner, E. K., Feldman, L. T. and Stevens, J. G. (1996).** A 348 base pair region in the latency associated transcript facilitates herpes simplex virus type 1 reactivation. *J. Virol.* **70**, 2449-2459.

**Boehmer, P. E. and Lehman, I. R. (1997).** Herpes simplex virus DNA replication. *Annu. Rev. Biochem.* **66**, 347-384.

**Booy, F. P., Newcomb, W. W., Trus, B. L., Brown, J. C., Baker, T. S. and Steven, A. C. (1991).** Liquid crystalline, phage-like packaging of encapsidated DNA in herpes simplex virus. *Cell* **64**, 1007-1015.

**Booy, F. P., Trus, B. L., Newcomb, W. W., Brown, J. C., Conway, J. F. and Steven, A. C. (1994).** Finding a needle in a haystack:detection of a small protein (the 12-kDa VP26) in a large complex (the 200-mDa capsid of herpes simplex virus). *Proc. Natl. Acad. Sci. USA.* **91**, 5652-5656.

**Booy, F. P., Trus, B. L., Davison, A. J. and Steven, A. C. (1996).** The capsid architecture of channel catfish virus, an evolutionarily distant herpesvirus, is largely conserved in the absence of discernible sequence homology with herpes simplex virus. *Virology* **215**, 134-141.

**Bowman, B. R., Baker, M. L., Rixon, F. J., Chui, W. and Quioco, F. A. (2003).** Structure of the herpesvirus major capsid protein. *EMBO.* **22**, 757-765.

**Brack, A. R., Dijkstra, J., Granzow, H., Klupp, B. G. and Mettenleiter, T. C. (1999).** Inhibition of virion maturation by simultaneous deletion of glycoproteins E, I and M of pseudorabies virus. *J. Virol.* **73**, 5364-5372.

**Brack, A. R., Klupp, B. G., Granzow, H., Tirabassi, R., Enquist, L. W. and Mettenleiter, T. C. (2000).** Role of the cytoplasmic tails of pseudorabies virus glycoprotein E in virion formation. *J. Virol.* **74**, 4004-4016.

**Bryant, H. E., Wadd, S. E., Lamond, A. I., Silverstein, S. J. and Clements J. B. (2001).** Herpes simplex virus IE63 (ICP27) protein interacts with spliceosome-associated protein 145 and inhibits splicing prior to the first catalytic step. *J Virol.* **75**, 4376-85.

**Buckmaster, A. E., Scott, S. D., Sanderson, M. J., Bournnell, M. E., Ross, N. L. and Binns, M. M. (1988).** Gene sequence and mapping data from Marek's disease virus and herpesvirus of turkeys: implications for herpesvirus classification. *J. Gen. Virol.* **69**, 2033-2042.

**Butcher, S. J., Aitken, J., Mitchell, J., Gowen, B. and Dargan, D. J. (1998).** Structure of the human cytomegalovirus B capsid by electron cryomicroscopy and image reconstruction. *J. Struct. Biol.* **124**, 70-76.

**Campadelli-Fiume, G., Cocchi, F., Menotti, L. and Lopez, M. (2000).** The novel receptors that mediate entry of herpes simplex viruses and animal herpesviruses into cells. *Rev. Med. Virol.* **10**, 305-319.

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

**Casjens, S., and King, J. (1975).** Virus Assembly. *Annual Review of Biochemistry* **44**, 555-771.

**Casjens, S. (1997).** Principles of virion structure, function, and assembly. *In* Structural Biology of Viruses. Edited by Chiu, W., Burnett, R. M. and Garcea, R. Oxford University Press, New York, N.Y., 3-37.

**Caspar, D. and Klug, A. (1962).** Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1-24.

**Catalano, C. E. (2000).** The terminase enzyme from bacteriophage lambda: a DNA-packaging machine. *Cell. Mol. Life Sci.* **57**, 128-148.

**Cebrian, J., Berthelot, N. and Laithier, M. (1989).** Genome structure of cottontail rabbit herpesvirus. *J. Virol.* **63**, 523-531.

**Cerritelli, M. E., Cheng, N., Rosenburg, A. H., McPherson, C. E., Booy, F. P. and Steven, A. C. (1997).** Encapsidated conformation of bacteriophage T7 DNA. *Cell* **91**, 271-280.

**Chang, Y. E., Van Sant, C., Krug, P. W., Sears, A. E. and Roizman, B. (1997).** The null mutant of the UL31 gene of herpes simplex virus 1: construction and phenotype in infected cells. *J. Virol.* **67**, 8307-8315.

**Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A., Kouzarides, T., Martignetti, J. A., Preddie, E., Satchwell, S. C., Tomlinson, P., Weston, K. M. and Barrell, B. G. (1990).** Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.* **154**, 125-169.

**Chen, D. H., Jiang, H., Lee, M., Liu, F. and Zhou, Z. H. (1999).** Three-dimensional visualization of tegument/capsid interactions in the intact human cytomegalovirus. *Virology* **260**, 10-16.

**Chen, D. H., Jakana, J., McNab, D., Mitchell, J., Zhou, Z. H., Dougherty, M., Chiu, W. and Rixon, F. J. (2001).** The pattern of tegument-capsid interaction in the herpes simplex virus type virion is not influenced by the small hexon-associated protein VP26. *J. Virol.* **75**, 11863-11867.

**Chen Hua, S., Lee, L. Y., Garber, D. A., Schaffer, P. A., Knipe, D. M. and Coen, D. M. (2002).** Neither LAT nor open reading frame P mutations increase expression of spliced or intron containing ICP0 transcripts in mouse ganglia latently infected with herpes simplex virus. *J. Virol.* **76**, 4764-4772.

**Chi, J. H. I. and Wilson, D. W. (2000).** ATP-dependant localization of the herpes simplex virus capsid protein VP26 to sites of procapsid maturation. *J. Virol.* **74**, 1468-1476.

- Chi, J. H., Harley, C. A., Mukhopadhyay, A. & Wilson, D. W. (2005).** The cytoplasmic tail of herpes simplex virus envelope glycoprotein D binds to the tegument protein VP22 and to capsids. *J. Gen. Virol.* **86**, 253–261.
- Chook, Y. M. and Blobel, G. (2001).** Karyopherins and nuclear import. *Curr. Opin. Struct. Biol.* **11**, 703-715.
- Church, G. A. and Wilson, D. W. (1997).** Study of herpes simplex virus maturation during a synchronous wave of assembly. *J. Virol.* **71**, 3603-3612.
- Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton. (2004).** The Jalview Java alignment editor. *Bioinformatics* **20**, 426-427.
- Cohen, G. H., Ponce de Leon, M., Diggelmann, H., Lawrence, W. C., Vernon, S. K. and Eisenberg, R. J. (1980).** Structural analysis of the capsid polypeptides of herpes simplex virus types 1 and 2. *J. Virol.* **34**, 67-87.
- Costa, R. H., Cohen, G., Eisenberg, R., Long, D. and Wagner, E. (1984).** Direct demonstration that the abundant 6-kilobase herpes simplex virus type 1 mRNA mapping between 0.23 and 0.27 map units encodes the major capsid protein VP5. *J. Virol.* **49**, 287-292.
- Costanzo F., Campadelli-Fiume G., Foa-Tomasi L. and Cassai E. (1977).** Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase B. *J Virol.* **21**, 996-1001.
- Crick, F. and Watson, J. (1956).** The structure of small viruses. *Nature* **177**, 473-475.
- Dasgupta, A., and Wilson, D. W. (1999).** ATP depletion blocks herpes simplex virus DNA packaging and capsid maturation. *J. Virol.* **73**, 2006-2015.

**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. (1992).** Channel catfish virus: a new type of herpesvirus. *Virology* **186**, 9-14.

**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, A. J. and Davison, M. D. (1995).** Identification of structural proteins of channel catfish virus by mass spectroscopy. *Virology* **206**, 1035-1043.

**Davison, A. J. and McGeoch, D. J. (1995).** Herpesviridae. *In* *Molecular Basis of Viral Evolution*. Edited by Garcia-Arenal, F. Cambridge University Press, Cambridge, UK, 290-309.

**Davison, A. J. (1998).** The genome of salmonid herpesvirus 1. *J. Virol.* **72**, 1974-1982.

**Davison, A. J. (2002).** Evolution of the herpesviruses. *Vet. Microbiol.* **86**, 69-88.

**Davison, A. J., Eberle, R., Hayward, G. S., McGeoch, D. J., Minson, A. C., Pellett, P. E., Roizman, B., Studdert, M. J. and Thiry, E. (2005a).** Herpesviridae. *In* *Virus Taxonomy*. Edited by Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. and Ball, L. A. Elsevier Academic Press, London, UK., 8<sup>th</sup> ed, 193-212.

**Davison, A. J., Trus, B. L., Cheng, N., Steven, A. C., Watson, M. S., Cunningham, C., Le Deuff, R. M. and Renault, T. (2005b).** A novel class of herpesvirus with bivalve hosts. *J. Gen. Virol.* **86**, 41-53.

**de Bruyn Kops A, Uprichard SL, Chen M, Knipe DM. (1998).** Comparison of the intranuclear distributions of herpes simplex virus proteins involved in various viral functions. *Virology* **252**, 162-78.

**del Rio, T., Ch'ng, T. H., Flood, E. A., Gross, S. P. and Enquist, L. W. (2005).** Heterogeneity of a Fluorescent Tegument Component in Single Pseudorabies Virus Virions and Enveloped Axonal Assemblies. *J. Virol.* **79**, 3903-3919.

**Desai, P., Deluca, N. A., Glorioso, J. C. and Person, S. (1993).** Mutations in herpes-simplex virus type-1 genes encoding VP5 and VP23 abrogate capsid formation and cleavage of replicated DNA. *J. Virol.* **67**, 1357-1364.

**Desai, P. and Person, S. (1996).** Molecular interactions between the HSV-1 capsid proteins as measured by the yeast two-hybrid system. *Virology* **220**, 516-521.

**Desai, P. and Person, S. (1998).** Incorporation of the green fluorescent protein into herpes simplex virus type 1 capsid. *J. Virol.* **72**, 7563-7568.

**Desai, P., Deluca, N. A. and Person, S. (1998).** Herpes simplex virus type 1 VP26 is not essential for replication in cell culture but influences production of infectious virus in the nervous system of infected mice. *Virology* **247**, 115-124.

**Desai, P. and Person, S. (1999).** Second site mutations in the N-terminus of the major capsid protein (VP5) overcome a block at the maturational cleavage site of the capsid scaffold proteins of herpes simplex virus type1. *Virology* **261**, 357-366.

**Desai, P. (2000).** A null mutation in the UL36 gene of herpes simplex virus type 1 results in an accumulation of unenveloped DNA-filled capsids in the cytoplasm of infected cells. *J. Virol.* **74**, 11608-11618.

**Desai, P., Sexton, G., McCaffery, J. and Person, S. (2001).** A null mutation in the gene encoding the UL37 polypeptide of herpes simplex virus type 1 abrogates virus maturation. *J. Virol.* **75**, 10259-10271.

**Desai, P., Akpa, J. C. and Person, S. (2003).** Residues of VP26 of Herpes Simplex Virus Type 1 That Are Required for Its Interaction with Capsids. *J. Virol.* **77**, 391-404.

**Diefenbach, B. J., Miranda-Saksena, M., Diefenbach, E., Holland, D. J., Boadle, R. A., Armati, P. J. and Cunningham, A. L. (2002).** Herpes simplex virus tegument protein US11 interacts with conventional kinesin heavy chain. *J. Virol.* **76**, 3282-3291.

**Dohner, K., Wolfstein, A., Prank, U., Echeverri, C., Dujardin, D., Vallee, R. and Sodeik, B. (2002).** Function of dynein and dynactin in herpes simplex capsid transport. *Mol. Biol. Cell.* **13**, 2795-2809.

**Dokland, T. (2000).** Freedom and restraint: themes in virus capsid assembly. *Structure* **8**, R157-R162.

**Dolan, A., Jamieson, F. E., Cunningham, C., Barnett, B. C. and McGeoch, D. J. (1998).** The genome sequence of herpes simplex virus type 2. *J. Virol.* **72**, 2010-2021.

**Douglas, M. W., Diefenbach, R. J., Homa F. L., Miranda-Saksena, M., Rixon, F. J., Vittone, V., Byth, K. and Cunningham, A. L. (2004).** Herpes simplex virus type 1 capsid protein VP26 interacts with dynein light chains RP3 and Tctex1 and plays a role in retrograde cellular transport. *J. Biol. Chem.* **279**, 28522-28530.

**Duda RL, Hendrix RW, Huang WM, Conway JF. (2006).** Shared architecture of bacteriophage SPO1 and herpesvirus capsids. *Curr. Biol.* **16**, R11-13.

**Elgadi, M. M., Hayes, C. E. and Smiley, J. R. (1999).** The herpes simplex virus Vhs protein induces endoribonucleolytic cleavage of target RNAs in cell extracts. *J. Virol.* **73**, 7153-7164.

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

**Elliott, G. and O'Hare, P. (1998).** Herpes simplex virus type 1 tegument protein VP22 induces the stabilization and hyperacetylation of microtubules. *J. Virol.* **72**, 6448-6455.

**Elliott, G. and O'Hare, P. (1999).** Live-cell analysis of a green fluorescent protein-tagged herpes simplex virus infection. *J. Virol.* **73**, 4110-4119.

**Enquist L. W., Husak P. J., Banfield B. W., Smith G. A. (1999).** Infection and spread of alphaherpesviruses in the nervous system. *Adv. Virus Res.* **51**, 237-347.

**Evans, A. S., and Niederman, J. C. (1991).** Epstein-barr virus. *In* Viral infections of humans, 3<sup>rd</sup> ed. Edited by Evans, A. Plenum publishing corporation, New York, N.Y., 256-292.

**Everett, R. D. and Murray, J. (1995).** ND10 Components Relocate to Sites Associated with Herpes Simplex Virus Type 1 Nucleoprotein Complexes during Virus Infection. *J. Virol.*, **79**, 5078-5089.

**Everly, D. N. Jr., Feng, P., Mian, I. S., Read, G. S. (2002).** mRNA Degradation by the Virion Host Shutoff (Vhs) Protein of Herpes Simplex Virus: Genetic and Biochemical Evidence that Vhs Is a Nuclease. *J. Virol.* **76**, 8560-8571.

**Farnsworth, A., Goldsmith, K. and Johnson, D. (2003).** Herpes simplex virus glycoproteins gD and gE/gI serve essential but redundant functions during acquisition of the virion envelope in the cytoplasm. *J Virol.* **77**, 8481-8494.

**Farrell, M. J., Dobson, A. T. and Feldman, L. T. (1991).** Herpes simplex virus latency associated transcript is a stable intron. *PNAS USA.* **88** 790-794.

**Fenwick M. L. and Everett R. D. (1990a).** 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 ML, Everett RD. (1990b).** Inactivation of the shutoff gene (UL41) of herpes simplex virus types 1 and 2. *J Gen Virol.* **71**, 2961–2967.

**Feng, P., Everly, D. N. Jr. and Read, G. S. (2001).** mRNA Decay during Herpesvirus Infections: Interaction between a Putative Viral Nuclease and a Cellular Translation Factor. *J. Virol.* **75**, 10272-10280.

**Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B. and Rossman, M. G. (2004).** Molecular architecture of the prolate head of bacteriophage T4. *Proc. Natl. Acad. Sci. USA.* **101**, 6003-6008.

**Foster, T. P. and Kousoulas, G. (1999).** Genetic analysis of the role of herpes simplex virus type 1 glycoprotein K in infectious virus production and egress. *J. Virol.* **73**, 8457-8468.

**Fuchs, W., Klupp, B. G., Granzow, H., Osterrieder, N. and Mettenleiter, T. C. (2002a).** The interacting UL31 and UL34 gene products of pseudorabies virus are involved in egress from the host-cell nucleus and represent components of primary enveloped but not of mature virions. *J. Virol.* **76**, 364-378.

**Fuchs, W., Klupp, B. G., Granzow, H., Hengartner, C., Brack, A., Mundt, A., Enquist, L. W. and Mettenleiter, T. C. (2002b).** Physical interaction between envelope glycoproteins E and M of pseudorabies virus and the major tegument protein UL49. *J. Virol.* **76**, 8208-8217.

**Fuchs, W., Klupp, B. G., Granzow, H. and Mettenleiter, T. C. (2004).** Essential Function of the Pseudorabies Virus UL36 Gene Product Is Independent of Its Interaction with the UL37 Protein. *J. Virol.* **78**, 11879-11889.

**Fujisawa, H. and M. Morita. (1997).** Phage DNA packaging. *Genes to Cells* **2**, 537-545.

**Furlong, D. (1978).** Direct evidence for 6-fold symmetry of the herpes virus hexon capsomeres. *Proc. Natl. Acad. Sci. USA.* **75**, 2764-2766.

**Gaietta, G., Deerinck, T. J., Adams, S. R., Bouwer, J., Tour, O., Laird, D. W., Sosinsky, G. E., Tsien, R. Y. and Ellisman, M. H. (2002).** Multicolor and Electron Microscopic Imaging of Connexin Trafficking. *Science* **296**, 503-507.

**Gao, M., Matsuick-Kumar, L., Hurlburt, W., Feuer Ditus, S., Newcomb, W. W., Brown, J. C., McCann, P. J., Deckman, I. and Colonno, R. J. (1994).** The protease of herpes simplex virus type 1 is essential for functional capsid formation and viral growth. *J. Virol.* **68**, 3702-3712.

**Geraghty, R. J., Krummenacher, C., Cohen, G. H., Eisenberg, R. J. and Spear, P. G. (1998).** Entry of alphaherpesviruses mediated by poliovirus receptor-related protein 1 and poliovirus receptor. *Science* **280**, 1618-1620.

**Gibson, W. and Roizman, B. (1972).** Proteins specified by herpes simplex virus. VIII. Characterisation of and composition of multiple capsid forms of subtypes 1 and 2. *J. Virol.* **10**, 1044-1052.

**Gibson, W., Baxter, M. K. and Clopper, K. S. (1996).** Cytomegalovirus "missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46. *J. Virol.* **70**, 7454-7461.

**Gold, E. and Nankervis, G. A. (1991).** Cytomegalovirus. *In* *Viral infections of humans*. Edited by Evans, A. Plenum publishing corporation, New York, NY., 3rd ed., 169-189.

**Goldstein, J. N. and Weller, S. K. (1998).** In vitro processing of HSV-1 replication intermediates by the viral alkaline nuclease, UL12. *J. Virol.* **72**, 8772-8781.

