The Role of the Papillomavirus E4 Protein During the Late Stage of the Viral Life Cycle Using *In Vivo* Model Systems

A THESIS SUBMITTED TO THE UNIVERSITY OF LONDON FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Woei Ling Peh

Division of Virology,
National Institute for Medical Research,
The Ridgeway, Mill Hill,
London, UK

2002
ACKNOWLEDGEMENT

The completion of this thesis is a dream come true. There have been many ups and downs, and numerous giving-ups along the way, but nonetheless, my PhD is finally finished! FINALLY! For those who had disbelieved (including myself during my dark days), this is for you.

I would like to dedicate this thesis especially to my parents, brother and sister, who are in Singapore. I wouldn’t have come so far without your endless and unconditional support. This day couldn’t have been possible without all of you.

To weiwei, nene and natalie, my dearest flatmates. You are my family in London. Thanks for being here and sharing your lives with me. You have contributed so much in the creation of this thesis.

A silent thank you, for those friends whom I’ve come to know during their time in London. Some of you have played a huge role (some knowingly, some unknowingly) in helping me to cope with the difficult times during the past few years of my PhD.

To my supervisor John, I hope it has been worth your time. Thanks for being so patient.

Not forgetting all those in the Papillomavirus Group at NIMR (Deb, Clare, Qian, Shirley, Pauline, Rina and Nicky). Thanks for keeping me as part of the group, even during my long absence. Also, thanks for proof-reading my chapters for me, now you know how bad my Singaporean english can be! Oh, and I'm back to claim a bench space!! And I know I still owe you guys a Chinese meal...

To Sir John Skehel, thank you for your continual support.

My PhD could never have begun, and be completed, without many prayers of some of my dearest friends.

"God grant me the serenity to accept the things I cannot change, the courage to change things I can, and the wisdom to know the difference."
Papillomaviruses make up a large family of viruses that can infect the epithelium of a range of human and animal hosts, and are capable of inducing a spectrum of different cutaneous and mucosal epithelial diseases. These viruses are highly tissue and host specific, and are dependent on epithelia differentiation for the completion of their life cycle. The late stages of the viral life cycle have been shown to occur only in differentiated epithelial cells which support viral DNA amplification and the expression of viral late proteins (E4, L1 and L2). These late events can be detected using in situ hybridisation and immunodetection methods in infected tissue sections. Specific E4 antibodies were made against 4 animal types (COPV, CRPV, ROPV and BPV-1) and 1 human type (HPV-11) papillomaviruses. Using specific E4 antibodies and viral DNA probes, the late stage of the viral life cycle were examined in tissue sections of naturally and experimentally induced infections. Different patterns of viral late events are described and compared. The usefulness of animal models as an alternative approach to study of the papillomavirus life cycle is also discussed.

The expression of the viral E4 protein and vegetative DNA replication had been shown to correlate exactly in vivo. However, the function of the E4 protein is still unclear. A series of E4 C-terminal truncation mutants were constructed along the E4 open reading frame by the introduction of stop codons using site-directed mutagenesis. Five different E4 truncation mutants were made in the CRPV and HPV-16 viral genomes. The shortest mutant for CRPV (labelled CRst9/0) and HPV-16 (labelled W12st15/6) was predicted to express an 8 or 14 amino acid E1\*E4 protein respectively. The CRPV wildtype and E4 mutant DNA were prepared and inoculated onto the backs of wild cottontail (natural) and New Zealand (experimental) rabbit hosts. This thesis presents the data obtained from the study of the E4 mutants in the CRPV-rabbit model systems. It is the first time the papillomavirus E4 protein has been shown to have an essential role in the initiation of viral DNA amplification in vivo.
CONTENTS

ABBREVIATIONS ........................................................................................................................................ 6

FIGURES AND TABLES .................................................................................................................................. 10

CHAPTER 1 - GENERAL INTRODUCTION ................................................................................................ 12
  1.1 Papillomaviruses .................................................................................................................................. 13
  1.2 Viral early (E) proteins ....................................................................................................................... 19
  1.3 Viral early (L) proteins ....................................................................................................................... 26
  1.4 The epithelium ..................................................................................................................................... 27
  1.5 PV-associated diseases ...................................................................................................................... 28
  1.6 The PV life cycle .................................................................................................................................. 31
  1.7 Immunity to PV infections .................................................................................................................. 35
  1.8 Experimental models for PV infection ................................................................................................. 38
  1.9 Project Objectives ................................................................................................................................ 45

CHAPTER 2 - GENERAL METHODS AND MATERIALS ................................................................................. 47

DNA Methods
  2.1 Bacterial transformation ..................................................................................................................... 50
  2.2 DNA preparations .............................................................................................................................. 52
  2.3 DNA cloning .......................................................................................................................................... 56
  2.4 Polymerase Chain Reaction ............................................................................................................... 57
  2.5 Oligonucleotides ................................................................................................................................... 58
  2.6 Synthesis of the E4 mutant genomes ................................................................................................. 59
  2.7 DNA Sequencing .................................................................................................................................... 69

Protein Methods
  2.8 Glutathione speharose-transferase and maltose binding protein fusion proteins ...................... 72
  2.9 Peptide conjugation ............................................................................................................................ 75
  2.10 Protein analysis .................................................................................................................................... 78
  2.11 Other protein methods ...................................................................................................................... 80

Antibodies and Tissue Processing Methods
  2.12 Preparation of specific antibodies ................................................................................................. 82
  2.13 Tissue preparations and histology .................................................................................................. 85
  2.14 Immuno-histochemistry .................................................................................................................. 86
  2.15 DNA fluorescent in situ hybridisation on tissue sections ............................................................... 90

CHAPTER 3 - E4 PROTEIN SEQUENCE ALIGNMENTS ............................................................................... 92
  3.1 Introduction ......................................................................................................................................... 93
  3.2 The E1^E4 transcripts and proteins ................................................................................................... 96
  3.3 E4 alignments between viruses from the same phylogenetic Supergroup/Group ...................... 103
  3.4 Other conserved features between the E4 proteins ......................................................................... 114
  3.5 Discussion .......................................................................................................................................... 121

CHAPTER 4 - E4 SPECIFIC ANTIBODIES ............................................................................................... 128
  4.1 Introduction ....................................................................................................................................... 129
  4.2 Preparation and purification of the E4 immunogens ...................................................................... 130
  4.3 Immunisation of animals .................................................................................................................... 136
  4.4 E4 specific antisera ............................................................................................................................ 138
  4.5 Discussion .......................................................................................................................................... 144

CHAPTER 5 - CHARACTERISATION OF THE EVENTS IN THE LATE STAGE OF THE VIRAL LIFE
  CYCLE: E4 AND L1 PROTEINS EXPRESSION AND VIRAL DNA AMPLIFICATION
  PATTERNS IN FORMALIN-FIXED, PARAFFIN-EMBEDDED PAPILLOMAVIRUS
  INFECTED TISSUES .................................................................................................................................. 146
  5.1 Introduction ....................................................................................................................................... 147
  5.2 Infected tissue specimens .................................................................................................................. 149
  5.3 Staining protocols and specifics ........................................................................................................ 150
5.4 E4 antibody staining and the general distribution of the E4 protein in infected tissue sections ................................................. 150
5.5 Viral DNA amplification and E4 expression correlate exactly in infected tissues .............................................................. 155
5.6 E4 expression precedes L1 in all human and animal PV infections and model systems .......................................................... 160
5.7 E4 shows different intracellular expression patterns in lesions caused by different PVs ........................................................... 162
5.8 The viral life cycle in mouse xenograft systems ....................................................................................................................... 167
5.9 Discussion ................................................................................................................................................................................. 170

CHAPTER 6 - EXPRESSION OF CELL CYCLE MARKERS IN INFECTED TISSUES IN RELATIONS TO THE ONSET OF THE VIRAL LATE EVENTS ...................................... 179
6.1 Introduction .................................................................................................................................................................................. 180
6.2 Staining protocol and antibodies .............................................................................................................................................. 183
6.3 Expression of PCNA and the onset of the viral late events ........................................................................................................ 183
6.4 Expression of G1/S transition cyclins and the onset of the viral late events ......................................................................... 189
6.5 Expression of cyclin B (G2/M transition) and onset of the viral late events ............................................................................. 192
6.6 Discussion .................................................................................................................................................................................... 196

CHAPTER 7 - CONSTRUCTION OF HPV-16 AND CRPV E4 C-TERMINAL TRUNCATION PROTEIN MUTANTS ................................................................................................................. 200
7.1 Introduction .................................................................................................................................................................................. 201
7.2 Design of the HPV-16 and CRPV E4 mutants ............................................................................................................................ 203
7.3 Construction of recombinant plasmids containing the E4 ORF ............................................................................................... 206
7.4 Kunkel site-directed mutagenesis .............................................................................................................................................. 207
7.5 QuikChange site-directed mutagenesis ................................................................................................................................... 208
7.6 Reconstruction of the HPV-16 E4 mutant viral genome ............................................................................................................. 209
7.7 Reconstruction of the CRPV E4 mutant viral genome ................................................................................................................ 215
7.8 Discussion .................................................................................................................................................................................... 220

CHAPTER 8 - ANALYSIS OF CRPV E4 MUTANTS IN NEW ZEALAND WHITE (NZW) AND COTTONTAIL (CT) RABBITS .................................................................................. 223
8.1 Introduction .................................................................................................................................................................................. 224
8.2 E4 mutant experiment outline .................................................................................................................................................... 226
8.3 Revised protocol for E4 mutant study in CT rabbits .................................................................................................................... 228
8.4 Loss of E4 did not affect the virus ability to induce papilloma development ........................................................................... 229
8.5 E4 knockout did not inhibit the development of rabbit papillomas ........................................................................................... 235
8.6 Papilloma morphology was not affected by the E4 knockout ................................................................................................... 240
8.7 The E4 knockout did not disrupt the early stage of the virus life cycle ...................................................................................... 242
8.8 The E4 knockout abolished the viral late events ......................................................................................................................... 245
8.9 Discussion .................................................................................................................................................................................... 249

CHAPTER 9 - FINAL DISCUSSION ....................................................................................................................................................... 253
APPENDIX 1 - SEQUENCES .................................................................................................................................................................. 265
APPENDIX 2 - PLASMID MAPS ............................................................................................................................................................ 270
APPENDIX 3 - SDS-PAGE ................................................................................................................................................................. 277
REFERENCES ..................................................................................................................................................................................... 278
PUBLICATION .................................................................................................................................................................................... 304
ABBREVIATIONS

A  adenosine
+ac  acetylated
-ac  non-acetylated
amp  ampicillin
AP  alkaline phosphatase
APP  appendix
APS  ammonium persulfate
ATP  adenosine triphosphate
BLAST  basic local alignment sequence
bp  basepairs
BPV  bovine papillomavirus
BSA  bovine serum albumin
C  cysteine (amino acid)
C  cytidine (nucleoside)
°C  degrees centigrade
CAP  chloramphenicol
CD  cluster of differentiation
cdk  cyclin-dependent kinase
CDKI  cyclin-dependent kinase inhibitor
CE  cornified envelope
CIN  cervical intraepithelial neoplasia
CKII  casein kinase II
cm  centimeter
COPV  canine oral papillomavirus
CR1(2)  conserved regions 1 (or 2)
CRPV  cottontail rabbit papillomavirus
CT  Cottontail
D  aspartic acid
DAB  3',3'-diaminobenzidine tetrahydrochloride
DAPI  4',6-Diamidino-2-phenylindole dihydrochloride
dATP  deoxyadenosine 5'-triphosphate
DBD  DNA-binding domain
dCTP  deoxycytidine 5'-triphosphate
dGTP  deoxyguanosine 5'-triphosphate
DIG  digoxigenin
DMF  dimethylformamide
DNA  deoxyribonucleic acid
dNTP  deoxynucleoside triphosphate
DPV  deer papillomavirus
DTNB  5,5-Dithio-bis-2-nitrobenzoic acid
dTTP  deoxythymidine 5'-triphosphate
E  early open reading frame
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>glutamic acid (amino acid)</td>
</tr>
<tr>
<td>E6AP</td>
<td>E6-associated protein</td>
</tr>
<tr>
<td>ECL</td>
<td>enhanced chemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Diaminoethanetetra-acetic acid disodium salt</td>
</tr>
<tr>
<td>EEPV</td>
<td>European elk papillomavirus</td>
</tr>
<tr>
<td>e.g.</td>
<td>example</td>
</tr>
<tr>
<td>EI</td>
<td>experimentally infected</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immuno sorbent assay</td>
</tr>
<tr>
<td>et al.</td>
<td>et alii (and others)</td>
</tr>
<tr>
<td>EthBr</td>
<td>ethidium bromide</td>
</tr>
<tr>
<td>EV</td>
<td>epidermodysplasia verruciformis</td>
</tr>
<tr>
<td>F</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>FB</td>
<td>final-bleed</td>
</tr>
<tr>
<td>FISH</td>
<td>fluorescent <em>in situ</em> hybridisation</td>
</tr>
<tr>
<td>FR</td>
<td>fast-red</td>
</tr>
<tr>
<td>x g</td>
<td>units for the centrifugal force in a rotator</td>
</tr>
<tr>
<td>G</td>
<td>glycine (amino acid)</td>
</tr>
<tr>
<td>G</td>
<td>guanosine (nucleoside)</td>
</tr>
<tr>
<td>G1</td>
<td>gap 1 in the cell cycle; a resting phase before cells enter into the replication (S) phase</td>
</tr>
<tr>
<td>G2</td>
<td>gap 2 in the cell cycle; a resting phase before cells enters into mitosis</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione sephrose transferase</td>
</tr>
<tr>
<td>H</td>
<td>histidine</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>HLA</td>
<td>human leukocyte antigen</td>
</tr>
<tr>
<td>HPV</td>
<td>human papillomavirus</td>
</tr>
<tr>
<td>HRP</td>
<td>horse-radish peroxidase</td>
</tr>
<tr>
<td>hrs</td>
<td>hours</td>
</tr>
<tr>
<td>I</td>
<td>isoleucine</td>
</tr>
<tr>
<td>ID</td>
<td>identity</td>
</tr>
<tr>
<td>i.e.</td>
<td>id est (that is)</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-1-thio-β-D-galactopyranoside</td>
</tr>
<tr>
<td>ISH</td>
<td><em>in situ</em> hybridisation</td>
</tr>
<tr>
<td>K</td>
<td>lysine</td>
</tr>
<tr>
<td>kbp</td>
<td>kilo base pair</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo daltons</td>
</tr>
<tr>
<td>KLH</td>
<td>keyhole limpet hemocyanin</td>
</tr>
<tr>
<td>L</td>
<td>late open reading frame</td>
</tr>
<tr>
<td>L</td>
<td>leucine (amino acid)</td>
</tr>
<tr>
<td>L</td>
<td>litres (measuring unit)</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani medium</td>
</tr>
<tr>
<td>LCR</td>
<td>long control region</td>
</tr>
<tr>
<td>μm</td>
<td>micron; micrometer</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>M</td>
<td>methionine (amino acid)</td>
</tr>
<tr>
<td>M</td>
<td>mitosis; a phase in the cell cycle</td>
</tr>
<tr>
<td>mA</td>
<td>milli amperes</td>
</tr>
</tbody>
</table>
MAPK  mitogen-activated protein kinase
MBP  maltose binding protein
MBS  \( m \)-maleimidedobenzoyl-N-hydroxysuccinimide ester
mg  milligrams
\( \text{MgCl}_2 \)  magnesium chloride
mins  minutes
mL  millilitres
mRNA  messenger ribonucleic acid
N  asparagine
NaCl  sodium chloride
NaOH  sodium hydroxide
NCR  non-coding region
ND10  nuclear domains 10
NES  nuclear export signal
ng  nanogram
NGS  normal goat serum
Ni  natural infection
NET-N  a solution that comprises sodium chloride, EDTA, Tris salt, and NP40 detergent.
NLS  nuclear localisation signal
NoLS  nucleolar localisation signal
NP40  nonidet P-40 detergent
NRE  negative regulatory element
NZW  New Zealand White
OD  optical density
ORC  origin recognition complex
ORF  open reading frame
ori  viral origin of replication
P  proline
PAE  papillomavirus early poly-adenylation site
PAGE  polyacrylamide gel electrophoresis
PAL  papillomavirus late poly-adenylation site
paps  papillomas
PBS  phosphate buffered saline
PCNA  proliferating cell nuclear antigen
PCR  polymerase chain reaction
PDZ  to describe an intracellular protein module (PDZ domain) which is identified by the presence of a conserved Gly-Leu-Gly-Phe sequence. Name derives from mammalian postsynaptic density protein 95kD (PDS-95), Drosophila disc large tumor suppressor (DlgA), and the mammalian tight junction protein Z01.
PEG  polyethylene glycol
p.f.u.  plaque forming units
pH  the symbol relating the hydrogen ion concentration or activity of a solution to that of a given standard solution
PI  propidium iodide
Pim  pre-Immune
PKA  protein kinase A
PKC  protein kinase C
pRB  retinoblastoma tumour suppressor protein
psi  pounds force of pressure per square inch
PV   papillomavirus
Q    glutamine
r    rat
R    arginine
Rb   rabbit
RFC  replication factor C
RLF  replication licensing factor
RM   protein rainbow marker
ROPV rabbit oral papillomavirus
RPA  replication protein A
rpm  rotations per minute
S    serine (amino acid)
S    the replication phase of the cell cycle
SCID severe combined immuno-deficiency
SDS  sodium dodecyl sulphate
SG   sytox green
SIL  squamous intraepithelial lesions
SSC  sodium citrate, sodium chloride buffer
SSCP single-strand conformation polymorphism
T    threonine
T    thymidine
Taq  Thermobius aquaticus
TB   test-bleed
TBE  tris borate EDTA buffer
TBS  tris buffered saline
TE   tris EDTA buffer
TEMED N,N',N'-Tetramethylethylenediamine
TM   tris magnesium chloride buffer
TMB  tetramethylbenzidine dihydrochloride
TSA  tyramide signal amplification
Urd  uridine
URR  upstream regulatory region
uv   ultra-violet
V    valine
V    volts
VLP  virus-like particle
v/v  volume by volume ratio
W    tryptophan
W    watts
wt   wildtype
w/v  weight by volume ratio
Y    valine
* termination (amino acid sequence)
FIGURES AND TABLES

Chapter 1
Figure 1.1. The PV phylogenetic tree.
Figure 1.2. Circular viral genome map of cottontail rabbit papillomavirus (CRPV).
Figure 1.3. The PV life cycle.
Figure 1.4. Experimental tools used in PV research and infections caused by animal PVs.