**Gompels, U. A., Nicholas, J., Lawrence, G., Jones, M., Thomson, B. J., Martin, M. E., Efstathiou, S., Craxton, M., and Macaulay, H. A. (1995).** The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. *Virology* **209**, 29-51

**Gorlich, D. and Kutay, U. (1999).** Transport between the cell nucleus and the cytoplasm. *Annu. Rev. Cell Dev. Biol.* **15**, 607–660.

**Goshima, F., Watanabe, D., Takakuwa, H., Wada, K., Daikoku, T., Yamada, M. and Nishiyama, Y. (2000).** Herpes simplex virus UL17 protein is associated with B capsids and colocalizes with ICP35 and VP5 in infected cells. *Arch. Virol.* **145**, 417-426.

**Griffin, B. A., Adams, S. R. and Tsien, R. Y. (1998).** Specific covalent labeling of recombinant protein molecules inside live cells. *Science* **281**, 269-272.

**Griffin, B. A., Adams, S. R., Jones, J. and Tsien, R.Y. (2000).** Fluorescent labeling of recombinant proteins in living cells with FLaSH 2001. *Methods in Enzymology.* **327**, 565-578.

**Gross, S. T., Harley, C. A. and Wilson, D. W. (2003).** The cytoplasmic tail of Herpes simplex virus glycoprotein H binds to the tegument protein VP16 in vitro and in vivo. *Virology* **317**, 1–12.

**Grünewald, K., Desai, P., Winkler, D. C., Heymann, J. B., Belnap, D. M., Baumeister, W. and Steven, A. C. (2003).** Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science* **302**, 1396-1398.

**Haarr, L. and Skulstad, S. (1994).** The herpes simplex virus type 1 particle: structure and molecular functions. Review article. *APMIS* **102**, 321-46.

**Haapa, S., Taira, S., Heikkinen, E. and Savilahti, H. (1999).** An efficient and accurate integration of mini-Mu transposons in vitro: a general methodology for functional genetic analysis and molecular biology applications. *Nucleic Acids Res.* **27**, 2777-2784.

**Hardy, W. R. and Sandri-Goldin, R. M. (1994).** Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. *J. Virol.* **68**, 7790-7799.

**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.

**He, J., Schmid, V. F., Zhou, Z. H., Rixon, F. and Chiu, W. (2001).** Finding and using local symmetry in identifying lower domain movements in hexon subunits of the herpes simplex virus type 1 B capsid. *J. Mol. Biol.* **309**, 903-914.

**Heilman, C., Zweig, M., Stephenson, J. and Hampar, B. (1979).** Isolation of a nucleocapsid polypeptide of herpes simplex virus types 1 and 2 possessing immunologically type-specific and cross-reactive determinants. *J. Virol.* **29**, 34-42.

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

**Hendrix, R. W. and Garcea, R. L. (1994).** Capsid assembly of dsDNA viruses. *Semin. Virol.* **5**, 15-26.

**Heymann, J. B., Cheng, N., Newcomb, W. W., Trus, B. L., Brown, J. C. and Steven, A. C. (2003).** Dynamics of herpes simplex virus capsid assembly maturation visualized by time-lapse cryo-electron microscopy. *Nat. Struct. Biol.* **10**, 334-341.

**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.

**Hodge, P. D. and Stow, N. D. (2001).** Effects of mutations within the herpes simplex virus type 1 DNA encapsidation signal on packaging efficiency. *J. Virol.* **75**, 8977-8986.

**Hoffmann, C., Gaietta, G., Bünemann, M., Adams, S. R., Oberdorff-Maass, S., Behr, B., Vilardaga, J. P., Tsien, R. Y., Ellisman, M. H. and Lohse, M. J. (2005).** A FIAsh-based FRET approach to determine G protein-coupled receptor activation in living cells. *Nature Methods* **2**, 171 – 176.

**Homa, F. L. and Brown, J. C. (1997).** Capsid assembly and DNA packaging in herpes simplex virus. *Rev. Med. Virol.* **7**, 107-122.

**Honess, R. W. and Roizman, B. (1974).** Regulation of herpesvirus macro-molecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. *J. Virol.* **14**, 8-19.

**Hong, Z., Beaudet-Miller, M., Durkin, J., Zhang, R. and Kwong, A. D. (1996).** Identification of a minimal hydrophobic domain in the herpes simplex virus type 1 scaffolding protein which is required for interaction with the major capsid protein. *J. Virol.* **70**, 533-540.

**Huszar, D. and Bacchetti, S. (1981).** Partial purification and characterization of the ribonucleotide reductase induced by herpes simplex virus infection of mammalian cells. *J Virol.* **37**, 580–588.

**Ishov, A. M. and Maul, G. G. (1996).** The periphery of nuclear domain 10 (ND10) as site of DNA virus deposition. *J. Cell. Biol.* **134**, 815-826.

**Jakel, S. and Gorlich, D. (1998).** Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO* **17**, 4491–4502.

**Jiang W., Li Z, Zhang Z, Booth CR, Baker ML, Chiu W. (2001).** Semi-automated icosahedral particle reconstruction at sub-nanometer resolution. *J. Struct. Biol.* **136**, 214–225.

**Jiang, W., Li, Z. L., Zhang, Z. X., Baker, M. L., Prevelige, P. E. and Chiu, W. (2003).** Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat. Struct. Biol.* **10**, 131-135.

**Jiang, W., Chang, J., Jakana, J., Weigele, P., King, J. and Chiu, W. (2006).** Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* **439**, 612-616.

**Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E. Jr. and Chiu, W. (2003).** Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat. Struct. Biol.* **10**, 131-135.

**Johnson, D. C. and Ligas, M. W. (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, J. E. and Speir, J. A. (1997).** Quasi-equivalent viruses: A paradigm for protein assemblies. *J. Mol. Biol.* **269**, 665-675.

**Jones, C. (2003).** Herpes Simplex Virus Type 1 and Bovine Herpesvirus 1 Latency. *Clin. Microbiol. Rev.* **16**: 79-95

**Jones, F. E., Smibert, C. A. and Smiley, J. R. (1995).** Mutational analysis of the herpes simplex virus virion host shutoff protein: evidence that vhs functions in the absence of other viral proteins. *J. Virol.* **69**, 4863-4871.

**Kaelin, K., Dezelee, S., Masse, M. J., Bras, F., and Flamand, A. (2000).** The UL25 protein of pseudorabies virus associates with capsids and localizes to the nucleus and to microtubules. *J. Virol.* **74**, 474-482.

**Kaiser, K., Sentry, J. W. and Finnegan, D. J. (1995).** *In* Mobile Genetic Elements. Edited by Sherratt, D. J. Oxford University Press, Oxford, UK, pp. 69-100.

**Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith. (1984).** Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* **311**, 33-38.

**Kennard, J., Rixon, F. J., McDougall, I. M., Tatman, J. D. and Preston, V. G. (1995).** The 25 amino acid residues at the carboxy terminus of the herpes simplex virus type 1 UL26.5 protein are required for the formation of the capsid shell around the scaffold. *J. Gen. Virol.* **76**, 4611-1621.

**King, J. and Chui, W. (1997).** The procapsid-to-capsid transition in double-stranded DNA bacteriophages. *In* Structural Biology of Viruses. Edited by Chiu, W., Burnett, R. M. and Garcea, R. Oxford University Press, New York, NY pp288-311.

**Kirkitadze, M. D., Barlow, P. N., Price, N. C., Kelly, S. M., Boutell, C. J., Rixon, F. J. and McClelland, D. A. (1998).** The Herpes Simplex Virus Triplex Protein, VP23, Exists as a Molten Globule. *J. Virol.* **72**, 10066-10072.

**Klupp, B. G., Granzow, H., Mettenleiter, T. C. (2000).** Primary envelopment of pseudorabies virus at the nuclear membrane requires the UL34 gene product. *J. Virol.* **74**, 10063-10073.

**Klupp, B. G., Granzow, H., Mettenleiter, T. C. (2001).** Effect of the pseudorabies virus US3 protein on nuclear membrane localization of the UL34 protein and virus egress from the nucleus. *J. Gen. Virol.* **82**, 2353-2371.

**Knipe, D. M., Batterson, W., Nosal, C., Roizman, B. and Buchan, A. (1981).** Molecular genetics of herpes simplex virus. VI. Characterization of a temperature-sensitive mutant defective in the expression of all early viral gene products. *J. Virol.* **38**, 539-547.

**Koch, H. G., Delius, H., Matz, B., Flugel, R. M., Clarke, J., and Darai, G. (1985).** Molecular cloning and physical mapping of the tupaia herpesvirus genome. *J. Virol.* **55**, 86-95.

**Kopp, M., Klupp, B. G., Granzow, H., Fuchs, W. and Mettenleiter, T. C. (2002).** Identification and characterization of the pseudorabies virus tegument proteins UL46 and UL47: role for UL47 in virion morphogenesis in the cytoplasm. *J. Virol.* **76**, 8820-8833.

**Kopp, M., Granzow, H., Fuchs, W., Klupp, B. G., Mundt, E., Karger, A. and Mettenleiter, T. C. (2003).** The pseudorabies virus UL11 protein is a virion component involved in secondary envelopment in the cytoplasm. *J. Virol.* **77**, 5339-5351.

**Koslowski, K. M., Shaver, P. R., Casey, J. T., Wilson, T., Yamanaka, G., Sheaffer, A. K., Tenney, D. J. and Pederson, N. E. (1999).** Physical and functional interactions between the herpes simplex virus UL15 and UL28 DNA cleavage and packaging proteins. *J. Virol.* **73**, 1704-1707.

**Kramer, M. F. and Coen, D. M. (1995).** Quantification of transcripts from the ICP4 and thymidine kinase genes in mouse ganglia latently infected with herpes simplex virus. *J. Virol.* **69**, 1389-1399.

**Kratz, P. A., Böttcher, B. and Nassal, M. (1999).** Native display of complete foreign protein domains on the surface of hepatitis B virus capsids. *Proc. Natl. Acad. Sci. USA.* **96**, 1915-1920.

**Krosky, P. M., Underwood, M. R., Turk, S. R., Feng, K. W.-H., Jain, R. K., Ptak, R. G., Westerman, A. C., Biron, K. K., Townsend, L. B. and Drach, J. C. (1998).** Resistance of human cytomegalovirus to benzimidazole ribonucleosides maps to two open reading frames: UL89 and UL56. *J. Virol.* **72**, 4721–4728.

**Kumar S, and Hedges SB. (1998).** A molecular timescale for vertebrate evolution. *Nature* **392**, 917-20.

**Kwong, A. D. and Frenkel, N. (1987).** Herpes simplex virus-infected cells contain a function(s) that destabilizes both host and viral mRNAs. *Proc. Natl. Acad. Sci. USA.* **84**, 1926-1930.

**Kwong, A. D., Kruper, J. A. and Frenkel, N. (1988).** Herpes simplex virus virion host shutoff function. *J. Virol.* **62**, 912-921.

**Latchman D. S. (1988).** Effect of herpes simplex virus type 2 infection on mitochondrial gene expression. *J. Gen. Virol.* **69**, 1405–1410.

**Lehman, I. R. and Boehmer, P. E. (1999).** Replication of herpes simplex virus DNA. *J. Biol. Chem.* **274**, 28059-28062.

**Leib, D., Bogard, C., Kosz-Vnenchak, M., Hicks, K., Coen, D., Knipe, D. and Schaffer, P. (1989).** A deletion mutant of the latency associated transcript of herpes simplex virus type 1 reactivates from the latent infection. *J. Virol.* **63**, 2893-2900.

**Leopardi, R., Ward, P. L., Ogle, W. O. and Roizman, B. (1997).** Association of herpes simplex virus regulatory protein ICP22 with transcriptional complexes containing EAP, ICP4, RNA polymerase II, and viral DNA requires posttranslational modification by the U<sub>L</sub>13 protein kinase. *J. Virol.* **71**, 1133-1139.

**Leslie, J., Rixon, F. J. and McLauchlan. J. (1996).** Overexpression of the herpes simplex virus type 1 tegument protein VP22 increases its incorporation into virus particles. *Virology* **220**, 60-68.

**Leuzinger, H., Ziegler, U., Schraner, E. M., Fraefel, C., Glauser, D. L., Heid, I., Ackermann, M., Mueller, M., Wild, P. (2005).** Herpes simplex virus 1 envelopment follows two diverse pathways. *J. Virol.* **79**, 13047-13059.

**Levy JA. (1997).** Three new human herpesviruses (HHV6,7, and 8). *Lancet* **349**, 558-563.

**Liu, F. and Roizman, B. (1991a).** The promoter transcriptional unit, and coding sequence of herpes simplex virus 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 contain within its coding domain the gene encoding the more abundant substrate. *J. Virol.* **65**, 5149-5156.

**Liu, F. and Roizman, B. (1993).** Characterization of the protease and other products of amino-terminus- proximal cleavage of the herpes simplex virus 1 UL26 protein. *J. Virol.* **67**, 1300-1309.

**Long, M. C., Leong, V., Schaffer, P. A., Spencer, C. A. and Rice, S. A. (1999).** ICP22 and the UL13 protein kinase are both required for herpes simplex virus-induced modification of the large subunit of RNA polymerase II. *J. Virol.* **73**, 5593-5604.

**Luby-Phelps, K. (2000).** Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int. Rev. Cytol.* **192**, 189-221.

**Lund, K. and Ziola, B. (1985).** Cell sonicates used in the analysis of how measles and herpes simplex type 1 virus infections influence Vero cell mitochondrial calcium uptake. *Can. J. Biochem. Cell Bio.* **63**, 1194-1197.

**Lund K. and Ziola B. (1986).** Synthesis of mitochondrial macromolecules in herpes simplex type 1 virus infected Vero cells. *Biochem Cell Biol.* **64**, 1303-9.

**MacLean, C. A., Dolan, A., Jamieson, F. E. and McGeoch, D. J. (1992).** The myristylated virion proteins of herpes simplex virus type 1: investigation of their role in the virus life cycle. *J. Gen. Virol.* **73**, 539-547.

**Macpherson, I. and Stoker, M. (1962).** Polyoma transformation of hamster cell clones: an investigation of genetic factors affecting cell competence. *Virology* **16**, 147-151.

**Mador, N., Goldenberg, D., Cohen, O., Panet, A. and Steiner, I. (1998).** Herpes simplex virus type 1 latency associated transcripts suppress viral replication and reduce immediate early gene mRNA levels in a neuronal cell line. *J. Virol.* **72**, 5067-5075.

**Mar Albà, M., Lee, D., Pearl, F. M. G., Shepherd, A. J., Martin, N., Orengo, C.A. and Kellam, P.** VIDA: a virus database system for the organisation of virus genome open reading frames. *Nucleic Acids Research* **29**, 133-136.

**Marinus, M. G. (1973).** Location of DNA methylation genes on the Escherichia coli K-12 genetic map. *Mol. Gen. Genet.* **127**, 47-55.

**Martin, B. R., Giepmans, B. N. G., Adams, S. R. and Tsien, R. Y. (2005).** Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nature Biotechnology* **23**, 1308-1314.

**Martinez, R., Sarisky, R. T., Weber, P. C. and Weller, S. K. (1996a).** Herpes simplex virus type 1 alkaline nuclease is required for efficient processing of viral DNA replication intermediates. *J. Virol.* **70**, 2075-2085.

**Martinez, R., Shao, L., Bronstein, J. C., Weber, P. C. and Weller, S. K. (1996b).** The product of a 1.9-kb mRNA which overlaps the HSV-1 alkaline nuclease (UL12) cannot relieve the growth defects of a null mutant. *Virology* **215**, 152-164.

**Martinez, R., Goldstein, J. N. and Weller, S. K. (2002).** The product of the UL12.5 gene of herpes simplex virus type 1 is not essential for lytic viral growth and is not specifically associated with capsids. *Virology* **298**, 248-257.

**Matusick-Kumar, L., Hurlburt, W., Weinheimer, S. P., Newcomb, W. W., Brown, J. C. and Gao, M. (1994).** Phenotype of the herpes simplex virus type 1 protease substrate ICP35 mutant virus. *J. Virol.* **68**, 5384–5394.

**Matusick-Kumar, L., Newcomb, W. W., Brown, J. C., McCann, P. J., Hurlburt, W., Weinheimer, S. P. and Gao, M. (1995).** The C-terminal 25 amino acids of the protease and its substrate ICP35 of herpes simplex virus type 1 are involved in formation of sealed capsids. *J. Virol.* **69**, 4347-4356.

**Maul, G. G., Ishov, A. M. and Everett, R. D. (1996).** Nuclear domain 10 as preexisting potential replication start sites of herpes simplex virus type-1. *Virology* **217**, 67-75.

**McClelland, D. A., Aitken, J. D., Bhella, D., McNab, D., Mitchell, J., Kelly, S. M., Price, N. C. and Rixon, F. J. (2002).** pH reduction as a trigger for dissociation of herpes simplex virus type 1 scaffolds. *J. Virol.* **76**, 7407-7417.

**McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E. and Taylor, P. (1988).** The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* **69**, 1531-1574.

**McGeoch, D. J., Gatherer, D. and Dolan, A. (2005).** On phylogenetic relationships among major lineages of the Gammaherpesvirinae. *J Gen Virol* **86**, 307-316.

**McGregor, F., Phelan, A., Dunlop, J. and Clements, J. B. (1996).** Regulation of herpes simplex virus poly(A) site usage and the action of immediate-early protein IE63 in the early–late switch. *J. Virol.* **70**, 1931–1940.

**McLauchlan, J. and Rixon, F. J. (1992).** Characterization of enveloped tegument structures (L particles) produced by alphaherpesviruses: integrity of the tegument does not depend on the presence of capsid or envelope. *J. Gen. Virol.* **73**, 269-276.

**McLauchlan, J. (1997).** The abundance of the herpes simplex virus type 1 UL37 tegument protein in virus particles is closely controlled. *J. Gen. Virol.* **78**, 189-194.

**McClintock, B. (1949).** Mutable loci in maize. *Carnegie Institute of Washington Yearbook.*, **48**, 142-154.

**McNab, A. R., Desai, P., Person, S., Roof, L. L. Thomsen, D. R., Newcomb, W. W., Brown, J. C. and Homa, F. L. (1998).** The product of the herpes simplex virus type 1 UL25 gene is required for encapsidation but not for cleavage of replicated viral DNA. *J. Virol.* **72**, 1060-1070.

**Mettenleiter, T. C. (2002).** Herpesvirus assembly and egress. *J. Virol.* **76**, 1537-1547.

**Mettenleiter, T. C. (2004).** Budding events in herpesvirus morphogenesis. *Virus Res.* **106**, 167-180.

**Mettenleiter, T. C. and Minson, T. (2006).** Egress of Alphaherpesviruses. *J. Virol.* **80**, 1610-1612.

**Meyer, H. H., Ripalti, A., Landini, M. P., Radsak, K., Kern, H. F. and Hensel, G. M. (1997).** Human cytomegalovirus late-phase maturation is blocked by stably expressed UL32 antisense mRNA in astrocytoma cells. *J. Gen. Virol.* **78**, 2621-2631.