Chapter 3
Table 3.1. Description of the E4 ORF (nucleotide positions), E1^E4 message (nucleotide positions) and protein (number of amino acids), and the virus classification.
Figure 3.2. E1^E4 protein sequences of BPV-1, CRPV, ROPV, COPV and HPV-11.
Figure 3.3. Alignments analysis of the E1^E4 protein sequences of the supergroup C PVs.
Figure 3.4. Alignment analysis of the E1^E4 protein sequences of the supergroup E PVs.
Figure 3.5. Alignments analysis of the E1^E4 protein sequences of the rabbit PVs.
Figure 3.6. Alignments analysis of the E1^E4 protein sequences of the Group A10 PVs.
Figure 3.7. Proline-rich region consensus in the E4 proteins of rabbit and Group A10 PVs.
Figure 3.8. Putative NLSs and NoLSs in E4.
Figure 3.9. Alpha-helical wheel representation of the putative leucine-rich NES motifs in the E1^E4 proteins of CRPV (a; positions 77-89), CRPVb (b & c; positions 77-89 & 103-109), ROPV (d; positions 99-119), COPV (e; positions 103-112), HPV-11 (f; positions 7-17) and HPV-1a (g; positions 89-96).

Chapter 4
Figure 4.1 Specific PCR primers for the synthesis of E1^E4 sequence fragments.
Figure 4.2 Preparation of the E1^E4 fusion proteins used for the generation of E1^E4 specific antibodies.
Table 4.3 Percentage of peptide conjugation as determined by the cysteine assay.
Table 4.4 Abbreviations used to denote the immunogens, animals and antisera during the generation of the E1^E4 specific antibodies.
Figure 4.5 E4 specific binding on western blots.
Figure 4.6 ELISA binding curves of antisera against GST or MBP fusion E1^E4 proteins.
Figure 4.7 ELISA binding curves of antisera against KLH-peptide immunogens.

Chapter 5
Table 5.1 Types and sources of infected tissue samples used for immuno-staining.
Figure 5.2 General distribution patterns of the viral E4 protein in infected tissue sections as detected by specific E4 antibodies.
Figure 5.3 Viral DNA amplification and E4 expression correlate exactly in infected tissues.
Figure 5.4 Rare events during the late stage of the viral life cycle.
Figure 5.5 E4 expression proceeds the expression of L1 in human and animal PV infections.
Figure 5.6 The different intracellular localisation patterns of E4 in (a) COPV, (b) ROPV and (c) CRPV infected lesions.
Figure 5.7 The PV life cycle in murine xenograft model systems.

Chapter 6
Table 6.1 Cell cycle protein antibodies and their species cross-reactivity on formalin-fixed paraffin-embedded tissue sections.
Figure 6.2 A population of PCNA positive cells in the differentiated layers of the epithelium coincides with the onset of viral late events.
Figure 6.3 Detection of cyclin E expression did not generally correlate with the onset of the viral late events, as represented by E4 expression or viral DNA amplification.
Cyclin B expression patterns showed variable degrees of correlation with the onset of the viral late events (E4 expression and viral DNA amplification).

Descriptions of the (a) HPV-16 and (b) CRPV E4 mutants.

Restriction enzyme sites chosen for the cloning of the HPV-16 and CRPV E4 fragments into pET-23b(+).

Mutagenesis statistics: Comparing the efficiency of the Kunkel and QuikChange site-directed mutagenesis methods.

Nucleotide variations found in the W12 HPV-16 genome revealed by sequencing of the pTZhW12 clones between positions 3094-4885 (coding strand).

Preparations of the re-circularised W12 HPV-16 wt and E4 mutant viral genome DNA.

Spontaneous nucleotide base changes in wt and E4 mutant pLACRPV plasmids revealed by bi-directional automated sequencing and restriction enzyme digests.

Preparation of the re-circularised CRPV wt and E4 mutants viral genome DNA.

CT rabbit papillomas collected at the end of the experiment.

Frequency data of papilloma formation twelve weeks post-inoculation of NZW rabbits collected by Janet and co-workers.

Papilloma development following DNA inoculation on NZW rabbits, Rb#1, Rb#2, Rb#3 and Rb#4. DNA "shots" were delivered onto the shaven areas on the right, left and rear of the NZW rabbit hosts.

Time course of papilloma development in the NZW rabbits.

Histology of the wt and E4 mutant (CRst9/0) DNA induced CT and NZW rabbit papillomas.

Immuno-detection of surrogate markers of E7 expression (PCNA and cycE) and viral DNA amplification in CRPV wt and E4 knockout (CRst9/0) DNA induced CT and NZW rabbit papillomas.

Detection of the viral major late events markers in the wt and E4 knockout (CRst9/0) DNA induced CT and NZW rabbit papillomas.

Selected list of proteins obtained from the basic local alignment sequence (BLAST2.0) entry of (a) BPV-1, (b) CRPV, (c) ROPV, (d) COPV, (e) HPV-11 E1^E4 protein sequences.

Nucleotide changes introduced into the HPV-16 E4 mutant genomes [E1^E4 ORF: (865 - 880)^(3357 - 3619)].

Nucleotide changes introduced into the CRPV E4 mutant genomes [E1^E4 ORF: (1362-1371 )A(3714-4015)].


Components of SDS-polyacrylamide gel solutions.
CHAPTER 1 - GENERAL INTRODUCTION

1.1 Papillomaviruses
1.1.1 Classification
1.1.2 Viral genome
1.1.3 Viral transcription

1.2 Viral early proteins
1.2.1 Viral oncoproteins - E6, E7 and E5
1.2.2 Viral replication and regulation proteins - E1 and E2
1.2.3 The E4 protein

1.3 Viral late proteins

1.4 The epithelium

1.5 PV-associated diseases
1.5.1 Cutaneous HPV diseases
1.5.2 Mucosal HPV diseases

1.6 The PV life cycle
1.6.1 Productive infection
1.6.2 Carcinoma

1.7 Immunity to PV infections
1.7.1 Cell-mediated immunity
1.7.2 Humoral immunity
1.7.3 Vaccine development

1.8 Experimental models for PV infection
1.8.1 Differentiating organo-raft cultures
1.8.2 CRPV-rabbit model
1.8.3 BPV-bovine model
1.8.4 COPV-canine model
1.8.5 ROPV-rabbit model
1.8.6 Murine models
1.8.7 Particle-mediated gene delivery technology (Genegun)

1.9 Project Objectives
Chapter 1 – General Introduction

1.1 Papillomaviruses

1.1.1 Classification

The first papillomavirus (PV) was isolated and identified from infected wild cottontail (CT) rabbits in 1933 (Shope, 1933). Since then, more than a hundred different types of human and animal PVs have been reported (de Villers, 1997). Even so, it is believed that a large number of PV types still remain undiscovered (Antonsson et al., 2000). PVs belong to the family of small, non-enveloped, double-stranded DNA viruses known as papovaviruses (reviewed in Lowy & Howley, 2001). They have been implicated in the development of numerous different epithelial tumours on the cutaneous and mucosal epithelia of human and animal hosts. These viruses are highly host and tissue specific, and are defined by their tissue tropism (either cutaneous or mucosal tropic viruses) and their pathology. For instance, the mucosal-tropic human papillomaviruses (HPVs) are divided into the low-risk or high-risk types, based on their association with benign to malignant conversion. PVs can also be classified on the basis of DNA sequence homology. This method of classification analyses short conserved regions of the E6, E7 and L1 genes. A new PV type is designated when a less than 90% similarity is found in the combined E6/E7/L1 segment alignment with existing PV types (de Villers, 1994). New PV types can then be placed in an appropriate lineage in the PV phylogeny, which reflects the evolutionary relationships between different viruses (Wiley, 1978 & 1981). Different PV phylogenies have been constructed based on the analysis of partial E6, E7 and L1 gene sequences (Chan et al., 1995). However, similar phylogeny structures consisting of five major supergroups (or clades), named Supergroup A to E, were produced either from individual or combined (E6/E7/L1) gene sequence analyses (Chan et al., 1995). A supergroup consists of all species that are derived from shared ancestral characters and their most recent common ancestor (Wiley, 1981). Within some supergroups, viruses are sub-divided into groups based on the calculated cladistic distances between the different PV types (Chan et al., 1995; Figure 1.1).

In general, most HPVs fall into one of two main phylogenetic supergroups. Supergroup A consists
of mucosal tropic HPV types, and Supergroup B consists of cutaneous tropic, and epidermodysplasia verruciformis (EV)-associated HPV types.

Animal PVs can be found in all five supergroups of the PV phylogeny. The primate specific PVs (ape or monkey) are the only animal PVs found in Supergroups A and B, while the non-primate animal PVs are located in other supergroups. Supergroup C consists of the ungulate fibropapillomaviruses such as bovine PV types 1 and 2 (BPV-1, BPV-2), deer PV (DPV) and european elk PV (EEPV). The epithelial-tropic bovine PVs such as BPV-3, BPV-4, and BPV-6 make up Supergroup D. Interestingly, Supergroup E consists of a unique combination of human (HPV-1, HPV-63 and HPV-41) and animal PVs of cutaneous and mucosal tissue tropism. In this supergroup, the closely related HPV-1 and HPV-63 virus types form Group E1. Another closely related group of viruses in group E are the two cottontail rabbit PV types (CRPV and CRPVb) and the rabbit oral PV (ROPV).
Figure 1.1. The PV phylogenetic tree.
This tree is based on the analysis of the 291bp segment of the L1 gene, and has been collated from the data presented by Chan et al. (1995), Myers et al. (1997), and Bernard & Chan (1997). The phylogenetic tree consists of five major supergroups (A-E). In some cases, PVs were further classified into smaller groups (e.g. A1, B1, and E1) within the supergroups. Interestingly, this method of classification showed that viruses that share similar tissue and cell type specificities evolved from common ancestors.
Figure 1.2. Circular viral genome map of cottontail rabbit papillomavirus (CRPV). The organisation of the PV genome is highly conserved. All viral genes are encoded from one coding strand which can be divided into the early (E) or late (L) gene regions. The transcription of the viral ORFs are regulated by elements located in the long control region (LCR). Viral transcripts can be expressed from multiple promoter sites, for example, four promoter sites (P₁, P₂, P₃ and P₄) have been identified in the cottontail rabbit PV (CRPV) genome. The early poly-adenylation site (PAE) is used by viral transcripts expressed during the early stage of the virus life cycle, while the late poly-adenylation site (PAL) is utilised by the viral transcripts expressed from the differentiation-dependent promoter (denoted by P₄ in the CRPV genome).
1.1.2. Viral genome

The genetic organisation of the PV genome is fairly conserved between different virus types. The viral genome consists of a double-stranded closed-circular DNA of approximately 8 kilobase pairs (kbp), and contains an average of nine open reading frames (ORFs) (reviewed in Lowy & Howley, 2001; Figure 1.2). All viral gene transcripts are expressed from one DNA strand. The viral genome is divided into the early (E) or late (L) gene regions based on their location in the genome (reviewed in Howley & Lowy, 2001). The early genes are designated E1 to E8 and are transcribed during the early stage of the virus life cycle. The late genes encode the major and minor viral capsid proteins, L1 and L2 respectively (Komly et al., 1986; Doorbar & Gallimore, 1987). An upstream regulatory region (URR), also known as the long control region (LCR), is located after the L1 ORF and before the start of the E6 ORF. The URRs differ in their sizes between different PVs, and have been found to exhibit a degree of tissue or cell-type specificity (Sailaja et al., 1999). This region contains a number of enhancer and promoter elements (Ahola et al., 1987; Cripe et al., 1987; Spalholz et al., 1987; Thierry et al., 1987), transcription factor binding motifs (Spalholz et al., 1988; Kanaya et al., 1997; Dong & Pfister, 1999), and the viral origin of replication (Waldeck et al., 1984; Stenlund et al. 1987). Thus the URR is the main region involved in the regulation of viral gene expression and viral DNA replication (Sousa et al., 1990; Chow & Broker, 1994).

1.1.3. Viral transcription

The PV transcription pattern is complex as it is regulated by the activity of multiple viral promoters, the viral transcription regulator protein E2, alternate and multiple splicing mechanisms, and keratinocyte differentiation (Spalholz et al., 1985; Barksdale & Baker, 1993 & 1995; Hummel et al., 1992 & 1995; Ozbun & Meyers, 1998). BPV-1 transcription is the most extensively studied gene expression model. Seven different BPV-1 transcription promoters and more than twenty different mRNA species have been identified in transformed cells, as well as in infected tissues (reviewed in Howley & Lowy, 2001). Subsequent gene expression
studies in HPV and CRPV infected cell lines and tissues have revealed similar characteristics, including the presence of multiple viral promoters in the URR and early gene region, and the generation of a range of different transcript species during the virus life cycle (Nasseri & Wettstein, 1984; Smotkin & Wettstein, 1986; Chow et al., 1987; Sherman et al., 1992; Ozbun & Meyers, 1997). PV gene expression is governed by two sets of promoter activities, namely the early and differentiation-dependent promoters. Viral transcripts produced during the early and late stages of the virus life cycle terminate at specific polyadenylation sites found at the end of the early gene region and at the end of the late genes respectively (Sousa et al., 1990). All PV genomes contain multiple early promoters within the URR or early gene regions. An early promoter in front of the E6 ORF is a common feature in all PVs (reviewed in Fuchs & Pfister, 1994). During the activation of the early promoters, the early genes including E6, E7, E1 and E2 are expressed (reviewed in Howley & Lowy, 2001). Numerous different spliced mRNAs species are generated in variable abundance as a result of different promoter activity and alternate splicing mechanisms (Hummel et al., 1992). A switch from the early promoters to the differentiation-dependent promoter(s) occurs during keratinocyte differentiation which results in a change of the main mRNA species expressed in the cells (Wettstein et al., 1987; Baker & Howley, 1987; Hummel et al., 1992; Ozbun & Meyers, 1997). Alternate splicing is the main mechanism used to generate different species of viral late mRNAs. Transcript mapping in different PV infected materials identified the main viral late transcripts to be the E1^E4, E5, L1 and L2 ORFs (Danos et al., 1985; Phelps et al., 1985; Palermo-Dilts et al., 1990; Rohlf et al., 1991; Hummel et al., 1995). The E1^E4 ORF is a spliced mRNA species which contains a short segment of the N-terminal E1 ORF joined to the spliced E4 ORF, generated from the conserved viral splice donor/acceptor sequences in the E1 and E4 ORFs.
1.2 Viral Early (E) Proteins

All PVs encode three main viral oncoproteins E5, E6 and E7, and three viral regulatory proteins E1, E2 and E4.

1.2.1 Viral oncoproteins - E6, E7 and E5

Although E7 is the major PV oncoprotein, the expression of high-risk HPVs E6 and E7 together is necessary to efficiently immortalise and transform primary keratinocytes or fibroblasts in vitro (Munger et al., 1989; Watanabe et al., 1989), as well as to cause epidermal cancer in transgenic mice (Lambert et al., 1993). Viral DNA integration into the host genome that disrupts the E2 ORF, results in the uncontrolled expression of E6 and E7, and is thought to be a contributing factor that leads to oncogenic conversion of HPV infection (Jeon & Lambert, 1995; Jeon et al., 1995). The oncogenic potential of E6 and E7 is more potent in the high-risk (e.g. HPV-16, HPV-18) than the low-risk HPVs (e.g. HPV-6, HPV-11), and is thought to be related to the association of high-risk HPVs with cancer (Schelgel et al., 1988; Storey et al., 1988; Nakagawa et al. 1996). E6 and E7 disrupt the host cell cycle by binding to different proteins involved in the cell cycle machinery. Both oncoproteins are small proteins that contain multiple zinc-fingers (two in E6 and one in E7), which are typical of nucleic-acid binding proteins (Barbosa et al., 1989; Grossman & Laimins, 1989; Grossman et al., 1989).