**Miyawaki, A., Sawano, A. and Kogure, T. (2003).** Lighting up cells: labelling proteins with fluorophores. *Nature Cell Biology* **5**, S1-S7.

**Mizuuchi, K. (1983).** *In vitro* transposition of bacteriophage Mu: A biochemical approach to a novel replication reaction. *Cell* **35**: 785-794.

**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.

**Moore, P. S., Gao, S. J., Dominguez, G., Cesarman, E., Lungu, O., Knowles, D. M., Garber, R., Pellett, P. E., McGeoch, D. J. and Chang, Y. (1996).** Primary characterization of a herpesvirus agent associated with Kaposi's sarcoma. *J. Virol.* **70**, 549-558.

**Morais, M. C., Choi, K. H., Koti, J. S., Chipman, P. R., Anderson, D. L. and Rossman, M. G. (2005).** Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of  $\Phi 29$ . *Mol. Cell* **18**, 149-159.

**Mossman, K., Sherburne, R., Lavery, C., Duncan, J. and Smiley, J. (2000).** Evidence that herpes simplex virus VP16 is required for viral egress downstream of the initial envelopment event. *J. Virol.* **74**, 6287-6299.

**Muranyi, W., Haas, J., Wagner, M., Krohne, G. and Koszinowski, U. H. (2002).** Cytomegalovirus Recruitment of Cellular Kinases to Dissolve the Nuclear Lamina. *Science* **297**, 854-857.

**Murata, T., Goshima, F., Daikoku, T., Inagaki-Ohara, K., Takakuwa, H., Kato, K. and Nishiyama, Y. (2000).** Mitochondrial distribution and function in herpes simplex virus-infected cells. *J. Gen. Virol.* **81**, 401-406.

**Nahmias, A. J., Keyserling, H. and Lee, F. K. (1991).** Herpes simplex viruses 1 and 2. *In* *Viral infections of humans*. Edited by Evans, A. A. Evans (Ed). Plenum publishing corporation, New York, N.Y. 3<sup>rd</sup> ed., 393-413.

**Namba, K., Pattanayek, R. and Stubbs, G. (1989).** Visualization of protein-nucleic acid interactions in a virus: refined structure of intact tobacco mosaic virus at 2.9Å resolution by X-ray fiber diffraction. *J. Mol. Biol.* **171**, 61-93.

**Nandhagopal, N., Simpson, A. A., Gurnon, J. R., Yan, X., Baker, T. S., Graves, M. V., Van Etten, J. L. and Rossmann, M. G. (2002).** The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proc. Natl Acad. Sci. USA.* **99**, 14758-14763.

**Nealon, K., Newcomb, W. W., Pray, T. R., Craik, C. S., Brown, J. C. and Kedes, D. H. (2001).** Lytic replication of Kaposi's sarcoma-associated herpesvirus results in the formation of multiple capsid species: isolation and molecular characterization of A, B, and C capsids from a gammaherpesvirus. *J. Virol.* **75**, 2866-78.

**Newcomb, W. W. and Brown, J. C. (1989).** Use of Ar<sup>+</sup> plasma etching to localize structural proteins in the capsid of herpes simplex virus type 1. *J. Virol.* **63**, 4697-4720.

**Newcomb, W. W. and Brown, J. C. (1991).** Structure of the herpes simplex virus capsid: effects of extraction with guanidine hydrochloride and partial reconstitution of extracted capsids. *J. Virol.* **65**, 613-620.

**Newcomb, W. W., Trus, B. L., Booy, F. P., Steven, A. C., Wall, J. S., and Brown, J. C. (1993).** Structure of the herpes simplex virus capsid: Molecular composition of the pentons and the triplexes. *J. Mol. Biol.* **232**, 499-511.

**Newcomb, W. W., Homa, F. L., Thomsen, D. R., Ye, Z. and Brown, J. C. (1994).** Cell-free assembly of the herpes simplex virus capsid. *J. Virol.* **68**, 6059-6063.

**Newcomb, W. W., Homa, F. L., Thomsen, D. R., Booy, F. P., Trus, B. L., Steven, A. C., Spencer, J. V. and Brown, J. C. (1996).** Assembly of the herpes simplex virus capsid: characterisation of intermediates observed during cell-free capsid formation. *J. Mol. Biol.* **263**, 432-436.

**Newcomb, W. W., Homa, F. L., Thomsen, D. R., Trus, B. L., Steven, A., Booy, F. P. and Brown, J. C. (1999).** Assembly of herpes simplex virus procapsid from purified components and identification of small complexes containing the major capsid and scaffolding proteins. *J. Virol.* **73**, 4239-4250.

**Newcomb, W. W., Trus, B. L., Cheng, N., Steven, A. C., Sheaffer, A. K., Tenney, D. J., Weller, S. K. and Brown, J. C. (2000).** Isolation of herpes simplex virus procapsids from cells infected with a protease-deficient mutant virus. *J. Virol.* **74**, 1663-1673.

**Newcomb, W. W., Juhas, R. M., Thomsen, D. R., Homa, F. L., Burch, A. D, Weller, S. K. and Brown, J. C. (2001a).** The UL6 gene product forms the portal for entry of the DNA into the herpes simplex virus capsid. *J. Virol.* **75**, 10923-10932.

**Newcomb, W. W., Homa, F. L., Thomsen, D. R., and Brown, J. C. (2001b).** *In vitro* assembly of the herpes simplex virus procapsid: formation of small procapsids at reduced scaffolding protein concentration. *J. Struct. Biol.* **133**, 23-31.

**Newcomb, W. W., Homa, F. L. and Brown, J. C. (2005).** Involvement of the portal at an early step in herpes simplex virus capsid assembly. *J. Virol.* **79**, 10540-10546.

**Newcomb, W. W., Homa, F. L. and Brown, J. C. (2006).** Herpes simplex virus capsid structure: DNA packaging protein UL25 is located on the external surface of the capsid near the vertices. *J. Virol.* **80**, 6286-6294.

**Nicholas, J. (1996).** Determination and analysis of the complete nucleotide sequence of human herpesvirus. *J. Virol.* **70**, 5975-5989.

**Nicholson, P., Addison, C., Cross, A. M., Kennard, J., Preston, V. G. and Rixon, F. J. (1994).** Localization of the herpes simplex virus type 1 major capsid protein VP5 to the cell nucleus requires the abundant scaffolding protein VP22a. *J. Gen. Virol.* **75**, 1091-1099.

**Nicola, A. V., McEnvoy, A. M. and Straus, S. E. (2003).** Roles for endocytosis and low pH in herpes simplex virus entry into HeLa and chinese hamster ovary cells. *J. Virol.* **77**, 5324-5332.

**O'Connor, C. M., B. Damania, and D. H. Kedes. (2003).** De novo infection with rhesus monkey rhadinovirus leads to the accumulation of multiple intranuclear capsid species during lytic replication but favors the release of genome-containing virions. *J. Virol.* **77**, 13439-13447.

**Ogasawara, M., Suzutani, T., Yoshida, I. and Azuma, M. (2001).** Role of the UL25 gene product in the herpes simplex virus capsid: location of UL25 product in the capsid and demonstration that it binds DNA. *J. Virol.* **75**, 1427-1436.

**Oien, N. L., Thomsen, D. R., Wathen, M. W., Newcomb, W. W., Brown, J. C. and Homa, F. L. (1997).** Assembly of herpes simplex virus capsids using the human cytomegalovirus scaffold protein: critical role of the C terminus. *J. Virol.* **71**, 1281-1291.

**Ojala, P. M., Sodiek, B., Ebersold, M. W., Kutay, U. and Helenius, A. (2000).** Herpes simplex virus type 1 entry into host cells: reconstitution of capsid binding and uncoating at the nuclear pore complex in vitro. *Mol. Cell. Biol.* **20**, 4922-4931.

**Oroskar, A. A. and Read, G. S. (1987).** A mutant of herpes simplex virus type 1 exhibits increased stability of immediate early (alpha) mRNAs. *J. Virol.* **61**, 604-606.

**Pak, A. S., Everly, D. N., Knight, K. and Read, G. S. (1995).** The virion host shutoff protein of herpes simplex virus inhibits reporter gene expression in the absence of other viral gene products. *Virology* **211**, 491-506.

**Palmeri, D. and Malim, M. H. (1999).** Importin beta can mediate the nuclear import of an arginine-rich nuclear localization signal in the absence of importin alpha. *Mol. Cell. Biol.* **19**, 1218-1225

**Patel, A. H. and Maclean, J. B. (1995).** The product of the UL6 gene of herpes simplex virus type 1 is associated with virus capsids. *Virology* **206**, 465-478.

**Pelletier, A., Do, J., Brisebois, L., Lagace, L. and Cordingley, M. (1997).** Self-association of herpes simplex virus type 1 ICP35 is via coiled-coil interactions and promotes stable interaction with the major capsid protein. *J. Virol.* **61**, 5197-5208.

**Perng, G. C., Jones, C., Ciacci-Zanella, J., Stone, M., Henderson, G., Yukt, A., Slanina, S. M., Hofman, F. M., Ghiasi, H., Nesburn, A. B. and Wechsler, S. L. (2000).** Virus-induced neuronal apoptosis blocked by the herpes simplex virus latency-associated transcript. *Science* **247**, 1500-1503.

**Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004).** UCSF Chimera – a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612.

**Phelan, A., Carmo-Fonseca, M., McLauchlan, J., Lamond, A. I. and Clements, J. B. (1993).** A herpes simplex virus type 1 immediate-early gene product, IE63, regulates small nuclear ribonucleoprotein distribution. *Proc. Natl. Acad. Sci. USA.* **90**, 9056-9060.

**Ponten, J and Saksela, E. (1967).** Two established in vitro cell lines from human mesenchymal tumours. *Int. J. Cancer* **2**, 434-447.

**Preston, V. G., Coates, J. A. 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 of UL26 reading frame. *Virology* **186**, 87-98.

**Preston, V. G., al Kobaisi, M. F., McDougall, I. M. and Rixon, F. J. (1994).** The herpes simplex virus gene UL26 proteinase in the presence of the UL26.5 gene product promotes the formation of scaffold-like structures. *J. Gen. Virol.* **75**, 2355-2366.

**Purves, F., Spector, D. and Roizman, B. (1992).** UL34, the target of the herpes simplex virus Us3 protein kinase, is a membrane protein which in its unphosphorylated state associates with novel phosphoproteins. *J. Virol.* **66**, 4295-4303.

**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 alpha (immediate-early) polypeptides. *J. Virol.* **46**, 498-512.

**Read, G. S., Karr, B. M. and Knight, K. (1993).** Isolation of a herpes simplex virus type 1 mutant with a deletion in the virion host shutoff gene and identification of multiple forms of the *vhs* (UL41) polypeptide. *J. Virol.* **67**, 7149-7160.

**Reynolds, A. E., Ryckman, B. J., Baines, J. D., Zhou, Y., Liang, L. and Roller, R. J. (2001).** UL31 and UL34 proteins of herpes simplex virus type 1 form a complex that accumulates at the nuclear rim and is required for envelopment of capsids. *J. Virol.* **75**, 8803-8817.

**Reynolds, A. E., Wills, E. G., Roller, R. J., Ryckman, B. J. and Baines, J. D. (2002).** Ultrastructural Localization of the Herpes Simplex Virus Type 1 UL31, UL34, and US3 Proteins Suggests Specific Roles in Primary Envelopment and Egress of Nucleocapsids. *J. Virol.* **76**, 8939-8952.

**Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J. E., Douglas, T. and Young, M. (2004).** The structure of a thermophilic archaeal virus shows that a dsDNA viral capsid type spans all domains of life. *Proc. Natl. Acad. Sci. USA.* **101**, 7716-7720.

**Rice, S. A., Long, M. C., Lam, V., Schaffer, P. A. and Spencer, C. A. (1995).** Herpes simplex virus immediate-early protein ICP22 is required for viral modification of host RNA polymerase II and establishment of the normal viral transcription program. *J. Virol.* **69**, 5550-5559.

**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., Davison, M. D. and Davison, A. J. (1990).** Identification of the genes encoding two capsid proteins of herpes simplex virus type 1 by direct amino acid sequencing. *J. Gen. Virol.* **71**, 1211-1214.

**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 type 1 infected cells. *J. Gen. Virol.* **73**, 277-284.

**Rixon, F. J. (1993).** Structure and assembly of herpesviruses. *Seminars in Virology* **4**, 135-144.

**Rixon, F. J., Addison, C., McGregor, A., Macnab, S. J., Nicholson, P., Preston, V. G. and Tatman, J. D. (1996).** Multiple interactions control the intracellular localization of the herpes simplex virus type 1 capsid proteins. *J. Gen. Virol.* **77**, 2251-2260.

**Rixon, F. J. and McNab, D. (1999).** Packaging-competent capsids of a herpes simplex virus temperature-sensitive mutant have properties similar to those of in vitro-assembled procapsids. *J. Virol.* **73**, 5714-5721.

**Rixon, F. J. and Chiu, W. (2003).** Studying large viruses. *In Advances in protein chemistry*, volume 64: Virus structure. Edited by Richards, R. M., Eisenberg, D. S. and Kuritan, J. Academic press, San Diego, CA., 379-408.

**Robertson, B. J., McCann, P. J., Matusick-Kumar, L., Newcomb, W. W., Brown, J. C., Colunno, R. J. and Gao, M. (1996).** Separate functional domains of the herpes simplex virus type 1 protease: evidence for cleavage inside capsids. *J. Virol.* **70**, 4317-4328.

**Roizman, B. (1979).** The structure and isomerization of herpes simplex virus genomes. *Cell* **16**, 481-494.

**Roizman, B., Desrosier, R. C., Fleckenstein, B., Lopez, C., Minson, A. C. and Studdert, M. J. (1992).** The family herpesviridae: an update. *Arch. Virol.* **123**, 425-449.

**Roizman, B. and Pellett, P. E. (2001).** The family *Herpesviridae*: A brief introduction. *In* *Fields Virology*, fourth edition. Edited by Knipe, D. M., Howley, P. M., Griffin, D. E., Martin, M. A., Lamb, R. A., Roizman, B. and Straus, S. E. Lippincott Williams and Wilkins, Philadelphia, PA., 2381-2397.

**Roizman, B. and Knipe, D. M. (2001).** Herpes simplex viruses and their replication. *In* *Fields Virology*, fourth edition. Edited by Knipe, D. M., Howley, P. M., Griffin, D. E., Martin, M. A., Lamb, R. A., Roizman, B. and Straus, S. E. Lippincott Williams and Wilkins, Philadelphia, PA., 2399-2442.

**Roller, R., Zhou, Y., Schnetzer, R., Ferguson, J. and DeSalvo, D. (2000).** Herpes simplex virus type 1 UL34 gene product is required for viral envelopment. *J. Virol.* **74**, 117-129.

**Rost, B., G. Yachdav, and J. Liu. (2004).** The PredictProtein server. *Nucleic Acids Res.* **32**, W321-W326.

**Rudner, L., Nydegger, S., Coren, L. V., Nagashima, K., Thali, M. and Ott, D. E. (2005).** Dynamic Fluorescent Imaging of Human Immunodeficiency Virus Type 1 Gag in Live Cells by Biarsenical Labeling. *J. Virol.* **79**, 4055-4065

**Ruthel, G., Demmin, G. L., Kallstrom, G., Javid, M. P., Badie, S. S., Will, A. B., Nelle, T., Schokman, R., Nguyen, T. L., Carra, J. H., Bavari, S. and Aman, M. J. (2005).** Association of Ebola Virus Matrix Protein VP40 with Microtubules. *J. Virol.* **79**, 4709-4719.

**Saad, A., Zhou, Z. H., Jakana, J., Chiu, W. and Rixon, F. J. (1999).** Roles of triplex and scaffolding proteins in herpes simplex virus type 1 capsid formation suggested by structures of recombinant particles. *J. Virol.* **73**, 6821-6830.

**Samaniego, L. A., Webb, A. L. and DeLuca, N. A. (1995).** Functional interactions between herpes simplex virus immediate-early proteins during infection: gene expression as a consequence of ICP27 and different domains of ICP4. *J. Virol.* **69**, 5705-5715.

**Sanchez, V., Greis, K. D., Sztul, E. and Britt, W. J. (2000).** Accumulation of virion tegument and envelope proteins in a stable cytoplasmic compartment during human cytomegalovirus replication: Characterization of a potential site of virus assembly. *J. Virol.* **74**, 975-986.

**Sandri-Goldin, R. M., Hibbard, M. K. and Hardwicke, M. A. (1995).** The C-terminal repressor region of herpes simplex virus type 1 ICP27 is required for the redistribution of small nuclear ribonucleoprotein particles and splicing factor SC35; however, these alterations are not sufficient to inhibit host cell splicing. *J. Virol.* **69**, 6063-6076.

**Schaley, J. E., Polonskaia, M. and Hearing, P. (2005).** The Adenovirus E4-6/7 Protein Directs Nuclear Localization of E2F-4 via an Arginine-Rich Motif. *J. Virol.* **2005**, **79**: 2301-2308.

**Schrag, J. D., Venkataram-Prasad, B. V., Rixon, F. J. and Chui, W. (1989).** Three-dimensional structure of the HSV-1 nucleocapsid. *Cell* **56**, 651-660.

**Sessler, R. J., and Noy, N. (2005).** A ligand-activated nuclear localization signal in cellular retinoic acid binding protein-II. *Mol. Cell* **18**, 343-353.

**Shao, L., Rapp, L. M. and Weller, S. K. (1993).** Herpes simplex virus 1 alkaline nuclease is required for efficient egress of capsids from the nucleus. *Virology* **196**, 146-162.

**Sheaffer, A. K., Newcomb, W. W., Brown, J. C., Gao, M., Weller, S. K. and Tenney, D. J. (2000).** Evidence for controlled incorporation of herpes simplex virus type 1 UL26 protease into capsids. *J. Virol.* **74**, 6838-6848.

**Sheaffer, A. K., Newcomb, W. W., Gao, M., Yu, D., Weller, S. K., Brown, J. C. and Tenney, D. J. (2001).** Herpes simplex virus DNA cleavage and packaging proteins associate with the procapsid prior to its maturation. *J. Virol.* **75**, 687-698.

**Shieh, M. T., WuDunn, D., Montgomery, R. I., Esko, J. D. & Spear, P. G. (1992).** Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *J. Cell. Biol.* **116**, 1273-1281.

**Shukla, D., Liu, J., Blaiklock, P., Shworak, N. W., Bai, X., Esko, J. D., Cohen, G. H., Eisenberg, R. J., Rosenberg, R. D. and Spear, P. G. (1999).** A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell.* **99**, 13-22.

**Sia, I. G. and Patel, R. (2000).** New Strategies for Prevention and Therapy of Cytomegalovirus Infection and Disease in Solid-Organ Transplant Recipients. *Clin. Microbiol. Rev.* **13**, 83-121.

**Silva, M. C., Yu, Q. C., Enquist, L. and Shenk, T. (2003).** Human cytomegalovirus UL99-encoded pp28 is required for the cytoplasmic envelopment of tegument-associated capsids. *J Virol.* **77**, 10594-605.

**Skepper, J. N., Whiteley, A., Browne, H. and Minson, A. (2001).** Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment → deenvelopment → reenvelopment pathway. *J. Virol.* **75**, 5697-5702.

**Sodeik, B., Ebersold, M. W. and Helenius, A. (1997).** Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. *J. Cell Biol.* **136**, 1007-1021.

**Sourvinos, G. and Everett, R. D. (2002).** Visualisation of parental HSV-1 genomes and replication compartments in association with ND10 in live infected cells. *EMBO* **21**, 4989-4997.

**Spaete, R. R. and Frenkel, N. (1982).** The herpes simplex virus amplicon: a new eukaryotic defective-virus cloning-amplifying vector. *Cell* **30**, 295-304.

**Spear, P. G. and Roizman, B. (1972).** Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpesvirion. *J. Virol.* **9**, 143-159.

**Spear, P. G. and Longnecker, R. (2003).** Herpesvirus entry: an update. *J. Virol.* **77**, 10179-10185.

**Spencer, J. V., Newcomb, W. W., Thomsen, D. R., Homa, F. L. and Brown, J. C. (1998).** Assembly of the herpes simplex virus capsid: preformed triplexes bind to the nascent capsid. *J. Virol.* **72**, 3944-3951.

**Stannard, L. M., Fuller, A. O. and Spear, P. G. (1987).** Herpes simplex virus glycoproteins associated with different morphological entities projecting from the virion envelope. *J. Gen. Virol.* **68**, 715-725.

**Steven, A. C., Couture, E., Aebi, U. and Showe, M. K. (1976).** Structure of T4 polyheads. II. A pathway of polyhead transformation as a model for T4 capsid maturation. *J Mol Biol.* **106**, 187-221.

- Steven, A. C., Roberts, C. R., Hay, J., Bisher, M. E., Pun, T. and Trus, B. L. (1986).** Hexavalent capsomers of herpes simplex virus type 2: symmetry, shape, dimensions, and oligomeric state. *J. Virol.* **57**, 578-584.
- Steven, A. C. and Spear, P. G. (1997).** Herpesvirus capsid assembly and envelopment. *In* *Structural Biology of Viruses*. Edited by Chiu, W., Burnett, R. M. and Garcea, R. Oxford University Press, New York, NY., p312-351.
- Stevens, J. G., Wagner, E. K., Devi-Rao, G. B., Cool, M. L. and Feldman, L. (1987).** RNA complementary to a herpesvirus alpha gene mRNA is predominant in latently infected neurones. *Science* **235**, 1056-1059.
- Stevenson, A. J., Morrison, E. E., Chaudhari, R., Yang, C. C. and Meredith, D. M. (1997).** Processing and intracellular localization of the herpes simplex type 1 proteinase. *J. Gen. Virol.* **78**, 671-675.
- Stothard, P. (2000).** The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* **28**, 1102-1104.
- Stow, N. D., McMonagle, E. C. and Davison, A. J. (1983).** Fragments from both termini of the herpes simplex virus type 1 genome contain signals required for the encapsidation of viral DNA. *Nucleic Acids Res.* **11**, 8205-8220
- Stow, N. D., Hammarsten, O., Arbuckle, M. I. and Elias, P. (1993).** Inhibition of herpes simplex virus type 1 DNA replication by mutant forms of the origin-binding protein. *Virology* **196**, 413-418.
- Stow, N. D. (2001).** Packaging of genomic and amplicon DNA by the herpes simplex virus type 1 UL25 null mutant, KUL25 NS. *J. Virol.* **75**, 10755-10765.
- Strauss, H. and King, J. (1984).** Steps in the stabilisation of newly packaged DNA during phage P22 morphogenesis. *J. Mol. Biol.* **172**, 523-543.

**Stroffekova, K., Proenza, C. and Beam, K. G. (2001).** The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Eur. J. Physiol.* **442**, 859-866.

**Subak-Sharpe, J. H. and Dargan, D. J. (1998).** HSV molecular biology: general aspects of herpes simplex virus molecular biology. *Virus Genes* **16**, 239-251.

**Szilágyi, J. F. and Cunningham, C. (1991).** Identification and characterisation of a novel non-infectious herpes simplex virus-related particle. *J. Gen. Virol.* **72**, 661-668.

**Szilágyi, J. F. and Berriman, J. (1994).** Herpes simplex virus L particles contain spherical membrane-enclosed inclusion vesicles. *J. Gen. Virol.* **75**, 1749-1753.

**Tatman, J. D., Preston, V. G., Nicholson, P., Elliott, R. M. and Rixon, F. J. (1994).** Assembly of herpes simplex virus type 1 capsids using a panel of recombinant baculoviruses. *J. Gen. Virol.* **75**, 1101-1113.

**Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994).** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

**Thompson, R. L., and Sawtell, N. M. (1997).** The herpes simplex virus type 1 latency-associated transcript gene regulates the establishment of latency. *J. Virol.* **71**, 5432-5440.

**Thomsen, D. R., Roof, L. L. and Homa, F. L. (1994).** Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. *J. Virol.* **68**, 2442-2457.

**Thomsen, D. R., Newcomb, W. W., Brown, J. C. and Homa, F. L. (1995).** Assembly of the herpes simplex virus capsid: Requirement for the carboxyl-terminal twenty-five amino acids of the proteins encoded by the UL26 and UL26.5 genes. *J. Virol.* **69**, 3690-3703.

**Thuman-Commike, P. A., Greene, B., Malinski, J. A., Burbea, M., McGough, A., Chiu, W. and Prevelige, P. E. Jr. (1999).** Mechanism of scaffolding-directed virus assembly suggested by comparison of scaffolding-containing and scaffolding-lacking P22 procapsids. *Biophys. J.* **76**, 3267-3277.

**Thurlow, J. K., Rixon, F. J., Murphy, M., Targett-Adams, P., Hughes, M. and Preston, V. G. (2005).** The herpes simplex virus type 1 DNA packaging protein UL17 is a virion protein that is present in both the capsid and the tegument compartments. *J. Virol.* **79**, 150-158.

**Thurlow, J. K., Murphy, M., Stow, N. D., Preston, V. G. (2006).** Herpes Simplex Virus Type 1 DNA-Packaging Protein UL17 Is Required for Efficient Binding of UL25 to Capsids. *J. Virol.* **80**, 2118-2126.

**Truant, R. and Cullen, B. R. (1999).** The Arginine-Rich Domains Present in Human Immunodeficiency Virus Type 1 Tat and Rev Function as Direct Importin beta - Dependent Nuclear Localization Signals. *Mol. Cell. Biol.* **19**, 1210-1217.

**Trus, B. L., Newcomb, W. W., Booy, F. P., Brown, J. C. and Steven, A. C. (1992).** Distinct monoclonal antibodies separately label the hexons or the pentons of herpes simplex virus capsid. *Proc. Natl. Acad. Sci. USA.* **89**, 11508-11512.

**Trus, B. L., Homa, F. L., Booy, F. P., Newcomb, W. W., Thomsen, D. R., Cheng, N., Brown, J. C. and Steven, A. C. (1995).** Herpes simplex virus capsids assembled in insect cells infected with recombinant baculoviruses: structural authenticity and localization. *J. Virol.* **69**, 7362-7366.

**Trus, B. L., Booy, F. P., Newcomb, W. W., Brown, J. C., Homa, F. L., Thomsen, D. R., and Steven, A. C. (1996).** The herpes virus procapsid: structure, conformational changes upon maturation, and roles of the triplex proteins VP19c and VP23 in assembly. *J. Mol. Biol.* **263**, 447-462.

**Trus, B. L., Gibson, W., Cheng, N. and Steven, A. C. (1999).** Capsid structure of simian cytomegalovirus from cryoelectron microscopy: evidence for tegument attachment sites [published erratum appears in *J Virol* 1999 73: 4530]. *J. Virol.* **73**, 2181-2192.

**Trus, B. L., Heymann, B., Nealon, K., Cheng, N., Newcomb, W. W., Brown, J. C., Kedes, D. H. and Steven, A. C. (2001).** Capsid structure of kaposi's sarcoma-associated herpesvirus, a gammaherpesvirus, compared to those of an alphaherpesvirus, herpes simplex virus type 1, and a betaherpesvirus, cytomegalovirus. *J. Virol.* **75**, 2879-2890.

**Trus, B. L., Cheng, N., Newcomb, W. W., Homa, F. L., Brown, J. C. and Steven, A. C. (2004).** Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J. Virol.* **78**, 12668-12671.

**Tsurumi, T. and Lehman, I. R. (1990).** Release of RNA polymerase from vero cell mitochondria after herpes simplex virus type 1 infection. *J. Virol.* **64**, 450-452.

**Uprichard, S. L. and D. M. Knipe. (1996).** Herpes simplex virus ICP27 mutant viruses exhibit reduced expression of specific DNA replication genes. *J. Virol.* **70**, 1969-1980.

**Valpuesta, J. M. and Carrascosa, J. L. (1994).** Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Q. Rev. Biophys.* **27**, 107-155.

**van Genderen, I. L., Brandimarti, R., Torrisi, M., Campadelli, G. and van Meer, G. (1994).** The phospholipid composition of extracellular herpes simplex virus differs from that of the host cell nuclei. *Virology* **200**, 831-836.

**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.

**Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. and McCawley, P. (1977).** The establishment of two insect lines from the insect *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *In Vitro* **13**, 213-217.

**Vernon, S., Lawrence, W. and Cohen, G. (1974).** Morphological components of herpesvirus 1. Intercapsomeric fibrils and the geometry of the capsid. *Intervirology* **4**, 237-248.

**Vernon, S. K., Ponce de Leon, M., Cohen, G. H., Eisenberg, R. J. and Rubin, B. A. (1981).** Morphological components of herpes virus. III. Localization of herpes simplex virus type 1 nucleocapsid polypeptides by immune electron microscopy. *J. Gen. Virol.* **54**, 39-46.

**Vittone, V., Diefenbach, E., Triffett, D., Douglas, M. W., Cunningham, A. L. and Diefenbach, R. J. (2005).** Determination of interactions between tegument proteins of herpes simplex virus type 1. *J. Virol.* **79**, 9566-9571.

**Vlazny, D. A., Kwong, A. and Frenkel, N. (1982).** Site-Specific Cleavage/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.

**Wagenaar, F., Pol, J. M., Peeters, B., Gielkins, A. L. J., de Wind, N. and Kimman, T. G. (1995).** The US3-encoded protein kinase from pseudorabies virus affects egress of virions from the nucleus. *J. Gen. Virol.* **76**, 1851-1859.

**Walters J. N., Sexton G. L., McCaffery J. M. and Desai, P. (2003).** Mutation of single hydrophobic residue I27, L35, F39, L58, L65, L67, or L71 in the N terminus of VP5 abolishes interaction with the scaffold protein and prevents closure of herpes simplex virus type 1 capsid shells. *J. Virol.* **77**, 4043-59.

**Ward, P. L., Ogle, W. O. and Roizman, B. (1996).** Assemblons: nuclear structures defined by aggregation of immature capsids and some tegument proteins of herpes simplex virus 1. *J. Virol.* **70**, 4623-4631.

**Warner, M. S., Geraghty, R. J., Martinez, W. M., Montgomery, R. I., Whitbeck, J. C., Xu, R., Eisenberg, R. J., Cohen, G. H. and Spear, P. G. (1998).** A cell surface protein with herpesvirus entry activity (HveB) confers susceptibility to infection by mutants of herpes simplex virus *type 1*, *herpes simplex virus type 2*, and pseudorabies virus. *Virology* **246**, 179–189.

**Warner, S. C., Desai, P. and Person, S. (2000).** Second site mutations encoding residues 34 and 78 of the major capsid protein (VP5) of herpes simplex virus type 1 are important for overcoming a blocked maturation cleavage site of the capsid scaffold proteins. *Virology* **278**, 217-226.

**Warner, S. C., Chytrova, G., Desai, P. and Person, S. (2001).** Mutations in the N-terminus of VP5 alter its interaction with the scaffold proteins of herpes simplex virus type 1. *Virology* **284**, 308-316.

**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.

**Webster, S. J. (2004).** Studies of the interactions between the capsid shell proteins of herpes simplex virus type 1. PhD thesis. University of Glasgow.

**Weir, J. P. (2001).** Regulation of herpes simplex virus gene expression. *Gene* **271**, 117-130.

- Weller, S. K., Carmichael, E. P., Aschman, D. P., Goldstein, D. L. and Schaffer, P. A. (1987).** Genetic and phenotypic characterization of mutants in four essential genes that map to the left half of the HSV-1 UL DNA. *Virology* **161**, 198-210.
- White, C. A., Stow, N. D., Patel, A. H., Hughes, M. and Preston, V. G. (2003).** Herpes Simplex Virus Type 1 Portal Protein UL6 Interacts with the Putative Terminase Subunits UL15 and UL28. *J. Virol.* **77**, 6351-6358.
- Whiteley, A., Bruun, B., Minson, T. and Browne, H. (1999).** Effects of Targeting Herpes Simplex Virus Type 1 gD to the Endoplasmic Reticulum and trans-Golgi Network. *J. Virol.* **73**, 9515-9520.
- Wikoff, W. R., Liljas, L., Duda, R. L., Tsuruta, H., Hendrix, R. W. and Johnson, J. E. (2000).** Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129-2133.
- Wildy, P., Russell, W. C. and Horne, R. W. (1960).** The morphology of herpes virus. *Virology* **12**, 204-222.
- Wingfield, P. T., Stahl, S. J., Thomsen, D. R., Homa, F. L., Booy, F. P., Trus, B. L. and Steven, A. C. (1997).** Hexon-only binding of VP26 reflects differences between the hexon and penton conformations of VP5, the major capsid protein of herpes simplex virus. *J. Virol.* **73**, 8955-8961.
- Wray, G. A., Levinton, J. S. and Shapiro, L. H. (1996).** Molecular Evidence for Deep Precambrian Divergences Among Metazoan Phyla. *Science* **274**, 568-73.
- Wu, L., Lo, P., Yu, X., Stoops, J. K., Forghani, B. and Zhou, Z. H. (2000).** Three-dimensional structure of the human herpesvirus 8 capsid. *J. Virol.* **74**, 9646-9654.
- WuDunn, D. and Spear, P. G. (1989).** Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J. Virol.* **63**, 52-58.

- Yang, Z. R., Thomson, R., McMeil, P and Esnouf, R. M. (2005).** RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **21**: 3369-3376.
- Ye, G. J., Vaughan, K. T., Vallee, R. B. and Roizman, B. (2000).** The herpes simplex virus 1 U(L)34 protein interacts with a cytoplasmic dynein intermediate chain and targets nuclear membrane. *J. Virol.* **74**, 1355-1363.
- Yu, D. and Weller, S. K. (1998).** Herpes simplex virus type 1 cleavage and packaging proteins UL15 and UL28 are associated with B but not C capsids during packaging. *J. Virol.* **72**, 7428-7439.
- Yu, X. K., O'Connor, C. M., Atanasov, I., Damania, B., Kedes, D. H. and Zhou, Z. H. (2003).** Three-dimensional structures of the A, B, and C capsids of rhesus monkey rhadinovirus: insights into gammaherpesvirus capsid assembly, maturation, and DNA packaging. *J. Virol.* **77**, 13182-13192.
- Zhang, X., Efstathiou, S and Simmons, A. (1994).** Identification of novel herpes simplex virus replicative intermediates by field inversion gel electrophoresis: implications for viral thymidine kinase expression. *J. Virol.* **67**, 1482-1492.
- Zhou, Z. H., Venkataram-Prasad, B. V., Jakana, J., Rixon, F. J. and Chui, W. (1994).** Protein subunit structures in the herpes simplex virus A-capsid determined from 400kV spot-scan electron cryomicroscopy. *J. Mol. Biol.* **242**, 456-469.
- Zhou, Z. H., He, J., Jakana, J., Tatman, J. D., Rixon, F. J. and Chui, W. (1995).** Assembly of VP26 in herpes simplex virus-1 inferred from structures of wild-type and recombinant capsids. *Nat. Struct. Biol.* **2**, 1026-1030.
- Zhou, Z. H., Chui, W., Haskell, K., Spears, H. Jr., Jakana, J., Rixon, F. J. and Scott, L. R. (1998a).** Refinement of herpesvirus B-capsid structure on parallel supercomputers. *Biophys. J.* **74**, 576-588.

**Zhou, Z. H., Macnab, S. J., Jakana, J., Scott, L. R., Chui, W. and Rixon, F. J. (1998b).** Identification of the sites of interaction between the scaffold and outer shell in herpes simplex virus-1 capsids by difference electron imaging. *Proc. Natl. Acad. Sci. USA.* **95**, 2778-2783.

**Zhou, Z. H., Chen, D. H., Jakana, J., Rixon, F. J. and Chui, W. (1999).** Visualisation of the tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. *J. Virol.* **73**, 3210-3218.

**Zhou, Z. H., Dougherty, M., Jakana, J., He, J., Rixon, F. J. and Chui, W. (2000).** Seeing the herpesvirus capsid at 8.5Å. *Science* **288**, 877-880.

**Zhu, Q. and Courtney, R. J. (1994).** Chemical cross-linking of virion envelope and tegument proteins of herpes simplex virus type 1. *Virology* **204**, 590–599.