The transformation and immortalisation of primary keratinocytes by high-risk HPV E6 and E7 is dependent on the expression of a full-length E6 protein (Sedman et al., 1991). Truncated forms of E6 have been detected in some high-risk HPV infected cells, and are thought to be expressed as a result of alternate splicing (Munger et al., 1989; Pim et al., 1997). These shorter E6 proteins were shown to inhibit the transforming activities of full-length high-risk HPV E6 proteins by acting as competitors for the E6 binding proteins (Munger et al., 1989; Pim et al., 1997). The transformation ability of E6 is limited to the high-risk HPVs, while low-risk HPV E6 proteins showed little or no transformation activity (Sedman et al., 1991). This is the case even
though sequence analysis of the high-risk and low-risk HPV E6 proteins did not reveal any major dissimilarities between them (McGlennen, 2000). The main E6-induced cellular effect in high-risk HPV infected cells is the rapid degradation of the tumour suppressor protein p53 (Scheffer et al, 1990). High-risk HPV E6 can associate with p53 via the ubiquitin protein ligase (E6AP), which promotes the ubiquitination of p53 and targets the protein for ubiquitin-dependent proteolysis in vitro (Huibregtse et al., 1991 & 1993; Scheffer et al, 1993). Low-risk HPV E6 does not bind p53 and cause its degradation (Scheffner et al., 1990). Expression of high-risk HPV E6 in cells also reduces the half-life of p53 and represses p53-mediated cellular responses to DNA damage (i.e. G1 and G2 growth arrest or apoptosis) (Hubbert et al., 1992; Lechner et al., 1992; Kessis et al., 1993; Innocente et al., 1999). In addition, high-risk HPV E6 has been shown to be able to induce p53-independent transformation activities in cancer cells using E6 mutants that can no longer bind p53 (Spitkovsky et al., 1996), and in p53-null background transgenic mice (Song et al., 1999). Other examples of p53-independent E6 activities include the activation of telomerase by HPV-16 E6 (Klingelhutz et al., 1996) and inhibition of Bak-induced apoptosis by HPV-18 E6 (Thomas & Banks, 1998). Recently, high-risk HPV E6 has also been found to bind numerous PDZ domain-containing cellular proteins causing them to be degraded via the ubiquitin pathway (Gardiol et al., 1999; Nakagawa et al., 2000). These cellular targets include the hD1g (human homologue of the Drosophila discs large tumour suppressor protein (Kiyono et al., 1997; Lee et al., 1997), MUPP1 (the multi-PDZ-containing protein; Lee et al., 2000) and hScrib (human homologue of Drosophila scribble; Nakagawa et al., 2000), which have been shown to be involved in negatively regulating cellular proliferation.

High-risk HPV E7 can transform cells in the presence of an activated ras oncogene, and abrogate cell cycle checkpoints at the G1/S phase boundary. Expression of HPV-18 E7 alone has been shown to be necessary and sufficient to induce S-phase entry in differentiated primary keratinocytes in raft cultures (Cheng et al., 1995). E7 can bind to the retinoblastoma tumour suppressor protein (pRB), and other related "pocket" proteins p107 and p130 (Dyson et al.,
1989), and promotes degradation of pRB (Dyson et al., 1989; Boyer et al., 1996). During the G1/S phase transition in the cell cycle, hypophosphorylated pRB normally binds to E2F and the protein complex inhibits cell cycle progression by acting as a transcription repressor of S phase genes (Harbour & Dean, 2000). E7 binding to hypophosphorylated pRB releases the E2F transcription factor, and leads to the expression of a number of genes necessary for S-phase entry and progression, including DNA polymerase α, thymidine kinase, PCNA, and cyclins A and E (Herwig & Strauss, 1997). The pRB binding, degradation and transformation properties of E7 are primarily mediated by the E7 conserved regions 1 and 2 (CR1, CR2) at the amino (N)-terminus of the protein (Dyson et al., 1989). Low-risk HPV E7 proteins bind pRB 10-fold weaker than the high-risk HPVs and are less efficient in transforming cells in vitro (Gage et al., 1990, Barbosa et al., 1991, Heck et al., 1992). E7-pRB binding affinity is dependent on the presence of a single conserved amino acid in the CR2 region motif (LXGXE). This amino acid is an aspartic acid residue in low-risk HPVs, and a glycine residue in high-risk HPVs (Sang & Barbosa, 1992).

Besides pRB binding, the carboxy (C)-terminus of E7 is also required for the disruption of the pRB/E2F pathway (Huang et al., 1993). This region contains a zinc-finger (Barbosa et al., 1989) and the E7 dimerisation domain, and is involved in E7 stability in cells (Edmonds & Vousden, 1989; Phelps et al., 1992). E7 also disrupts other cell cycle regulatory pathways by binding to the cyclin-dependent kinase inhibitors (CDKI) p21 and p27 (Zerfass-Thome et al., 1996; Funk et al., 1997; Jones et al., 1997). p21 and p27 are broad acting inhibitors that can affect the activities of cyclins D, E and A-dependent kinases (reviewed in Sherr & Roberts, 1999). E7 has been shown to interact with p27 and inhibit its activity (Zerfass-Thome et al., 1996). E7 association with p21 prevents the inhibition of cyclin-dependent kinase activity and PCNA-dependent DNA replication (Funk et al., 1997; Jones et al., 1997).

E5 is a transmembrane protein which has been shown to have the potential to transform fibroblasts and keratinocytes in vitro (Schiller et al., 1988; Chen & Mounts, 1990; Leptak et al., 1991). Although the E5 gene is conserved in both fibropapillomaviruses and PVs, the
transformation activity of its gene product is thought to be more important for the proliferation of dermal fibroblasts (Kulke & DiMaio, 1991). The role of E5 in epithelial infections is less well-understood, but it does not appear to be required for carcinogenesis (Brandsma et al., 1992; McGlennen, 2000). The protein is localised primarily in the membranes of the Golgi apparatus, endosomes and to some extent the cell membrane (Burkhardt et al., 1989; Burnett et al., 1992; Conrad et al., 1993). E5-induced cell transformation is mediated by its ability to associate with and activate tyrosine-kinase-associated growth factor receptors, such as those for platelet-derived growth factor, colony stimulating factor, and epidermal-growth factor (Martin et al., 1989; Petti & DiMaio, 1992). E5-receptor binding activates the receptor pathways in a ligand-independent manner (Straight et al., 1993; Petti & DiMaio, 1994), leading to the constitutive induction of a host of signal transduction effectors including the activation of endogenous ras, induction of c-fos, and NF-1 (Chen et al., 1996; Ghai et al., 1996). E5 can also inhibit the acidification of endosomes by binding to the 16kDa subunit of the vacuolar ATPase thereby abrogating the down-regulation of the growth factor receptors (Conrad et al., 1993; Straight et al., 1995). Recently, E5 has also been shown to downregulate surface MHC class I molecule expression by inducing MHC class I heavy chain degradation in the Golgi apparatus (Ashrafi et al., 2002).

1.2.2 Viral replication and regulation proteins - E1 and E2

E1 and E2 are essential for PV DNA replication (Ustav & Stenlund, 1991). E1 is directly involved in plasmid replication, while E2 has an auxiliary role. Besides E1 and E2, additional viral and cellular factors are required for viral DNA replication. These include the viral origin of replication (ori), replication protein A (RPA), replication factor C (RFC), DNA polymerase-α primase, DNA polymerase δ, proliferating cell nuclear antigen (PCNA), topoisomerases I and II (Waldeck et al., 1984; Park et al., 1994, Melendy et al., 1995).
The E1 protein is a DNA helicase with DNA-dependent ATPase activity (Seo et al., 1993; Yang et al., 1993; Jenkins et al., 1996). It is required for the initiation and elongation of viral DNA synthesis (Liu et al., 1995). The E1 protein is highly conserved among the PV types (reviewed in Howley & Lowy, 2001). The N-terminal domain of E1 contains the nuclear localisation signal (NLS) and DNA-binding domain (DBD), whilst the ATPase and DNA helicase activities are found in the C-terminal part of the protein (Lentz et al., 1993; Thorner et al., 1993; Sarafi & McBride, 1995; Leng et al., 1997). In the presence of E2, E1 binds to the ori and initiates DNA replication. Assembly of the E1-E2-ori initiation precursor involves several specific interactions between E1 and E2, and between E1 dimers in a highly ordered fashion (Chen & Stenlund, 1998). The active E1 DNA helicase complex is formed when the E2 proteins are displaced from the precursor and then replaced with two E1 proteins to form a hexamer complex (reviewed in Howley & Lowy, 2001). In addition, E1 has also been shown to bind other cellular proteins such as the S-phase specific cyclin E/cdk2 (cyclin-dependent kinase) complex (Cueille et al., 1998), and modification protein Ubc9 (Rangasamy & Wilson, 2000). Interaction between E1 and the cyclin E/cdk2 complex determines the efficiency of PV replication (Cueille et al., 1998), and the association and SUMO-1 modification of E1 by Ubc9 is required for efficient origin-dependent replication and for the intranuclear accumulation of E1 (Rangasamy et al., 2000).

E2 does not have any enzymatic function in viral DNA replication, and its primary role in replication is to recruit E1 to the ori (Mohr et al., 1990; Lusky et al., 1993 & 1994). Besides this role, E2 is also involved in origin recognition (Yang et al., 1991), early gene transcription regulation (McBride et al., 1991), the recruitment and binding of cellular replication factors (Sedman & Stenlund, 1995), and plasmid maintenance (Skidopoulos & McBride, 1998). The structure of E2 is conserved among all PVs. The N-terminal part of the protein contains the trans-activation domain, and the C-terminal region the DNA-binding and dimerisation domains (Giri & Yaniv, 1988; McBride et al., 1988). Three different species of the E2 protein have been
identified. E2 can be expressed either as a full-length protein, or as two shorter N-terminal truncated forms (Hubbert et al., 1989; Chiang et al., 1991). E2 dimers regulate viral promoter activity through E2-responsive elements located within the viral genome (Spalholz et al., 1987). The trans-activation or repression activity of the full length E2 protein depends on the location of the E2-binding sites within the enhancer or promoter region (Androphy et al., 1987; Stenlund & Botchan, 1990; Thierry & Howley, 1991). Repression of viral transcriptional activity by the two truncated forms of E2 is mediated by competitive binding to the E2-binding sites and by forming inactive heterodimers with the full-length E2 protein (Lambert et al., 1987; Barsoum et al., 1992).

1.2.3 The E4 protein

The E4 gene product is expressed from a spliced E1^E4 mRNA from the differentiation-dependent promoter (Doorbar et al., 1988; Pray & Laimins, 1995). Although the E4 gene is found in the early gene region of the viral genome, the E1^E4 ORF has been found in numerous late transcript species in BPV and HPV infected tissues, and has been described as a major late transcript in HPV-1 and HPV-11 infections (Chow et al., 1987; Nasseri et al., 1987; Palermo-Dilts et al., 1990). E1^E4 is a small phosphoprotein (Grand et al., 1989; Bryan et al., 2000) that is expressed during the late stage of the virus life cycle. For this reason, the E1^E4 protein is often considered as a late protein and has been implicated in roles during the viral genome amplification and virus production stages of the virus life cycle (Doorbar et al., 1986 & 1997; Breitburd et al., 1987).

The E4 amino acid sequence is highly divergent between different PV types (see Chapter 3). However, conserved characteristics are apparent in the E4 protein of different PV types. All E1^E4 proteins have proline-rich domains near the N-terminus, and at least one charged region consisting of a stretch of basic and acidic amino acid residues (Doorbar and Myers et al., 1996; Chapter 3). Multimerisation is a characteristic of numerous HPV types,
including HPV-1, HPV-11, and HPV-31b (Doorbar et al., 1988; Tomita et al., 1991; Pray & Laimins, 1995). HPV-1 E1^E4 multimerisation is found to be mediated by the hydrophobic residues in the C-terminus of the protein (Doorbar et al., 1996; Ashmole et al., 1998). Furthermore, the C-terminal regions of HPV-1 and HPV-16 have also been shown to be involved in the binding of E1^E4 to the cytokeratin network (Roberts et al., 1997; Ashmole et al., 1998). Recently, the C-terminus of HPV-16 E1^E4 was shown to bind a putative RNA helicase (Doorbar et al., 2000). Phosphorylation is also a feature of E1^E4 proteins. Numerous putative phosphorylation sites have been identified in the E1^E4 protein sequence of different PV types (Breitburd et al., 1987; Grand et al., 1989; Bryan et al., 2000; Chapter 3). Moreover, HPV-1 E1^E4 was shown to be progressively phosphorylated as cells differentiate in the infected tissue (Grand et al., 1989).

PVs that exhibit the same tissue tropism also share some conserved motifs in their E1^E4 proteins. The mucosal HPVs have a leucine cluster motif (LLXLL) at the N-terminus of E1^E4, which is usually followed by a central proline-rich region and a charged domain near the C-terminus (Doorbar & Myers, 1996). A leucine-rich region is also identified in the N-terminus of cutaneous HPV type E1^E4 proteins, although only HPV-1, HPV-2 and HPV-63 E1^E4 proteins contain the LLXLL motif (Roberts et al., 1994; Doorbar et al., 1996). A C-terminal (DLXLYW) motif has been found in most cutaneous HPVs, while the mucosal HPVs show very little conservation in their C-terminal E4 protein sequence (Doorbar & Myers, 1996).

In numerous HPV and animal PV infected tissues, immuno-detection of E1^E4 has showed that the protein is usually expressed at high levels during the late stage of the infection (Doorbar et al., Nicholls et al., 2001; Peh et al., in press; Chapter 5). When expressed in cells, E1^E4 exhibits a predominantly cytoplasmic localisation (Doorbar et al., 1986; Brown et al., 1991; Jareborg & Burnett, 1991). Formation of E1^E4 inclusions is apparent in some HPV types infections (Doorbar et al., 1989). For example, HPV-1 E1^E4 forms cytoplasmic and nuclear
inclusions of varying shapes and sizes in infected tissues, as well as in monolayer cell cultures (Rogel-Gaillard et al., 1993). Cytokeratin association is also observed in a number of HPV E1^E4 expressing cells in vitro. For some HPV types, E1^E4/cytokeratin association leads to the collapse of the cytokeratin network (e.g. HPV-16, HPV-31b) (Doorbar et al., 1991; Pray & Laimins, 1995), while others do not result in this dramatic cytopathic effect (e.g. HPV-1) (Rogel-Gaillard et al., 1993). Mutational studies of HPV-16 E1^E4 showed that the leucine cluster (LLKLL) has an important role for the association of E4 with the cytokeratin network (Roberts et al., 1994). In addition, the N and C-terminal sequences of the cutaneous HPV-1 E1^E4 proteins were also implicated in the interaction and binding affinity of the protein with cytokeratins respectively (Roberts et al., 1994). Recently, E1^E4 has been found to be able to associate with other cellular structures. HPV-11 E1^E4 associates with the cornified cell envelope (Brown & Bryan, 2000; Bryan & Brown, 2000), and COPV E1^E4 with the nucleolus (Nicholls et al., 2000; Peh et al., in press; Chapter 5).

1.3 Viral Late (L) Proteins

The viral late proteins, L1 and L2, are also known as the major and minor capsid proteins respectively. Together they form the icosahedral capsid (approximately 55nm in diameter) that contains the viral genome (Baker et al., 1991). Although the L1 protein alone can self-assemble into virus-like particles (VLPs), the L2 protein seems to be essential for the infectivity of virions (Stauffer et al., 1998). Expression of late transcripts containing the L1 and L2 ORFs is restricted to the highly differentiated layers of infected epithelium (reviewed in Howley and Lowey, 2001). Different regulatory mechanisms are believed to be involved in the control of L1 and L2 transcripts and protein expression in the uppermost layers of the epithelium. Production of the L1 and L2 containing mRNAs are controlled by the viral differentiation-dependent promoter, which is activated only upon keratinocyte differentiation. L1 and L2 protein expression can also
be regulated post-transcriptionally, such as at the level of codon usage (Zhou et al., 1999). In addition, negative regulatory RNA sequences on papillomavirus late mRNAs have been identified in the coding and non-coding regions of different PV late transcripts (Schwartz, 1998 & 2000). HPV-1 late transcripts contain a 57-nucleotide AU-rich inhibitory element in their late gene 3' untranslated region that reduces mRNA stability and inhibits translation (Tan & Schwartz, 1995; Sokolowski et al., 1997). Similar inhibitory sequences were also identified in the 3' untranslated region in BPV-1 (Furth & Baker, 1991) and HPV-16 late transcripts (Kennedy et al., 1990 & 1991). In BPV-1, this sequence acts by inhibiting polyadenylation (Furth et al., 1994), while in HPV-16 it reduces mRNA stability in vitro (Kennedy et al., 1991). It has also been found that specific cellular factors can interact, and adversely affect the inhibitory activities of these AU-rich elements (Furth et al., 1994; Zhao et al., 1996; Koffa et al., 2000). Together, these regulatory mechanisms act to ensure that expression of the L1 and L2 proteins are confined to the upper layers of the infected epithelium. It has been suggested that the expression of L2 proceeds L1 in infected tissues (Florin et al., 2002). Day et al. (1998) have also shown that the L2 protein localises to nuclear domains 10 (ND10) and recruits L1 into these domains, where assembly of infectious virions is thought to occur.