## Mutational Analysis of the Herpes Simplex Virus Triplex Protein VP19C

Walt E. Adamson, David McNab, Valerie G. Preston, and Frazer J. Rixon\*

*MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, United Kingdom*

Received 1 September 2005/Accepted 13 November 2005

**Herpes simplex virus type 1 (HSV-1) capsids have an icosahedral structure with capsomers formed by the major capsid protein, VP5, linked in groups of three by distinctive structures called triplexes. Triplexes are heterotrimers formed by two proteins in a 1:2 stoichiometry. The single-copy protein is called VP19C, and the dimeric protein is VP23. We have carried out insertional and deletional mutagenesis on VP19C and have examined the effects of the mutations on virus growth and capsid assembly. Insertional mutagenesis showed that the N-terminal ~100 amino acids of the protein, which correspond to a region that is poorly conserved among herpesviruses, are insensitive to disruption and that insertions into the rest of the protein had various effects on virus growth. Some, but not all, severely disabled mutants were compromised in the ability to bind VP23 or VP5. Analysis of deletion mutants revealed the presence of a nuclear localization signal (NLS) near the N terminus of VP19C, and this was mapped to a 33-amino-acid region by fusion of specific sequences to a green fluorescent protein marker. By replacing the endogenous NLS with that from the simian virus 40 large T antigen, we were able to show that the first 45 amino acids of VP19C were not essential for assembly of functional capsids and infectious virus particles. However, removing the first 63 amino acids resulted in formation of aberrant capsids and prevented virus growth, suggesting that the poorly conserved N-terminal sequences have some as-yet-unidentified function.**

Herpesviruses particles have complex structures consisting of four components: envelope, tegument, capsid, and core (7, 24, 41). The core consists of the double-stranded DNA genome, which is packaged into a preformed capsid within the nuclei of infected cells. The capsid is surrounded by a proteinaceous layer of variable thickness called the tegument, and the entire entity is enclosed by a glycoprotein-containing lipid envelope.

The herpes simplex virus type 1 (HSV-1) capsid has been the subject of intense study, and its structure and composition are relatively well characterized compared to the other virion compartments (31). It consists of a protein shell, which has  $T = 16$  icosahedral symmetry. Three distinctive structural elements, designated pentons, hexons, and triplexes, make up the bulk of this shell; 150 hexons form the faces and edges of the icosahedron, while the vertices are formed by the pentons. The triplexes occupy the local threefold positions between hexons and between hexons and pentons. The icosahedral geometry of the capsid largely determines the number of each component, but the precise numbers of pentons and triplexes are uncertain. This is because 1 of the 12 vertices is believed to be the location of a portal complex through which the genome enters and leaves the capsid (16). It seems probable that the portal replaces 1 of the pentons, leaving 11 occupying the remaining icosahedral vertices. Similarly, for the triplex, there are 320 local threefold positions in the capsid, including 5 surrounding each vertex. However, the relationship between the triplexes and the portal is unknown, making it unclear whether triplexes are present at the five positions surrounding the portal.

Assembly of functional capsids is thought to initiate through an interaction between the internal scaffolding protein and the

portal (17). Stepwise addition of shell and scaffold complexes then results in the formation of an unstable spherical procapsid (14, 15). This undergoes a maturational reconfiguration (8, 37), during which the scaffolding protein is proteolytically cleaved and removed and the DNA is packaged. The reconfiguration results in altered interactions between capsid proteins and generates a stable polyhedral capsid shell (14, 19, 26).

The HSV-1 capsid shell is primarily composed of four types of protein. Five copies of the major capsid protein, VP5 (149 kDa), form the pentons, and six copies form the bulk of the hexons (18, 46), while six copies of VP26 (12 kDa) occupy the outer surfaces of the hexons (39, 44). The remaining two proteins, VP19C (50 kDa) and VP23 (34 kDa), make up the triplexes. The triplex is a characteristic structural feature of herpesvirus capsids and is a heterotrimer with  $\alpha\beta_2$  organization, which in HSV-1 is formed by a single molecule of VP19C ( $\alpha$ -subunit) and two copies of VP23 ( $\beta$ -subunit) (18, 29). The triplex forms connections with two of the three surrounding capsomers through domains that have been designated the head and tail (11), and all three triplex proteins interact with the capsid floor (14, 43). VP19C is required for efficient transport of VP23 to the nucleus, which is the site of capsid assembly (25). It is also able to transport VP5 into the nucleus, although in HSV-1-infected cells this function is predominantly provided by the scaffolding protein, preVP22a (20), a transient capsid component that is required during assembly.

The structure of the HSV-1 capsid has been determined to ~8.5 Å by electron cryomicroscopy (43). Detailed examination of this structure, combined with the analysis of particles formed by subsets of the capsid proteins (29, 39, 44, 45), has allowed the boundaries of the component protein molecules to be defined with good precision. However, with the exception of a 60-kDa domain of VP5 that forms the tops of the hexons and pentons (2), the structures of the capsid proteins are unknown at the atomic level, and relatively little is known about the functional organiza-

\* Corresponding author. Mailing address: MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, United Kingdom. Phone: 44 141 330 4025. Fax: 44 141 337 2236. E-mail: f.rixon@vir.gla.ac.uk.

TABLE 1. PCR primers used in cloning UL38 mutants

Primer	Sequence <sup>a</sup>
UL38-45F	GTCTAGAGGATCCATATGGGCCAGGGCTC CCAGGGGTC
UL38-63F	GTCTAGAGGATCCATATGTTGCTGGCCGTG GACGGCACAGAC
UL38R	GACAAGATTCACGGCGCATGCCGCCACTC GCCGGGG
UL38-45PCR F1	GAAGAAGAAGCGCAAGTTCGACGGCCAGG GCTCCCAAGGGGTCG
UL38-63PCR F1	GAAGAAGAAGCGCAAGTTCGACTTGTCTGG CCTGGACGGCACAGAC
38R-Xba	TCTAGATCAGCGCATGCCGCCACTCGC
UL38NLS-45PCR F3	CGCAATGGCTCCGAAGAAGAAGCGCAAGGT CGACGGCCAGGGGCTCCCAAGGGGTCGG
UL38ATG	GACAGAATTCGCAATGAAGACCAATCCGC
UL38AA46	ACAGTTCGACCCGTCGGCGGAGAGATCCG
UL38AA56	ACAGTTCGACCGGGTCCCGACCCCTGGG
UL38AA66	ACAGTTCGACCCGCAAGCAACACAGCGTGC
UL38AA76	ACAGTTCGACCCCGCCAGGGGGCGGCTCTG
UL38AA83	ACAGTTCGACGATTCGTCGTTGGGGGTCAGC
UL38AA118	ACAGTTCGACAGATCCGTCAGGATCACTTGG
UL38-24 NTERM	GCAATGACACGGATACCGCGGGACAG
UL38-33 NTERM	GCAATGCTTCGGCGCTCTCGGCCCC
GFP UPSTREAM	GCTGATTATGATCTAGAGTCGC

<sup>a</sup> The recognition sequences for BamHI (GGTACC), BglII (AGATCT), Sall (GTCGAC), and XbaI (TCTAGA) used in cloning are singly or doubly underlined.

tion of the protein sequences. In this paper, we describe mutagenesis of VP19C to analyze the roles of different parts of the protein in virus growth and capsid assembly.

#### MATERIALS AND METHODS

**Plasmids.** (i) **VP19C insertional mutants.** pFBpCl was made by inserting a BglII-EcoRI fragment containing the human cytomegalovirus (HCMV) promoter/enhancer from pCl-neo (Promega) into BamHI/EcoRI-digested pFASTBAC HTa (Invitrogen). As a first step, a single NotI site in pFASTBAC HTa was removed by digesting the DNA with NotI, making the ends flush with T4 polymerase, and religation. The UL38 open reading frame (ORF) encoding VP19C was isolated from pBJ382 (33) by digestion with XbaI and HincII and ligated into NheI/StuI-digested pFBpCl to produce pUL38FBpCl, which allows expression of VP19C under the control of the HCMV immediate-early (IE) promoter. Insertional mutants were generated using the transposon-based Mutation Generation System (Finnzymes). Following the in vitro transposition reaction, pUL38FBpCl clones containing the entrance (MI-Kan<sup>r</sup>) sequence inserted at random positions, were isolated by selection on kanamycin (50 µg/ml) agar plates. Individual colonies were grown, and the position of the insertion within the pUL38FBpCl DNA was determined by digestion with PstI, which released a fragment of 2,918 bp if the transposon had inserted into or near the UL38 ORF or a fragment of 1,790 bp if it had inserted elsewhere in the plasmid. Plasmid DNAs containing inserts within the UL38 ORF were digested with NotI and religated. This removed the bulk of the transposon sequence, leaving a 15-bp insertion, which introduced five additional amino acids into the protein. The positions of insertions were determined by DNA sequencing.

(ii) **VP19C N-terminal truncations.** Truncated UL38 ORFs were generated by PCR using the forward primers UL38-45F and UL38-63F and the reverse primer UL38R (Table 1). The PCR products were purified and cloned into pGEM-T Easy (Promega). Individual colonies were grown and screened by restriction enzyme digestion. Clones with the inserts in the correct orientation were designated pUL38-45T Easy and pUL38-63T Easy. For expression in insect cells, the UL38 sequences were isolated from pUL38-45T Easy and pUL38-63T Easy by digestion with XbaI and PstI and cloned downstream from the polyhedrin promoter in XbaI/PstI-digested pFASTBAC1 (Invitrogen) to generate pUL38-45FB1 and pUL38-63FB1. For expression in mammalian cells, the UL38 sequences were isolated from pUL38-45T Easy and pUL38-63T Easy by digestion with EcoRI and SpeI and cloned downstream from the HCMV IE1 promoter in EcoRI/XbaI-digested pFBpCl to generate pUL38-45FBpCl and pUL38-63FBpCl.

(iii) **VP19C-NLS fusions.** PCRs were carried out using the forward primers UL38-45PCR F1 and UL38-63PCR F1 and the reverse primer 38R-Xba (Table 1). UL38-45PCR F1 and UL38-63PCR F1 introduce a Sall site immediately upstream of amino acids 45 and 63 in the UL38 ORF. The PCR products were

purified and cloned into pGEM-T Easy to generate pUL38-45PCR-T Easy and pUL38-63PCR-T Easy, respectively. An oligonucleotide, UL38NLS-45PCR F3, was designed that specifies the nuclear localization signal (NLS) of the simian virus 40 (SV40) large T antigen (sequence, PKKKRKV) (10) fused to the N terminus of the UL38 ORF truncated at amino acid 45 (Table 1). A PCR was performed using primers UL38NLS-45PCR F3 and 38R-Xba, with pUL38-45PCR-T Easy as the template. The PCR fragment was purified and cloned into pGEM-T Easy to generate pUL38-45NLS-T Easy. The UL38 insert was isolated by digestion with EcoRI and XbaI and ligated into EcoRI/XbaI-digested pFBpCl to generate pUL38-45NLSFBpCl. The Sall/XbaI fragment from pUL38-63PCR-T Easy was then purified and ligated into Sall/XbaI digested pUL38-45NLSFBpCl to generate pUL38-63NLSFBpCl.

(iv) **VP19C-green fluorescent protein (GFP) fusion.** DNA fragments containing increasing numbers of residues from the 5' end of the VP19C ORF were produced by PCR using the forward primer UL38ATG and the reverse primers UL38AA46, UL38AA56, UL38AA66, UL38AA76, UL38AA83, and UL38AA118 (Table 1). The appropriate PCR products were purified and cloned into pGEM-T Easy. Individual colonies were grown and screened by digestion with EcoRI. The UL38 inserts were isolated by digestion with EcoRI and Sall and ligated into EcoRI/Sall-digested pEGFP-N1 (Clontech) to generate p66AA-UL38-GFP, p66AA-UL38-GFP, p66AA-UL38-GFP, p66AA-UL38-GFP, p66AA-UL38-GFP, and p18AA-UL38-GFP. To delete sequences from the N terminus, PCR was carried out with p66AA-UL38-GFP DNA as a template and using the forward primers UL38-24 NTERM and UL38-33 NTERM and the reverse primer GFP UPSTREAM. The resulting PCR products were cloned into pGEM-T Easy, and then the inserts were isolated by digestion with EcoRI and NotI and ligated into EcoRI/NotI-digested p66AA-UL38-GFP.

All plasmids were checked by diagnostic restriction enzyme digestion and DNA sequencing.

**Cells.** BHK-21 clone 13 cells were grown in Glasgow minimal essential medium (MEM) supplemented with 10% tryptose phosphate broth and 10% newborn calf serum. UL38RSC<sup>+</sup> cells (36) were grown in Dulbecco's MEM supplemented with 10% fetal calf serum and 500 µg/ml Geneticin G418 (Invitrogen) at every fourth pass. U2OS cells were grown in Dulbecco's MEM supplemented with 10% fetal calf serum. Sf21 (*Spodoptera frugiperda*) cells were cultured in TC100 medium supplemented with 5% fetal calf serum.

**Viruses.** The HSV-1 UL38-minus mutant  $\Delta$ 38YFP was propagated on UL38RSC cells (36).

**Baculoviruses.** AcUL38 (expressing full-length VP19C), AcUL19 (VP5), and AcUL26.5 (preVP22a) have been described previously (23, 33). AcAB3.12 (expressing VP23, VP26, and the UL26 protease) was made by cloning the ORFs into the triple expression transfer plasmid pAcAB3 (BD Biosciences, Pharmingen) and recombining them into AcPAK6 (1). Recombinant baculoviruses expressing other VP19C proteins were produced using the BAC-TO-BAC Baculovirus Expression System (Invitrogen) following the manufacturer's protocol. Briefly, pUL38-45FB1, pUL38-63FB1, pUL38FBpCl, pUL38-45NLSFBpCl, and pUL38-63NLSFBpCl were transfected into competent DH10Bac *Escherichia coli* cells. White colonies were selected on X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside)-containing agar plates, and the recombinant bacmid DNAs were isolated and transfected into Sf21 cells. The progeny baculoviruses were screened for the presence of inserts by PCR. Positive isolates were designated UL38-45FBac, UL38-63FBac, UL38FBpCl, UL38-45NLSFBpCl, and UL38-63NLSFBpCl and were grown to high titer.

**Plasmid transfection and baculovirus transduction.** BHK cells for immunofluorescence were grown on 13 mm glass coverslips and transfected with plasmid DNAs using Lipofectin PLUS reagent (Invitrogen) according to the manufacturer's instructions. After 3 h of incubation at 37°C, the transfection reagent was replaced with fresh culture medium and incubation was continued for a further 16 h before the cells were fixed for immunofluorescence.

BHK cells for virus rescue were grown directly in 24-well tissue culture dishes and transfected with plasmid DNA as described above. At 5 h posttransfection, the cells were infected with  $\Delta$ 38YFP. After 1 h at 37°C, unabsorbed virus was inactivated by incubating the cells in Tris-glycine, pH 3, for 45 s (27). The culture medium was replaced, and the cells were incubated for a further 40 h at 37°C before being harvested into the medium and frozen at -70°C prior to titration on UL38RSC cells.

U2OS cells were overlaid with 50 PFU/cell (titrated on Sf21 cells) of the appropriate baculovirus and incubated at 20°C. After 1 h, the inoculum was removed and the cells were infected with 5 PFU/cell of  $\Delta$ 38YFP. Incubation was continued at 37°C for 16 h, and the cells were harvested and processed for electron microscopy.

**Immunofluorescence.** Transfected BHK cells were fixed for 1 h in methanol at -20°C, washed in phosphate-buffered saline (PBS), and incubated for 10 min in

PBS containing 5% newborn calf serum and 0.05% Tween 20 (solution A). The coverslips were drained, inverted onto 20  $\mu$ l of primary antibody, and incubated at room temperature for 45 min. After three washes in solution A, the coverslips were drained and inverted onto 20  $\mu$ l of secondary antibody for a further 45 min. The coverslips were washed four times in solution A and twice in deionized water before being mounted in Mowiol 4-88 (Harco) plus 2.5% DABCO (Sigma) and examined on a Zeiss LSM 510 confocal microscope.

**Western blot analysis.** Proteins separated by electrophoresis on 10% sodium dodecyl sulfate-polyacrylamide gels were transferred to Hybond E:CL nitrocellulose membranes and analyzed using the enhanced-chemiluminescence method (Amersham). Antibody dilutions and membrane washes were carried out in PBS-T (PBS containing 0.05% [vol/vol] Tween 20 [Sigma]). Before addition of the first antibody, the membrane was blocked in PBS-T containing 5% milk powder. Protein A-peroxidase (Sigma) was diluted 1:1,000 in PBS-T.

**Antibodies.** Mouse monoclonal antibody mAb02040, raised against purified VP19C, was diluted 1:1,000 in solution A (immunofluorescence) or PBS-T (Western blotting). Rabbit antisera rAb184 and rAb186, against purified VP5 and VP23, respectively, were diluted 1:1,000 in solution A. Secondary antibodies, goat anti-mouse (GAM) fluorescein isothiocyanate (FITC) conjugate and goat anti-rabbit (GAR) tetramethyl rhodamine isothiocyanate (TRITC) conjugate (Sigma), were diluted 1:100 in solution A.

**Electron microscopy.** Cells were harvested and pelleted in embedding capsules (TAAB). The cell pellets were fixed with 2.5% glutaraldehyde and 1% osmium tetroxide, dehydrated through a graded alcohol series, and embedded in Epon 812 resin (TAAB). Thin sections were cut using a LEICA ULTRACUT E, stained with uranyl acetate and lead citrate, and examined on a JEOL 100S electron microscope.

## RESULTS

### Generation and analysis of insertional mutants in VP19C.

Fifty clonally distinct mutants containing 5-amino-acid insertions in the UL38 ORF were generated by transposon-based insertional mutagenesis into pUL38FbpCI, which expresses VP19C from the HCMV IE promoter. The insertions were distributed throughout the ORF with an average spacing of ~9 amino acids and a maximum separation of 38. Two pairs of insertions (in242A-in242B and in358A-in358B) were at different sites in the same codon. In these cases, the mutant proteins differed only in the identities of the inserted amino acids (Table 2). One insertion (inSTOP) disrupted the TGA stop codon, resulting in the addition of 5 amino acids at the end of the protein.

To analyze the effects of the mutations on the properties of VP19C, their abilities to complement growth of the HSV-1 UL38 deletion mutant  $\nu$ Δ38YFP (36) were examined. DNA of each insertional mutant was transfected into BHK cells, which were incubated at 37°C for 5 h before being infected with 2 PFU/cell of  $\nu$ Δ38YFP. After 1 h, the unabsorbed input virus was neutralized by an acid wash. Incubation was continued at 37°C for 40 h, at which time the progeny virus was harvested and titrated on UL38RSC rabbit skin cells expressing VP19C. Figure 1 shows the relative yield obtained for each of the insertional mutants compared to that obtained with the wild-type control.