1.4 The Epithelium

Epithelial tissues are classified into keratinizing (cornified) stratified squamous epithelia, stratified non-keratinizing epithelia, and simple (non-stratified) epithelia, based on their morphology and differentiation-specific expression patterns (Presland & Dale, 2000). The skin, palate and oral gingival tissue are examples of keratinizing epithelia, whereas the oesophagus and buccal oral mucosa are made of stratified non-keratinizing epithelial cells. Only cells of the stratified keratinizing epithelia undergo terminal differentiation to form a highly dehydrated protective surface composed of cornified cells (or corneocytes). These cells are toughened by a
proteinaceous cornified envelope (CE), are filled with keratin filaments, and lack nuclei and cytoplasmic organelles (Melino et al., 1998; Nemes & Steinert, 1999). The differentiation program of the epithelium is a carefully orchestrated process whereby the dividing basal cells lose their ability to proliferate as they migrate upwards from the basal layer (Eckert, 1989; Coffman & Studzinski, 1999). During this process, the cells undergo a series of morphological and biochemical transformations which require the coordinated expression of host of cellular genes (Eckert et al., 1997). This enables the epithelial structure to be divided into different layers according to the cellular changes that occur during differentiation (Presland & Dale, 2000). The keratinizing epithelium can be divided into the basal, spinous, granular and cornified layers, in an ascending order. The differentiated cells in the non-keratinizing epithelium are morphologically less well-defined, and are divided into the basal, intermediate and superficial layers. Cutaneous PVs cause specific infections in the skin epidermis, whereas mucosal PVs target non-keratinizing epithelia. Since the PV life cycle is dependent on epithelial differentiation, the activation of life cycle events is likely to be directly influenced by the cellular changes and regulatory mechanisms that occur during cellular differentiation.

1.5 PV-Associated Diseases

The main biological characteristics of PVs include their species and tissue type (mucosal or cutaneous) specificity, their preference to cause lesions in particular regions of the body, and their ability to induce hyperproliferation of epithelial cells and to form papillomas. Controversial data surround the investigation of the PV cell receptor(s), whereby several candidate receptors have been identified, including alpha-6 integrin, heparan sulfate and probably a secondary receptor that remains to be identified (Evander et al., 1997; Joyce et al., 1999; Sibbet et al., 2000; Giroglou et al., 2001). In the face of the present controversies regarding this issue, the importance of the PV cell receptor in determining the infection of specific epithelial cell types...
cannot be addressed. However, it has been suggested that the host and tissue type specificity exhibited in PV infections may be dependent on the activity of the LCR of the PV genome, which regulates viral gene expression, and not by the uptake of virus into cells (Roden et al., 1994; Muller et al., 1995; Sailaja et al., 1999; reviewed in Lowy & Howley, 2001). The compatibility between the intracellular virus-host interactions may explain the association of specific HPV types with particular lesions.

1.5.1 Cutaneous HPV diseases

Cutaneous type HPVs can infect non-genital skin throughout the body. However, specific HPV types appear to be consistent in causing infections at favoured locations, resulting in the development of typical wart morphology. For example, HPV-2 and HPV-4 are usually the causative agents of common warts (multiple) found on the hands, HPV-1 and HPV-63 cause the development of solitary deep plantar warts at the sole of the foot, and HPV-3, HPV-10, HPV-28 and HPV-41 cause flat warts (multiple) around the arms, face and knee regions. Most cutaneous warts regress spontaneously following a self-limiting virus infection, and seldom persist indefinitely (Messing & Epstein, 1963).

A class of cutaneous HPV, known as the EV HPVs, cause widespread skin lesions in individuals with a rare disorder known as epidermodysplasia verruciformis (EV), or in immunosuppressed individuals (Berger et al., 1991; Leigh et al., 1999). EV is an inherited disorder characterised by the unique susceptibility to chronic cutaneous HPV infections (Majewski et al., 1997). The types of skin lesions in EV patients are usually a mixture of flat warts (commonly caused by HPV-3 and HPV-10) and scaly red-brown macules (pityriasis-like lesions caused by the EV-associated HPVs, frequently HPV-5 and HPV-8). About one third of the patients suffering from pityriasis-like lesions develop skin cancer in association to their lesions. The most commonly isolated HPV types in these skin cancers are HPV-5 and HPV-8.
1.5.2 Mucosal HPV diseases

Mucosal type HPVs can not only infect the genital (vaginal tract, cervix and anal canal) and nongenital (oral, conjunctival and respiratory tract) mucosa, but also the genital skin (reviewed in Lowy & Howley, 2001). Genital warts (condylomata acuminata) are the most commonly diagnosed HPV lesions found in the external female genitalia, male penis and anus regions of sexually active individuals. These are usually associated with HPV-6 and HPV-11 (low-risk HPVs) infections and can regress spontaneously or persist if the infections are left untreated (Beutner & Trying, 1997). On the other hand, some HPV types, such as HPV-16, HPV-18, HPV-31 and HPV-45, are known as high-risk type HPVs. Although these high-risk HPVs can cause benign lesions, they are also frequently found in invasive cervical cancer in the over 60s female population (Nakagawa et al., 1996; Koutsky, 1997). HPV-16 DNA can be found in approximately 40-60% of cervical cancer, and HPV-18 DNA in 10-20% of malignancies (Beutner & Trying, 1997). The development of cervical cancer is likely to involve additional cofactors besides infection by high-risk HPVs. Smoking, oral contraceptives, pregnancy, parity, nutrition, other sexually transmitted diseases, and certain HLA class II alleles, are some of the factors that have been suggested to be associated with increased risk of developing cervical cancer (Daling et al., 1992; Stern, 1996; Giuliano et al., 1997; Kjellberg et al., 2000). The key area for the development of cervical cancer is in the transformation zone of the cervix. Traditionally, the different stages of cancer progression in the cervix can be classified as cervical intraepithelial neoplasia (CIN) grades 1, 2 and 3 (reviewed in Lowy & Howley, 2001). The severity of the lesion is determined by the degree to which the squamous epithelium is replace by basaloid cells, for instance in a CIN-1 lesion, about one third of the epithelial thickness will be replaced by basaloid cells, whereas the entire epithelium is replaced by basaloid cells in a CIN-3 lesion.

HPV lesions in the oral cavity are common and are frequently associated with the genital HPV types (reviewed in Praetorius; 1997). Oral lesions can be classified into three main types (reviewed in Lowy & Howley; 2001). Oral squamous cell papilloma are solitary benign tumours,
mostly associated with HPV-6, HPV-11, HPV-16 and HPV-57. Oral condyloma accuminatum are often multiple and derived from HPV-6 and HPV-11 infections. Oral verruca vulgaris at the vermillion border of the lip is caused mainly by HPV-2 infections. HPV-associated papillomatosis in the respiratory tract and conjunctiva is caused primarily by HPV-6 and HPV-11 (Mounts & Shah, 1984; McDonnel et al., 1987). In the respiratory tract, HPVs can infect the laryngeal, tracheal, lung and nasal epithelia, although the development of papilloma in these regions is uncommon (Mounts & Shah, 1984; Wu et al., 1993). Conjunctival papillomas are also generally rare (McDonnel et al., 1987).

1.6 The PV Life Cycle

1.6.1 Productive infection

The productive life cycle of PVs is supported only in differentiating epithelium, and can be divided into the early and late stages (Figure 1.3). The early stage of the virus life cycle occurs in the lower layers of the infected epithelium where episomal viral DNA is maintained at low copy number, and the viral E6, E7, E1 and E2 proteins are expressed at low levels. The E6 and E7 proteins work together to prolong the proliferative state of the infected cells, whilst E1 and E2 are required for viral DNA replication and episomal maintenance. Histologically, all PV infected papillomas display several characteristics typical of PV infection. One of the earliest morphological changes observed is the thickening of the epithelial tissue (acanthosis), and in some types of epithelia, papillae formation (e.g. in HPV-1 induced warts). Other pathological features include an increase in the number of basal and parabasal cell mitoses, an increase in tissue pigmentation and intracellular inclusion granules, the presence of koilocytes, the preservation of nuclei in higher differentiating layers, and the retention of nuclei in corneocytes (parakeratosis) (Lowy & Holwey, 2000). These histological characteristics are likely to be induced by viral early events since the expression of CRPV E6, E7, E1 and E2 proteins in New
Zealand White (NZW) rabbits is sufficient to cause the development of rabbit papillomas that showed typical wild-type (wt) viral papilloma morphology (Wu et al., 1994). However, the events that lead to the development of such pathological observations remain poorly understood.