Eighteen of the 50 insertional mutants had little or no effect on the ability of VP19C to support virus growth, producing yields of over 80% of the wild-type VP19C control level. The remaining 32 showed lower levels of complementation, with two producing yields of <10% and 16 being completely negative for virus growth. The results shown in Fig. 1 suggest that VP19C can be divided into three broadly defined regions, with the majority of the severely disabling insertions (<10% of wild-type yield) occurring in the central region of the protein, between in113 and in350. The yields obtained with all 12 of the insertions into the first 107 amino acids were >40% of the

TABLE 2. Locations and sequences of five amino acid insertions into VP19C

AA no. <sup>a</sup>	AA sequence <sup>b</sup>
16	ggCGRSGst
23	ppMRPHPt
24	ptAAAAPTr
26	trCGRTRdt
37	vlLRPHLrp
51	lpSAAAPrg
80	tpCGRTPnd
84	dtDAAAATeq
86	cqVRPQOal
102	alMRPHLig
106	spAAAASPh
107	prAAAAPRhh
113	rqCGRSQvi
143	hlVRPHLah
182	raVRPHLagl
183	agAAAAGlv
193	aaALVAAca
200	daLRPHAr
207	avAAAAVra
217	rgCGRSGtr
234	amCGRTMvh
235	mvAAAMVht
242A	phAAAAPhev
242B	phCGRTHlev
243	heCGRNEvm
248	fgCGRIGgl
286	avAAAAVil
289	lpVRPHIpac
310	flLRPHLyl
327	cvCGRSVyv
328	vyAAAAYvl
329	vyMRPHVik
330	viNAAAIs
338	rgCGRSGile
343	alVRPHLer
350	lrCGRIRit
358A	hgCGRNGte
358B	hgAAAHIgte
365	ppVRPHIPap
368	pnCGRTNrn
390	agCGRTGqv
396	pqCGRTQfa
426	lgAAALGmm
436	clLRPHLhr
438	hrMRPHIRte
439	rtDAAAATer
449	pvAAAAPVvi
459	rpGAAAAPge
463	wrVRPQRac
STOP	caC'AAAA

<sup>a</sup> AA, amino acid. The number denotes the last unchanged amino acid before the insertion.

<sup>b</sup> The five inserted amino acids are shown in uppercase, and the unchanged flanking residues are in lowercase.

wild-type level, with only in38 in the middle of this region and in106 and in107 at its C terminus giving <80%. Similarly, all but 2 (in390 and in463) of the 14 insertions into the C-terminal 107 amino acids gave complementation levels of >30%. In contrast, 16 of the 18 insertions that gave complementation values of <10% mapped in the central region of the protein between amino acids 113 and 350, and of the 24 insertions in this region, only 4 gave values of >30%.

**Interaction of capsid proteins.** In order to ensure the formation of functional capsids and hence infectious virions, VP19C

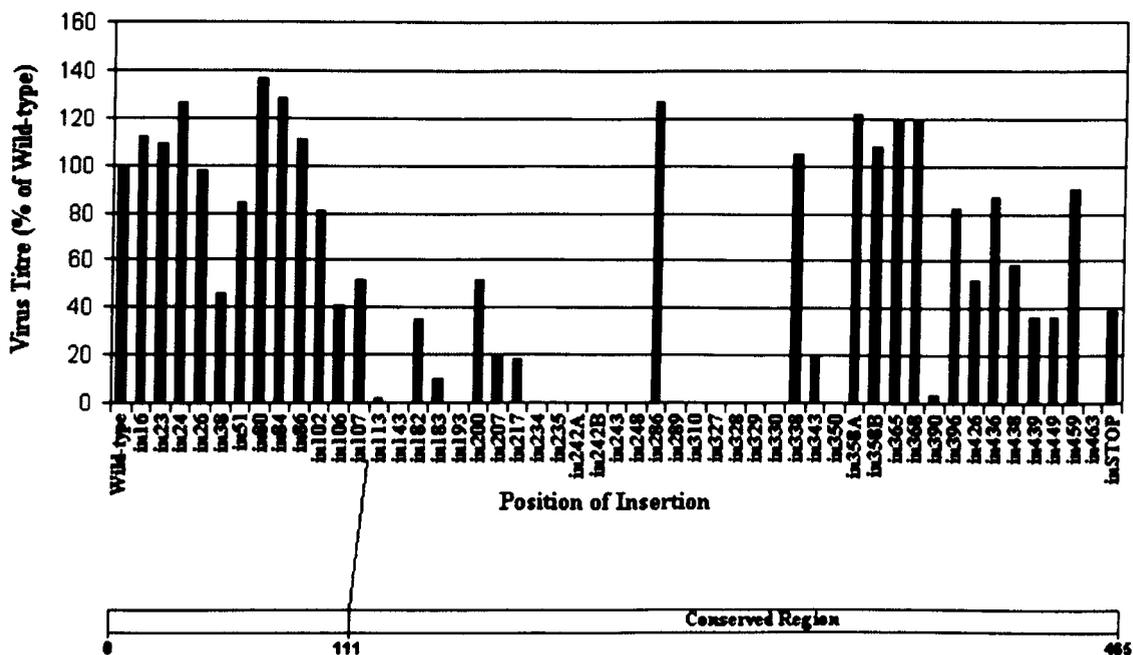


FIG. 1. Functional analysis of VP19C insertional mutants. (Top) Complementation of growth of the VP19C-minus mutant  $\nu\Delta 38$ YFP by transfected plasmids expressing the VP19C insertional mutants was carried out in BHK cells. The progeny virus was titrated on UL38RSC cells. The results are shown as percentages of the titer obtained with the wild-type VP19C control plasmid. (Bottom) Scale drawing of the UL38 ORF showing the extent of the poorly conserved N-terminal sequences (Fig. 7a).

must interact correctly with the other triplex protein, VP23, and with the major capsid protein, VP5. Since these interactions are essential, they can be assumed to have formed correctly in those VP19C mutants that complemented  $\nu\Delta 38$ YFP. Therefore, only the 18 mutants that gave less than 10% of the yield of wild-type VP19C and were considered to be severely disabled were examined. Previous studies have shown that VP19C is a nuclear protein. VP23 and VP5, when expressed in the absence of other virus proteins, are uniformly distributed throughout the cell but become predominantly nuclear in the presence of VP19C (25). Furthermore, immunofluorescence analysis of  $\nu\Delta 38$ YFP-infected BHK cells revealed that VP23 remained distributed throughout the cell (data not shown), demonstrating that its nuclear localization was entirely dependent on the presence of VP19C.

To examine these interactions, BHK cells were cotransfected with plasmids expressing the VP19C mutants and with pE18 (expressing VP23) (20) or with pE19 (VP5) (25), and the distributions of the proteins were determined by immunofluorescence. One of the mutant VP19Cs (in330) was no longer able to locate to the nucleus, making it impossible to determine its effect on the distribution of the other proteins (Fig. 2 and 3).

When the interaction between VP23 and the remaining 17 severely disabled VP19C mutants was examined in this assay, two distinct patterns were seen (Fig. 2). Nine mutants (in143, in234, in235, in327, in328, in329, in350, in390, and in463) exhibited a wild-type pattern, with VP19C and VP23 colocalizing to the nucleus. In the case of in390, the VP19C antibody was unable to bind to the mutant protein. However, VP23 was clearly transported to the nucleus, implying that the insertion

had disrupted the antibody-binding site but had not interfered with the ability of VP19C to interact with VP23 or to localize to the nucleus. In the remaining eight mutants (in113, in193, in242A, in242B, in243, in248, in289, and in310), VP19C localized to the nucleus as expected but VP23 remained in the cytoplasm, implying that the inserted sequences had disrupted the interaction between VP19C and VP23.

Seven severely disabled VP19C mutants (in113, in242A, in242B, in289, in328, in329, and in463) were unable to transport VP5 to the nucleus (Fig. 3). Of these, four (in113, in242A, in242B, and in289) also failed to transport VP23 to the nucleus, while the other three (in328, in329, and in463) interacted successfully with VP23. The results of the VP23 and VP5 analyses are summarized in Table 3.

**Role of the VP19C N terminus.** Studies examining HSV-1 capsid formation using a recombinant-baculoviruses-based model system, had shown that the N-terminal 90 amino acids of VP19C were not absolutely required for capsid assembly (30). The results of the complementation analysis (Fig. 1) are consistent with this conclusion, as all of the insertions into the N-terminal 107 codons of the UL38 ORF were able to support virus growth. To investigate the roles of these N-terminal sequences further, plasmids *pUL38-45FBpCl* and *pUL38-63FBpCl*, expressing VP19C proteins lacking the N-terminal 45 (VP19C-45) and 63 (VP19C-63) amino acids, respectively, were constructed and tested for the ability to complement growth of  $\nu\Delta 38$ YFP. As shown in Fig. 4a, the yield obtained with VP19C-45 was approximately 50-fold lower than for the VP19C-WT control, while removal of the first 63 amino acids completely abolished complemen-

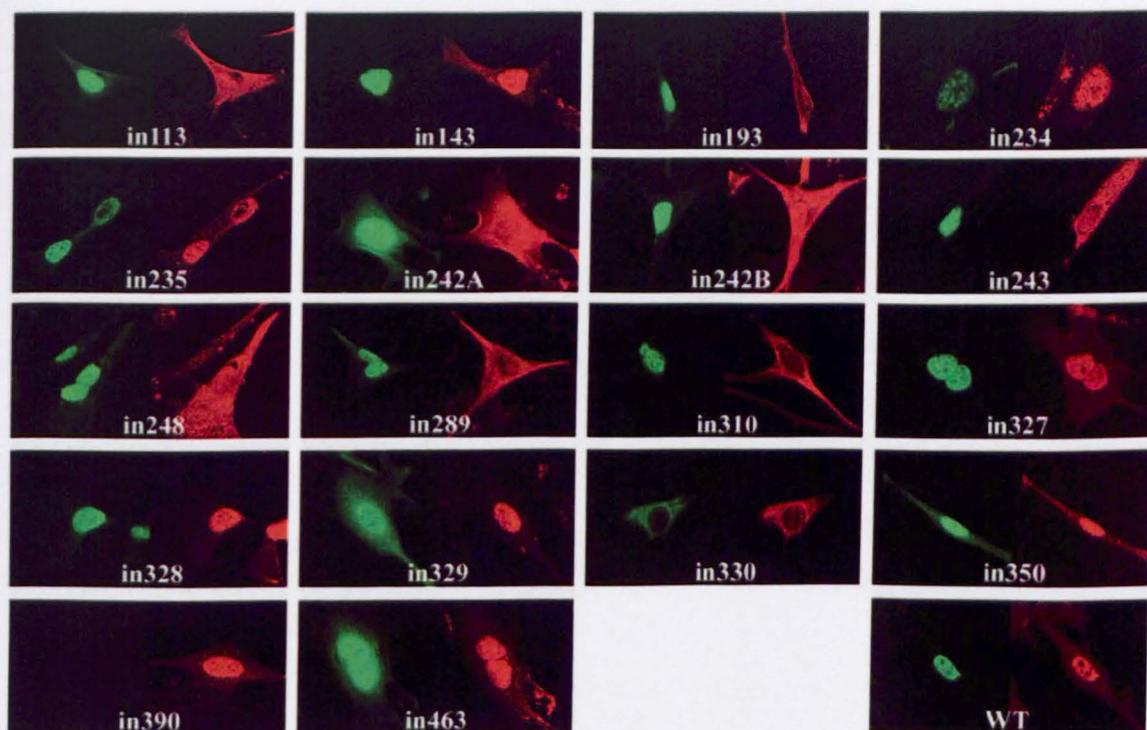


FIG. 2. Influence of VP19C insertional mutants on the distribution of VP23. BHK cells were cotransfected with pE18 (expressing VP23) and either pUL38FBpCI (wild-type [WT] VP19C) or one of the VP19C insertional mutants as indicated. VP19C was detected with the monoclonal antibody mAb02040 and visualized using FITC GAM (green). VP23 was detected with the antiserum rAb186 and visualized using TRITC-conjugated GAR (red).

tation, thereby demonstrating that the N-terminal sequences of VP19C are important for production of wild-type levels of infectious virus.

To determine the effects of the N-terminal truncations on the behavior of VP19C, the distributions of VP19C-45 and VP19C-63 in transfected cells were examined. In marked contrast to the pattern seen with VP19C-WT (Fig. 2 and 3), neither VP19C-45 (Fig. 4b) nor VP19C-63 (Fig. 4e) was specifically nuclear. These results suggested that these sequences contain an NLS. Since the nucleus is the site of capsid assembly, the smaller amounts of the mutant VP19C entering this compartment could account for the reduction in virus yield seen in the complementation assay. To determine whether this was the case, the SV40 large T antigen NLS (10) was fused to the N-termini of VP19C-45 and VP19C-63. The resulting proteins, VP19C-45NLS (Fig. 4c) and VP19C-63NLS (Fig. 4f), both localized efficiently to the nucleus, and in addition, both transported VP23 (Fig. 4d and g) and VP5 (not shown) to the nucleus, confirming that the exogenous NLS was functioning as expected. When the abilities of these constructs to support virus growth were tested, VP19C-45NLS was found to complement  $\nu\Delta 38$ YFP as effectively as VP19C-WT (Fig. 4a), thereby demonstrating that the reduction in growth caused by the removal of the N-terminal 45 amino acids was primarily a result of the defect in nuclear localization. In contrast, no complementation occurred with VP19C-63NLS (Fig. 4a), suggesting

that amino acids 45 to 63 have an additional role in the formation of infectious virus.

**Mapping the VP19C NLS.** Although VP19C is transported to the nucleus in the absence of other viral proteins, no NLS was identified when the complete amino acid sequence was submitted to the NLS prediction program (<http://cubic.bioc.columbia.edu/predictNLS/> and <http://psort.nibb.ac.jp/form2.html>). To confirm that the N-terminal region of VP19C did contain an NLS, increasing amounts of sequence from the N terminus were fused to the GFP, and the intracellular distributions of the chimeric proteins were determined. In cells expressing GFP alone, the fluorescence was uniformly distributed between the cytoplasm and the nucleus (Fig. 5a). However, fusing GFP to the N-terminal 118 (not shown), 83, 76, 66, or 56 amino acids of VP19C converted it to a nuclear protein (Fig. 5b to e). In contrast, the distribution seen with the N-terminal 46 amino acids resembled the GFP control (Fig. 5f). Thus, it is clear that the NLS is located within the first 56 residues of VP19C. Examination of the amino acid sequence in this region revealed a high concentration of arginine residues extending from amino acids 25 to 57 (Fig. 5i). Since stretches of basic amino acids are characteristic of many NLSs, the entire arginine-rich sequence from amino acids 24 to 66 was fused to GFP. This also resulted in nuclear localization of GFP (Fig. 5g), confirming that the NLS mapped in this region. In an attempt to define the NLS more precisely, amino acids 33 to 66

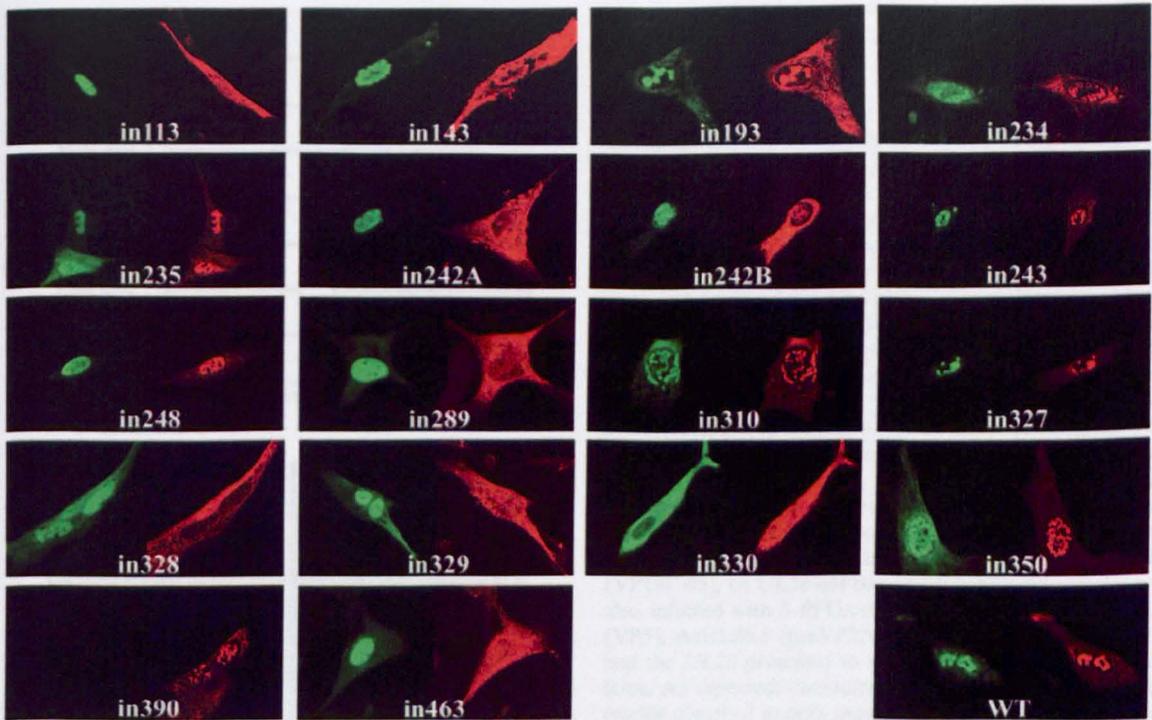


FIG. 3. Influence of VP19C insertional mutants on the distribution of VP5. BHK cells were cotransfected with pE19 (expressing VP5) and either pUL38FBpCI (wild-type [WT] VP19C) or one of the VP19C insertional mutants as indicated. VP19C was detected with the monoclonal antibody mAb02040 and visualized using FITC GAM (green). VP5 was detected with the antiserum rAb184 and visualized using TRITC-conjugated GAR (red).

were also fused to GFP. However, the resulting chimera showed an unexpected and distinctive cytoplasmic pattern that resembled the distribution expected for mitochondria (Fig. 5h). Submission of the VP19C sequence to the subcellular

localization site prediction program iPSORT (<http://biocaml.org/ipsort/iPSORT/#predict>) identified amino acids 32 to 40 (LLRRVLRPP) as a potential mitochondrial localization signal. Given what is known about the properties and functions of VP19C, it seems unlikely that this mitochondrial localization signal is functional in the context of the intact protein. Presumably, the NLS, which occupies the same region, is dominant and the mitochondrial signal becomes apparent only after the increasingly large deletions have rendered the NLS inactive. Although further analysis is required to determine the precise composition of the VP19C NLS, the results presented here map it unambiguously to an arginine-rich sequence between residues 24 and 56.

#### Nature of the growth defect associated with VP19C-63NLS.

To determine at what stage during infection the block on replication due to deletion of the N-terminal 63 amino acids was occurring, duplicate 35-mm plates of U2OS cells (one containing a glass coverslip) were incubated with 50 PFU/cell of baculoviruses, UL38FBacpCI (expressing VP19C), UL38-45NLSFBacpCI (VP19C-45NLS), and UL38-63NLSFBacpCI (VP19C-63NLS), for 1 h at 37°C. The baculovirus inoculum was removed, the cells were infected with 5 PFU/cell of  $\nu\Delta 38$ YFP, and incubation was continued at 37°C. After 16 h, the coverslip from one of the duplicate plates was removed and processed for immunofluorescence microscopy, while the remaining cells were used for Western blotting. The second plate

TABLE 3. Intracellular distribution of severely disabled VP19C insertional mutants

VP19C mutant	Localizes to nucleus <sup>a</sup>	Takes VP23 to nucleus <sup>a</sup>	Takes VP5 to nucleus <sup>a</sup>
in113	+	-	-
in143	+	+	+
in193	+	-	+
in234	+	+	+
in235	+	+	+
in242A	+	-	-
in242B	+	-	-
in243	+	-	+
in248	+	-	+
in289	+	-	-
in310	+	-	+
in327	+	+	+
in328	+	+	-
in329	+	+	-
in330	-	-	-
in350	+	+	+
in390	+	+	+
in463	+	+	-

<sup>a</sup> +, yes; -, no.

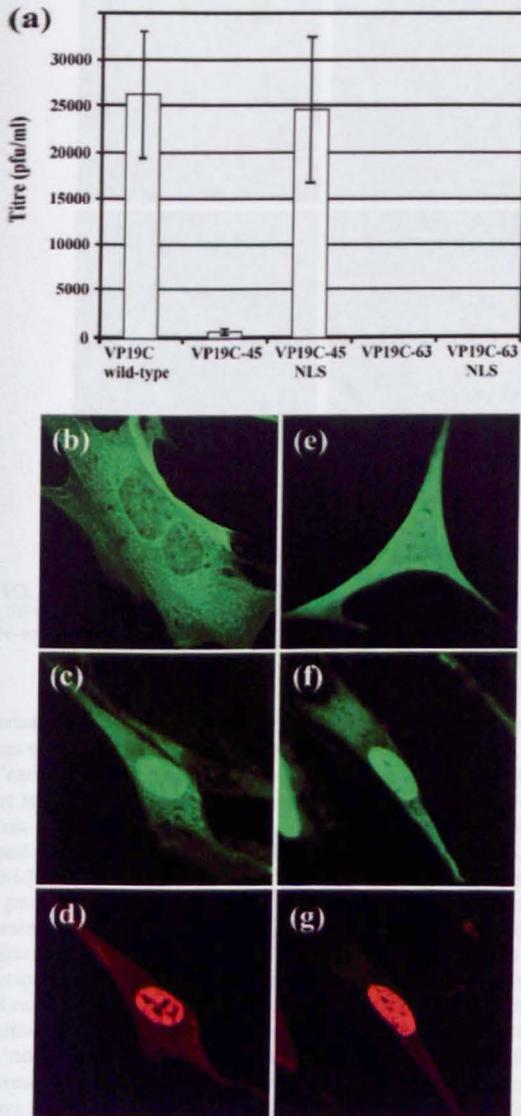


FIG. 4. Roles of the VP19C N-terminal sequences. (a) Complementation of growth of the VP19C-minus mutant v $\Delta$ 38YFP by transfected plasmid pUL38-45FBpCI (expressing VP19C-45), pUL38-63FBpCI (VP19C-63), pUL38-45NLSFBpCI (VP19C-45NLS), or pUL38-63NLSFBpCI (VP19C-63NLS) was carried out in BHK cells. The progeny virus was titrated on UL38RSC cells. The error bars indicate the standard errors of the means. (b to g) Intracellular localization of N-terminal mutants of VP19C. BHK cells were transfected singly with plasmids (b) pUL38-45FBpCI (expressing VP19C-45) and (c) pUL38-63FBpCI (VP19C-63) or were cotransfected with (d) pUL38-45NLSFBpCI (VP19C-45NLS) and pE18 (VP23) or with (f and g) pUL38-63NLSFBpCI (VP19C-63NLS) and pE18. VP19C (b, c, e, and f) was detected with the monoclonal antibody mAb02040 and visualized using FITC-conjugated GAM (green). VP23 (d and g) was detected with the antiserum rAb186 and visualized using TRITC-conjugated GAR (red).

was harvested and processed for electron microscopy. Immunofluorescence showed similar numbers of cells expressing the three forms of VP19C (data not shown), and Western blotting (Fig. 6f) confirmed that all three proteins were expressed at similar levels. Examination of thin sections revealed the presence of capsids in the nuclei of 48% of the UL38FBacpCI-transduced and 57% of the UL38-45NLSFBacpCI-transduced samples. As expected for VP19C proteins that support productive infection, both DNA-containing and empty capsids were seen in cells expressing wild-type VP19C (Fig. 6a) and VP19C-45NLS (Fig. 6b). In contrast, no capsid-like particles of any type were seen in over 100 nuclei of the UL38-63NLSFBacpCI-transduced samples. Published studies had reported that capsid assembly in a baculovirus model was not prevented by the removal of up to 90 amino acids from the N terminus of VP19C, although it was less efficient than with full-length VP19C (30). Therefore, the failure to detect capsids in the UL38-63NLSFBacpCI-transduced U2OS cells was unexpected. To investigate this further, capsid assembly was analyzed using the much more efficient baculovirus model system. Duplicate plates of Sf21 cells (one containing a coverslip) were infected separately with 5 PFU/cell of baculovirus AcUL38 (expressing full-length VP19C), UL38-45FBac (VP19C-45), or UL38-63FBac (VP19C-63). All samples were also infected with 5 PFU/cell each of baculoviruses AcUL19 (VP5), AcUL26.5 (preVP22a), and AcAB3.12 (VP23, VP26, and the UL26 protease) to supply the remaining capsid proteins. As expected, characteristic herpesvirus B capsids were readily observed in cells expressing wild-type VP19C (Fig. 6c) and in smaller amounts in cells expressing VP19C-45 (Fig. 6d). The reduced number of intact capsids seen with VP19C-45 is probably a result of inefficient nuclear localization due to the absence of the NLS. However, in cells expressing VP19C-63, only incomplete capsid shells were seen (Fig. 6e). Once again, immunofluorescence (data not shown) and Western blotting (Fig. 6f) confirmed that all three proteins were being expressed at similar levels, thereby demonstrating that the block on particle formation was a direct result of the VP19C-63 deletion.

## DISCUSSION

Members of the three subfamilies of mammalian and avian herpesviruses (*Alpha*-, *Beta*-, and *Gammaherpesviridae*) share a common capsid architecture, and in all examples analyzed to date, the triplex is composed of one copy of a VP19C homologue (the  $\alpha$ -subunit) and two copies of a VP23 homologue (the  $\beta$ -subunit) (3, 6, 40). In HSV-1 and most other alphaherpesviruses, the  $\alpha$ -subunit is considerably larger than the  $\beta$ -subunit. However, in beta- and gammaherpesviruses, the  $\alpha$ -subunit is significantly smaller than VP19C (6, 13, 21) and is of a size similar to that of the  $\beta$ -subunit. Therefore, it seems likely that the additional sequences in the alphaherpesvirus VP19C homologues represent relatively recent evolutionary developments. Comparison of  $\alpha$ -subunit sequences reveals that the N-terminal 111 amino acids of VP19C are not conserved in the other alphaherpesviruses (Fig. 7a). However, differences in the N-terminal region are not entirely responsible for the size variation among the different herpesvirus subfamilies. An alignment of sequences from alpha-, beta-, and gammaherpesviruses is in broad agreement with earlier findings (38) and indicates that the greater size of the alphaherpesvirus proteins is largely accounted for by two

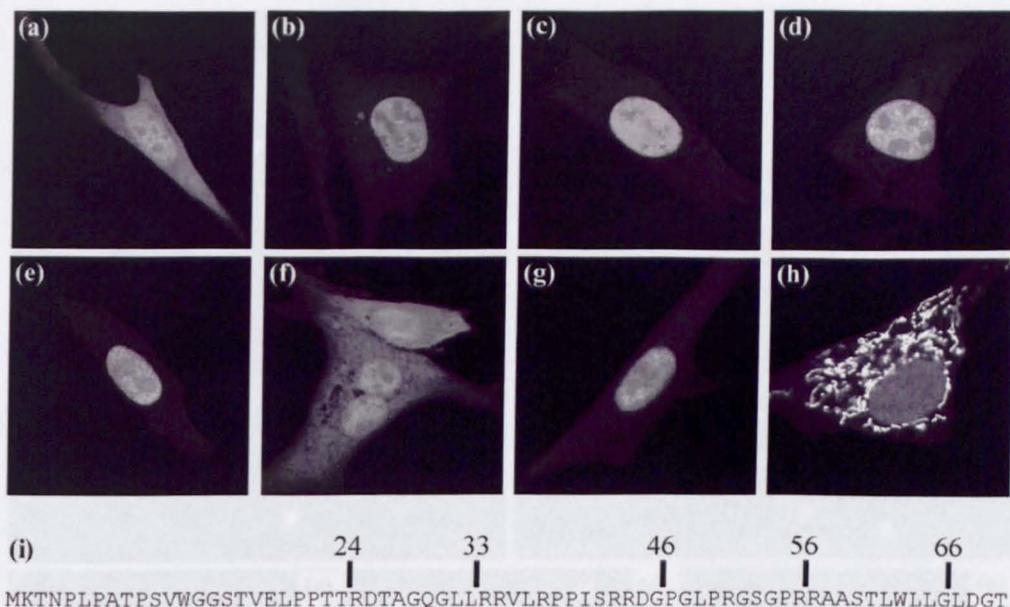


FIG. 5. Mapping the VP19C NLS. BHK cells were transfected with plasmids expressing the GFP protein fused in frame to the N-terminal 83 (b), 76 (c), 66 (d), 56 (e), or 46 (f) amino acids of VP19C. (g and h) GFP is fused to amino acids 24 to 66 and 33 to 66 of VP19C. (a) GFP control. (i) N-terminal 70 amino acids of VP19C with the positions of various truncations indicated. Arginine residues are underlined.

internal insertions, each of ~40 to 50 residues (insert 1 and insert 2), as well as by extra sequences at the N terminus (Fig. 7b).

Transposon-based insertional mutagenesis is a rapid and efficient method of screening a protein sequence for functional domains. However, the results of mutational analysis must be interpreted with caution, particularly when considering the behavior of individual mutants. In some cases, pairs of mutants with contrasting properties may be very close together. Thus, the plasmids containing insertions in286 and in289 differ markedly in their abilities to complement growth of v $\Delta$ 38YFP. In addition, not all functions are equally sensitive to insertion. For example, although in38 maps within the sequences identified as containing a nuclear localization signal, the ability of this mutant to enter the nucleus was not affected (data not shown). Therefore, without detailed information on the structure of the target protein or specific assays to measure particular properties, it is usually not possible to make other than a general interpretation of results. Three-dimensional structural analysis of the capsid has shown that the component proteins interact in different ways at distinct locations in the capsid (42). Furthermore, the reconfiguration that occurs during capsid maturation involves domain movements leading to changes in the patterns of interaction (8, 37). This type of behavior means that the importance of particular regions of a protein may vary with maturation status and location. Nevertheless, the distribution of severely disabling mutations across VP19C was nonuniform, with 16 in the central region of VP19C but none in the N-terminal 107 residues and only 2 in the C-terminal 107 residues. The most N-terminal severely disabled mutant, in113, occurs almost immediately after the first evidence of conservation between the sequences of the alphaherpesvirus  $\alpha$ -subunits (Fig. 7a). There is no equivalent distinction marking the boundary of

the C-terminal 107 residues, although interestingly, its start at in358 coincides with the beginning of the second large insert unique to alphaherpesviruses (insert 2) (Fig. 7b). However, unlike the poorly conserved N-terminal region, insert 2 contains a severely disabling mutation, in390. Only one insertion (in143) mapped within the insert 1 locus, and it was also severely disabling, indicating that this inserted sequence is functionally important. The distribution of severely disabled mutants suggests that VP19C can be divided into three regions based on the ability to tolerate insertions and may suggest that the three regions have differing roles. However, this interpretation must be treated with caution, as the distribution of insertions is not uniform and clustering of mutations with similar phenotypes, such as in327 to -330, tends to distort the pattern. Statistical analysis by chi-square analysis indicates that, although the N-terminal region is more tolerant of insertions, the perceived difference between the central and C-terminal regions is marginal.

The sensitivity to insertion of the central region suggests that it may be important for folding of VP19C in a manner similar to that suggested for the central region (VP5ud) of the major capsid protein, VP5, which has been proposed to act as a folding nucleus for the more flexible N- and C-terminal regions (2). The large number of tolerated insertions in the C-terminal region is rather surprising, as deletional analysis had shown that removal of as few as 15 amino acids from the C terminus destroyed the ability of VP19C to support capsid assembly (30). When considered alongside the behavior of in463, which completely fails to support virus growth, this seems to indicate that sequences very near the C terminus of VP19C play essential roles in capsid assembly. However, the viability of most of the mutants in the C-terminal region suggests that it has con-

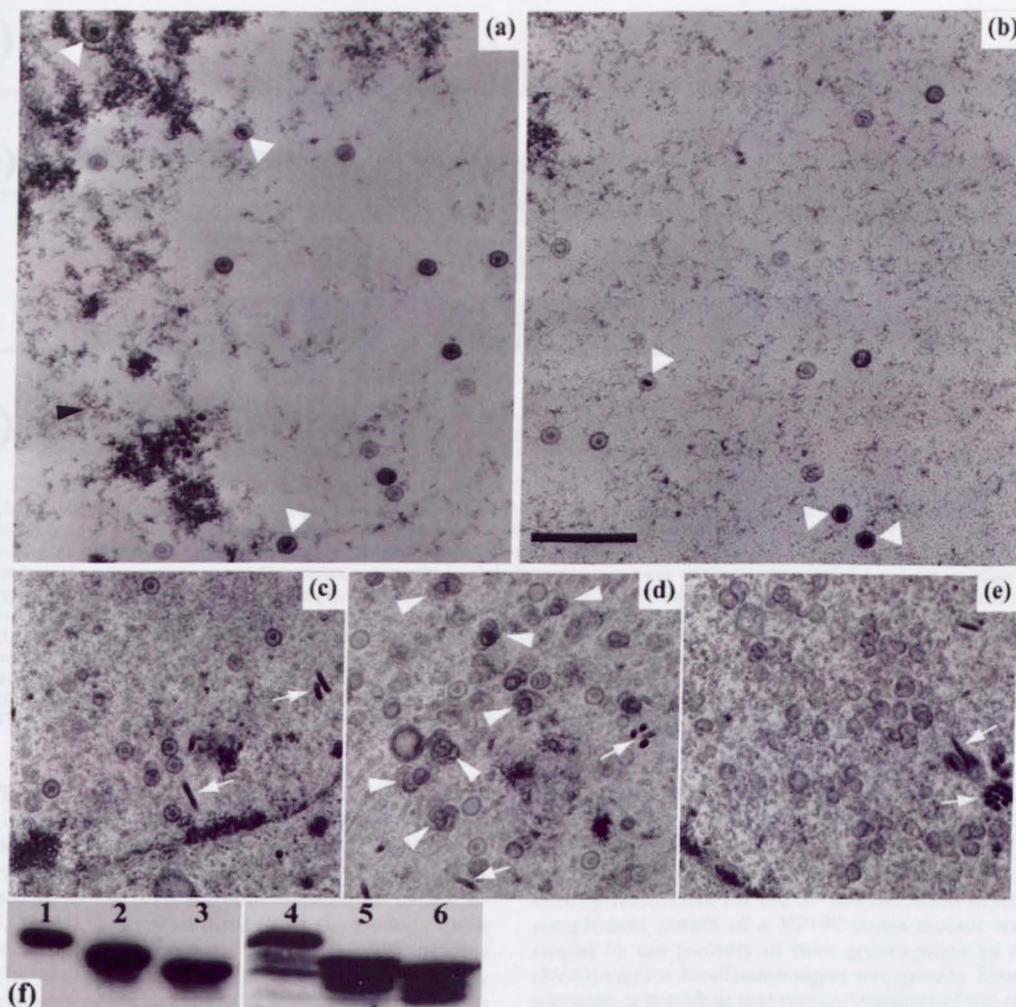


FIG. 6. Effects of VP19C N-terminal mutations on capsid assembly. (a and b) Thirty-five-millimeter plates of U2OS cells were transduced with 50 PFU/cell of baculovirus UL38FBacpCI (a) or UL38-45NLSFBacpCI (b) and incubated at 20°C for 1 h before being infected with 5 PFU/cell of  $\Delta$ 38YFP. Infection was continued at 37°C for 16 h. (c to e) Thirty-five-millimeter plates of Sf21 cells were infected with 5 PFU/cell each of baculoviruses AcUL19, AcUL26.5, and AcAB3.12 and with 5 PFU/cell of AcUL38 (c), UL38-45FBac (d), or UL38-63FBac (e) and incubated at 28°C for 48 h. At the appropriate times after infection, the cells were harvested and prepared for electron microscopy. The majority of capsids in panels a to d are B capsids and are unlabeled. In panels a and b, the C (DNA-containing) capsids are indicated by triangles. In panel d, concentrations of incomplete capsids are indicated by arrowheads. Only incomplete capsids are present in panel e, and they are not labeled. Baculovirus capsids (c to e) are indicated by arrows. Scale bar = 500 nm. (f) Expression of VP19C in the infected cells was confirmed by Western blotting. Protein samples were separated on a 10% sodium dodecyl sulfate-polyacrylamide gel and detected by enhanced chemiluminescence using mAb02040 and protein A-peroxidase. Lanes 1 and 2 show U2OS cell samples corresponding to panels a and b. Lane 3 shows a sample from U2OS cells transduced with UL38-63NLSFBacpCI and infected with  $\Delta$ 38YFP. Lanes 4 to 6 show Sf21 cell samples corresponding to panels c to e.

siderable structural flexibility despite the secondary-structure predictions, which appear to show that conservation of secondary structure increases toward the C terminus (Fig. 7c).

It is difficult to discern any pattern in the interaction properties of the severely disabled insertional mutants. The 18 severely disabled mutants can be divided into four categories based on the ability to interact with VP5 and VP23 (Table 3). Five (including in242A and in242B, which are in the same codon) failed to transport either VP23 or VP5 to the nucleus,

suggesting that these insertions may have altered the overall folding of the protein. These insertions (residues 113, 242, 289, and 330) are evenly spread across the central region and do not appear to demarcate any obvious domain, although they may lie close together in the folded protein. Four of the mutants interact with VP5 but not with VP23, while in three, this pattern is reversed. The specificity of these mutants suggests that the insertions do not cause global misfolding of the protein but have a more localized effect. The mutants that failed

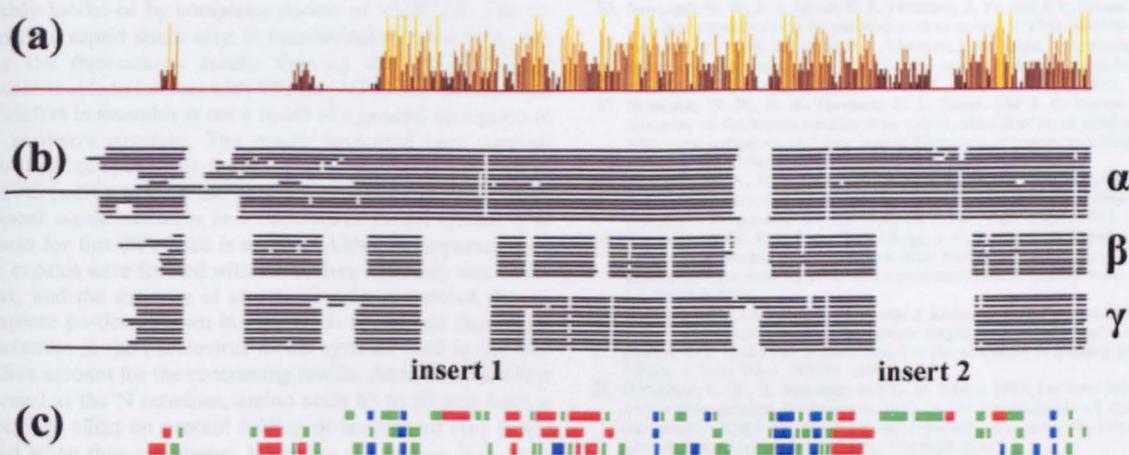


FIG. 7. Sequence comparison of VP19C homologues. We used the CLUSTAL W (<http://www.ebi.ac.uk/clustalw/>) multiple sequence alignment program (34) to align the triplex  $\alpha$ -subunit sequences of 15 alphaherpesviruses, 9 betaherpesviruses, and 12 gammaherpesviruses identified by a BLAST search from the VIDA database ([http://www.biochem.ucl.ac.uk/bsm/virus\\_database/VIDA.html](http://www.biochem.ucl.ac.uk/bsm/virus_database/VIDA.html)) using the HSV-1, HCMV, and Kaposi's sarcoma-associated herpesvirus triplex  $\alpha$ -subunit sequences as input. The CLUSTAL W output was used as the basis for subsequent analyses. (a) Levels of sequence conservation among the triplex  $\alpha$ -subunits in alphaherpesviruses were analyzed and plotted using the JalView (<http://www.jalview.org/>) multiple sequence editor program (4). (b) Schematic representation of the consensus alignment between alpha-, beta- and gamma-herpesvirus triplex  $\alpha$ -subunit sequences, showing the locations of additional sequences (insert 1 and insert 2 comprising residues 140 to 176 and 357 to 405, respectively, in HSV-1) in the alphaherpesvirus proteins. The consensus sequence was calculated using the program Consensus (<http://www.bork.embl-heidelberg.de/Alignment/consensus.html>). (c) Secondary-structure predictions for triplex  $\alpha$ -subunits of alpha- (top), beta- (middle), and gammaherpesviruses (bottom). Secondary-structure predictions were carried out on each individual sequence using PredictProtein (28) (<http://cubic.bioc.columbia.edu/predictprotein/>). Structural elements present in >50% of examples were plotted relative to the alignments shown in Fig. 7b for each subfamily. The positions of  $\alpha$ -helices are shown in red,  $\beta$ -sheet in blue, and coil in green.

to take VP23 to the nucleus were located N terminal (between residues 193 and 310) to those that failed to bind VP5 (between 328 and 463), suggesting that the regions involved are particularly important for binding of VP23 and VP5, respectively. Since VP23 and VP5 are both essential for capsid assembly (5, 33, 35), it is not surprising that the mutants which fail to interact with one or both of them do not support virus growth. The nature of the defect in the remaining severely disabled mutants is unclear. Further analysis is required to determine whether they support the assembly of structurally normal capsids. However, examination of mutants in143, in234, in235, in327, in350, and in390 in a plasmid-based DNA-packaging assay (9) showed that they were unable to support DNA packaging in  $\nu\Delta 38$ YFP-infected cells (data not shown). This suggests either that assembly of the capsid shell was disrupted, presumably due to some subtle alteration in the interactions with VP23 or VP5, or that an interaction involving one of the other virion components had been affected. VP19C has been reported to interact with the UL25 protein (22). The tegument protein present at the vertices in mature virions is in contact with the peripentonal triplexes and may also interact with VP19C (41). Failure to bind tegument would be lethal but would not be expected to prevent DNA packaging. However, UL25 mutants show aberrant DNA packaging (12, 32), and disruption of this interaction might result in the type of phenotype seen here.

The N-terminal regions of all three herpesvirus subfamilies are particularly poorly conserved in both length and sequence, with effectively no homology evident until after amino acid 111 of VP19C (Fig. 7a). This raises questions regarding the func-

tion of this highly variable region. Part of that function was revealed by the deletion of 45 residues from the N terminus of VP19C, which removed a nuclear localization signal. That the failure of this deleted protein to concentrate in the nucleus was entirely responsible for the 50-fold reduction in its ability to complement growth of a VP19C-minus mutant was demonstrated by the recovery in virus growth when an exogenous (SV40) nuclear localization signal was present. From this observation, it is evident that correct transport of the component proteins to the site of capsid assembly is an important function of VP19C. Although the precise sequence of the VP19C NLS has not been identified, it has been mapped to a region containing 33 amino acids, and database searching has revealed that it does not belong to any of the known classes of NLS. Interestingly, the pattern of arginine residues seen in the VP19C NLS is not present in the N-terminal sequences of any other alphaherpesvirus VP19C homologue, apart from that of the very closely related HSV-2. Furthermore, examination of triplex  $\alpha$ -subunit sequences from all three herpesvirus subfamilies using the predictNLS program failed to identify any recognizable NLSs. Therefore, it is unclear whether a role in nuclear localization represents the usual function of the poorly conserved alphaherpesvirus N termini or where NLSs, if any are present, might be located in the  $\alpha$ -subunits of beta- and gammaherpesviruses.

Since the functional NLS is contained within the first 56 amino acids of VP19C, the function of the rest of the poorly conserved N-terminal region remains unclear. Electron microscopic analysis failed to identify capsid formation by the VP19C-63 truncation, either in a baculovirus-based capsid as-

sembly model or by complementation of  $\nu\Delta 38YFP$ . The incomplete capsid shells seen in baculovirus-infected cells confirm the fluorescence results showing that the VP19C-63 protein is able to interact with VP5 and VP23 and indicate that the defect in assembly is not a result of a general disruption of the protein's structure. The results presented here contrast with those of Spencer et al. (30), who showed that a mutant of VP19C (nd90) lacking the N-terminal 90 amino acids could support capsid assembly in a baculovirus model system. The reason for this difference is unclear. Although apparently intact capsids were formed with nd90, their assembly was inefficient, and the majority of structures seen resembled the incomplete particles shown in Fig. 5e. It is possible that slight differences in the baculovirus model systems used in the two studies account for the contrasting results. Alternatively, when exposed at the N terminus, amino acids 63 to 90 may have a disruptive effect on protein folding or interaction that is negated when they are absent. Whatever the reason, it is clear that sequences between residues 45 and 63 are necessary for efficient capsid formation, and further analysis will be needed to determine precisely what roles they play.

#### ACKNOWLEDGMENTS

We thank D. McGeoch for critical reading of the manuscript and J. Mitchell for excellent technical assistance. pFBpC1 was provided by R. Everett.

This work was supported by the United Kingdom Medical Research Council. W. E. Adamson was funded by a MRC Research Training studentship.

#### REFERENCES

- Bishop, D. H. L. 1992. Baculovirus expression vectors. *Semin. Virol.* 3:253–264.
- Bowman, B. R., M. L. Baker, F. J. Rixon, W. Chiu, and F. A. Quijcho. 2003. Structure of the herpesvirus major capsid protein. *EMBO J.* 22:757–765.
- Chen, D. H., H. Jiang, M. Lee, F. Y. Liu, and Z. H. Zhou. 1999. Three-dimensional visualization of tegument/capsid interactions in the intact human cytomegalovirus. *Virology* 260:10–16.
- Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton. 2004. The Jalview Java alignment editor. *Bioinformatics* 20:426–427.
- Desai, P., N. A. Deluca, J. C. Glorioso, and S. Person. 1993. Mutations in herpes simplex virus type 1 genes encoding VP5 and VP23 abrogate capsid formation and cleavage of replicated DNA. *J. Virol.* 67:1357–1364.
- Gibson, W., M. K. Baxter, and K. S. Clopper. 1996. Cytomegalovirus "missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46. *J. Virol.* 70:7454–7461.
- Grunewald, K., P. Desai, D. C. Winkler, J. B. Heymann, D. M. Belnap, W. Baumeister, and A. C. Steven. 2003. Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science* 302:1396–1398.
- Heymann, J. B., N. Q. Cheng, W. W. Newcomb, B. L. Trus, J. C. Brown, and A. C. Steven. 2003. Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nat. Struct. Biol.* 10:334–341.
- Hodge, P. D., and N. D. Stow. 2001. Effects of mutations within the herpes simplex virus type 1 DNA encapsidation signal on packaging efficiency. *J. Virol.* 75:8977–8986.
- Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith. 1984. Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 311:33–38.
- Lu, G. Y., Z. H. Zhou, M. L. Baker, J. Jakana, D. Y. Cai, X. C. Wei, S. X. Chen, X. C. Gu, and W. Chiu. 1998. Structure of double-shelled rice dwarf virus. *J. Virol.* 72:8541–8549.
- McNab, A. R., P. Desai, S. Person, L. L. Roof, D. R. Thomsen, W. W. Newcomb, J. C. Brown, and F. L. Homa. 1998. The product of the herpes simplex virus type 1 UL25 gene is required for encapsidation but not for cleavage of replicated viral DNA. *J. Virol.* 72:1060–1070.
- Nealon, K., W. W. Newcomb, T. R. Pray, C. S. Craik, J. C. Brown, and D. H. Kedes. 2001. Lytic replication of Kaposi's sarcoma-associated herpesvirus results in the formation of multiple capsid species: isolation and molecular characterization of A, B, and C capsids from a gammaherpesvirus. *J. Virol.* 75:2866–2878.
- Newcomb, W. W., F. L. Homa, D. R. Thomsen, F. P. Booy, B. L. Trus, A. C. Steven, J. V. Spencer, and J. C. Brown. 1996. Assembly of the herpes simplex virus capsid: characterization of intermediates observed during cell-free capsid formation. *J. Mol. Biol.* 263:432–446.
- Newcomb, W. W., F. L. Homa, D. R. Thomsen, Z. Ye, and J. C. Brown. 1994. Cell-free assembly of the herpes simplex virus capsid. *J. Virol.* 68:6059–6063.
- Newcomb, W. W., R. M. Juhás, D. R. Thomsen, F. L. Homa, A. D. Burch, S. K. Weller, and J. C. Brown. 2001. The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *J. Virol.* 75:10923–10932.
- Newcomb, W. W., D. R. Thomsen, F. L. Homa, and J. C. Brown. 2003. Assembly of the herpes simplex virus capsid: identification of soluble scaffold-portal complexes and their role in formation of portal-containing capsids. *J. Virol.* 77:9862–9871.
- Newcomb, W. W., B. L. Trus, F. P. Booy, A. C. Steven, J. S. Wall, and J. C. Brown. 1993. Structure of the herpes simplex virus capsid: molecular composition of the pentons and the triplexes. *J. Mol. Biol.* 232:499–511.
- Newcomb, W. W., B. L. Trus, N. Q. Cheng, A. C. Steven, A. K. Sheaffer, D. J. Tenney, S. K. Weller, and J. C. Brown. 2000. Isolation of herpes simplex virus procapsids from cells infected with a protease-deficient mutant virus. *J. Virol.* 74:1663–1673.
- Nicholson, P., C. Addison, A. M. Cross, J. Kennard, V. G. Preston, and F. J. Rixon. 1994. Localization of the herpes simplex virus type 1 major capsid protein VP5 to the cell nucleus requires the abundant scaffolding protein VP22a. *J. Gen. Virol.* 75:1091–1099.
- O'Connor, C. M., B. Damania, and D. H. Kedes. 2003. De novo infection with rhesus monkey rhadinovirus leads to the accumulation of multiple intranuclear capsid species during lytic replication but favors the release of genome-containing virions. *J. Virol.* 77:13439–13447.
- Ogasawara, M., T. Suzutani, I. Yoshida, and M. Azuma. 2001. Role of the UL25 gene product in packaging DNA into the herpes simplex virus capsid: location of UL25 product in the capsid and demonstration that it binds DNA. *J. Virol.* 75:1427–1436.
- Preston, V. G., M. F. Al-Kobaisi, I. M. McDougall, and F. J. Rixon. 1994. The herpes simplex virus gene UL26 proteinase in the presence of the UL26.5 gene product promotes the formation of scaffold-like structures. *J. Gen. Virol.* 75:2355–2366.
- Rixon, F. J. 1993. Structure and assembly of herpesviruses. *Semin. Virol.* 4:135–144.
- Rixon, F. J., C. Addison, A. McGregor, S. J. Macnab, P. Nicholson, V. G. Preston, and J. D. Tatman. 1996. Multiple interactions control the intracellular localization of the herpes simplex virus type 1 capsid proteins. *J. Gen. Virol.* 77:2251–2260.
- Rixon, F. J., and D. McNab. 1999. Packaging-competent capsids of a herpes simplex virus temperature-sensitive mutant have properties similar to those of in vitro-assembled procapsids. *J. Virol.* 73:5714–5721.
- Rosenthal, K. S., M. D. Leather, and B. G. Barisais. 1984. Herpes simplex virus binding and entry modulate cell surface protein mobility. *J. Virol.* 49:980–983.
- Rost, B., G. Yachdav, and J. Liu. 2004. The PredictProtein server. *Nucleic Acids Res.* 32:W321–W326.
- Saad, A., Z. H. Zhou, J. Jakana, W. Chiu, and F. J. Rixon. 1999. Roles of triplex and scaffolding proteins in herpes simplex virus type 1 capsid formation suggested by structures of recombinant particles. *J. Virol.* 73:6821–6830.
- Spencer, J. V., W. W. Newcomb, D. R. Thomsen, F. L. Homa, and J. C. Brown. 1998. Assembly of the herpes simplex virus capsid: preformed triplexes bind to the nascent capsid. *J. Virol.* 72:3944–3951.
- Steven, A. C., and P. G. Spear. 1997. Herpesvirus capsid assembly and envelopment, p. 312–351. *In* W. Chiu, R. M. Burnett, and R. Garcea (ed.), *Structural biology of viruses*. Oxford University Press, New York, N.Y.
- Stow, N. D. 2001. Packaging of genomic and amplicon DNA by the herpes simplex virus type 1 UL25-null mutant KUL25NS. *J. Virol.* 75:10755–10765.
- Tatman, J. D., V. G. Preston, P. Nicholson, R. M. Elliott, and F. J. Rixon. 1994. Assembly of herpes simplex virus type 1 capsids using a panel of recombinant baculoviruses. *J. Gen. Virol.* 75:1101–1113.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.
- Thomsen, D. R., L. L. Roof, and F. L. Homa. 1994. Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. *J. Virol.* 68:2442–2457.
- Thurlow, J. K., F. J. Rixon, M. Murphy, P. Targett-Adams, M. Hughes, and V. G. Preston. 2005. The herpes simplex virus type 1 DNA packaging protein UL17 is a virion protein that is present in both the capsid and the tegument compartments. *J. Virol.* 79:150–158.
- Trus, B. L., F. P. Booy, W. W. Newcomb, J. C. Brown, F. L. Homa, D. R. Thomsen, and A. C. Steven. 1996. The herpes simplex virus procapsid: structure, conformational changes upon maturation, and roles of the triplex proteins VP19c and VP23 in assembly. *J. Mol. Biol.* 263:447–462.
- Trus, B. L., J. B. Heymann, K. Nealon, N. Q. Cheng, W. W. Newcomb, J. C. Brown, D. H. Kedes, and A. C. Steven. 2001. Capsid structure of Kaposi's sarcoma-associated herpesvirus, a gammaherpesvirus, compared to those of an alphaherpesvirus, herpes simplex virus type 1, and a betaherpesvirus, cytomegalovirus. *J. Virol.* 75:2879–2890.
- Trus, B. L., F. L. Homa, F. P. Booy, W. W. Newcomb, D. R. Thomsen, N. Q. Cheng, J. C. Brown, and A. C. Steven. 1995. Herpes simplex virus capsids

- assembled in insect cells infected with recombinant baculoviruses: structural authenticity and localization of VP26. *J. Virol.* **69**:7362–7366.
40. **Yu, X. K., C. M. O'Connor, V. Atanasov, B. Damania, D. H. Kedes, and Z. H. Zhou.** 2003. Three-dimensional structures of the A, B, and C capsids of rhesus monkey rhadinovirus: insights into gammaherpesvirus capsid assembly, maturation, and DNA packaging. *J. Virol.* **77**:13182–13193.
41. **Zhou, Z. H., D. H. Chen, J. Jakana, F. J. Rixon, and W. Chiu.** 1999. Visualization of tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. *J. Virol.* **73**:3210–3218.
42. **Zhou, Z. H., W. Chiu, K. Haskell, H. Spears, J. Jakana, F. J. Rixon, and L. R. Scott.** 1998. Refinement of herpesvirus B-capsid structure on parallel supercomputers. *Biophys. J.* **74**:576–588.
43. **Zhou, Z. H., M. Dougherty, J. Jakana, J. He, F. J. Rixon, and W. Chiu.** 2000. Seeing the herpesvirus capsid at 8.5 Å. *Science* **288**:877–880.
44. **Zhou, Z. H., J. He, J. Jakana, J. D. Tatman, F. J. Rixon, and W. Chiu.** 1995. Assembly of VP26 in herpes simplex virus 1 inferred from structures of wild-type and recombinant capsids. *Nat. Struct. Biol.* **2**:1026–1030.
45. **Zhou, Z. H., S. J. Macnab, J. Jakana, L. R. Scott, W. Chiu, and F. J. Rixon.** 1998. Identification of the sites of interaction between the scald and outer shell in herpes simplex virus-1 capsids by difference electron imaging. *Proc. Natl. Acad. Sci. USA* **95**:2778–2783.
46. **Zhou, Z. H., B. V. V. Prasad, J. Jakana, F. J. Rixon, and W. Chiu.** 1994. Protein subunit structures in the herpes simplex virus A-capsid determined from 400 kV spot-scan electron cryomicroscopy. *J. Mol. Biol.* **242**:456–469.