The activation of the late stage of the PV life cycle is linked to cell differentiation. Late events are triggered following the activation of the viral differentiation-dependent promoter, and include viral DNA amplification, expression of E4, L1 and L2 proteins, as well as virus particle assembly (reviewed in Howley & Lowy, 2001). A change in viral transcription pattern results in the expression of viral transcripts containing the E1, E2, E4, L1 and L2 ORFs (Baker & Howley, 1987; Hummel et al., 1992). From immuno-histochemistry studies of PV infected tissue sections, expression of E4 and viral DNA amplification have been shown to correlate closely in human and animal lesions (Doorbar et al., 1997; Nicholls et al., 2000; Peh et al., in press; Chapter 5). This is followed by the synthesis of the viral capsid proteins (L1 and L2) in the nucleus (Frattini et al., 1997). Increased levels of E1 and E2 are likely to be required for the amplification of the viral DNA although a role for E4 has also been implicated (Doorbar et al., 1997; Chapter 8). The restricted expression of the viral capsid proteins in the highly differentiated cells has been shown to be negatively regulated by inhibitory elements in the viral late mRNAs (Furth et al., 1994; Schwartz, 1998; Koffa et al., 2000; Collier et al., 2002). The exact mechanisms that govern the expression of L1 and L2 are not well-understood.

1.6.2 Carcinoma

Malignant conversions of benign papillomas at cervical sites can occur. Histologically, this involves the gradual increase in the ratio of undifferentiated to differentiated epithelial cells in the transformation zone, until the entire thickness of the epithelium is replaced with undifferentiated basaloid cells (reviewed in Lowy & Howley, 2001). This dysplastic change typically occurs over several years before the malignant carcinoma is established (reviewed in Lowy & Howley, 2001). Unlike in productive infections, most malignant cells harbour integrated
HPV DNA in the host chromosome (Durst et al., 1985; Cullen et al., 1991). The event of viral DNA integration results in the deletion of large segments of the virus genome, and the constitutive expression of the E6 and E7 ORFs (Schwarz et al., 1985; Baker et al., 1987). Disruption of the virus genome usually occurs in the E2 ORF (Daniel et al., 1997). The loss of E2 as a transcriptional repressor enables the increase in E6 and E7 expression in the cell, which is a major contributing factor leading to the hyper-proliferation of the undifferentiated basaloid cells (Jeon & Lambert et al., 1995). Loss of epithelial differentiation in the zone of dysplasia also results in the disruption of the late stage of the virus life cycle (Laimins, 1996). Thus during the development of malignancy, the virus gradually loses its ability to complete its productive life cycle due to the decrease in the differentiated layers of the epithelium.
Figure 1.3. The PV life cycle.
The life cycle of PVs is dependent on the differentiation of the infected epithelium. PVs infect the epithelial basal cells via a micro-wound. Activation of the viral early promoter in the infected basal cells triggers the expression of viral transcripts which terminate at the early poly-adenylation site (PAE). These transcripts allow the expression of the viral early proteins, including E6, E7, E1 and E2. E6 and E7 proteins disrupt normal cell cycle progression and cause cells to be in a hyperproliferative state. E1 and E2 proteins are required for the replication of virus DNA, and maintaining viral genome levels at about 10-20 episomal copies per cell. As the infected cells differentiate, the differentiation-dependent promoter is activated. This changes the viral transcription pattern to the expression of transcripts that terminate at the late poly-adenylation site (PAL). At the late stage of the virus life cycle, E4 protein expression and viral genome amplification are triggered in the same cells. Expression of the viral capsid proteins occur in the highly differentiated superficial layers of the epithelium.
1.7 Immunity to PV Infections

Both the cell-mediated and humoral (antibodies) immune responses are important in eradicating PV infected cells and infection (Frazer, 1996). Cell-mediated immunity is responsible for the destruction of PV-infected cells and regression of PV-associated lesions (reviewed in Malejczyk et al., 1997), while humoral immunity inactivates PV particles and protects against reinfection (reviewed in Carter & Galloway, 1997). However, specific immunity generated against PV infections appears to be type-specific, since patients can be infected with multiple HPV types at one time, and the infection by one type does not influence the infections by other types (Thomas et al., 2000). In addition, individuals who are resistant to one HPV type may not be resistant to re-infection by other HPV types (reviewed in Lowy & Howley, 2001).

1.7.1 Cell-mediated immunity

Cell-mediated immunity is mediated by infiltrating lymphoid cells, such as T-helper cells, cytotoxic T-cells and natural killer cells, as well as local antigen presenting cells (e.g. Langerhans cells) and keratinocytes (Malejczyk et al., 1997). The mechanisms involved include cytokine production, the activation of specific immune cell responses by antigen presenting cells, the recruitment of immune cells to the site of infection, and cell-mediated cytotoxicity. Cytokines are implicated in the down-regulation of HPV gene expression (Kyo et al., 1994), growth inhibition of HPV-infected cells (Khan et al., 1993), and triggering infiltrating leukocyte responses (Malejczyk et al., 1991). The antigen presenting property of langerhans cells and keratinocytes, and their subsequent interactions with the cytotoxic T-cells, determines the regression or persistence of the lesions (Malejczyk et al., 1997). The authors suggested that, in addition to the presentation of viral antigens via the MHC class II molecules to the T-cell receptor, the interaction between the B27 (on antigen presenting cells or keratinocytes) and CD28 (on T-cell) surface molecules may play a crucial role in directing the outcome of the infection (i.e. lesion regression or persistence). One of the most significant characteristics of lesion regression is the
active infiltration of immune cells, such as T-cells, natural killer cells, and monocytes/macrophages, to the site of infection (Coleman et al., 1994). Cytotoxic T-cells are the main cells involved in eradicating HPV-infected cells, and mediating lesion regression (Coleman et al., 1994). The importance of cellular immunity can also be illustrated in immuno-suppressed and immuno-deficient patients (Benton et al., 1992; Petry et al., 1994). In these patients, the abrogation of the specific or non-specific cellular immune responses allows lesion growth and persistence, which may result in the accumulation of chromosomal mutations and subsequent loss of antigenicity of HPV-immortalised cells, and lead to an increased risk of HPV-associated carcinoma (Malejczyk et al., 1997).

1.7.2 Humoral immunity

Specific anti-HPV antibodies can be detected in individuals with previous, current or subclinical HPV infections (reviewed in Stern, 1996). Although the presence of humoral immunity is central to the prevention of re-infection (as a result of virus neutralisation and eradication of infected cells via antibody-dependent cellular cytotoxicity (Stern, 1996)), experimental infections of animals with PV demonstrated that presence of antibodies was not sufficient to induce the regression of established infections (Breitburd et al., 1995; Suzich et al., 1995; reviewed in Nicholls & Stanley, 2000). More recently, it has been showed that antibodies to conformational L1 epitopes are effective in protecting dogs and ox from further infections with COPV and BPV-4 respectively (Kirnbauer et al., 1996; Ghim et al., 1997). Antibodies can work by neutralising virus infection at different stages, such as by preventing virus-cell attachment, or at the post-attachment stage (Roden et al., 1994a & 1994b; Christensen et al., 1995). Since virus-cell receptor binding and internalisation is a complex multi-step process (Haywood, 1994), specific antibodies may work at the secondary binding, virus entry, or uncoating step(s) other than preventing virus attachment to the cell (Unckell et al., 1997).
1.7.3 Vaccine development

Due to the high prevalence of HPV infections in anogenital and skin cancers, the search for a vaccine for HPV infections has become increasingly important (Schiller & Okun, 1996). Vaccines can be designed to work at three different stages during virus infection. Firstly, a prophylactic vaccine that can induce a neutralising antibody response, effective in protecting the host against future infections. Secondly, a therapeutic vaccine that can eliminate an ongoing productive infection, and thirdly, a therapeutic vaccine that will control the growth of invasive tumours (Tindle, 1996). Most vaccine studies of PV infections are carried out in animal models, particularly the CRPV-rabbit, BPV-cattle, and COPV-dog models (Nicholls & Stanley, 2000). The majority of these studies focus on the development of prophylactic vaccines using purified virus particles, virus-like particles (VLPs), recombinant proteins, virus DNA preparations and homogenised wart extracts (reviewed in Nicholls & Stanley, 2000). Although individual viral proteins have been shown to be effective as prophylactic and therapeutic vaccines (Campo et al., 1993; Lin et al., 1993; Selvakumar et al., 1995), the use of VLPs and DNA-based vaccines are also currently being explored. Immunisation with VLPs generated long-term protection against re-infections in ox, rabbits, and dogs (Kirnbauer et al., 1992; Breitburd et al., 1995; Ghim et al., 1995; Christensen et al., 1996). Immunisation of rabbits using DNA-based vaccines showed that the viral early and late genes can be used to preferentially elicit the cell-mediated and humoral immune responses respectively (Sundaram et al., 1997 & 1998; Han et al., 1999a & 1999b). No doubt, the use of recombinant VLPs and DNA vaccines will provide a greater flexibility in the development of vaccines against HPV infections.
1.8 Experimental Models for PV Infection

PV infections are highly species specific. This strict virus-host relationship has hindered the study of HPV infections in experimental animals until the recent development of in vitro differentiating organoraf cultures. Animal PVs and transgenic mice models have also been used for PV research. Among the most commonly used animal PVs are CRPV, BPV and canine oral PV (COPV). The application of particle-mediated gene delivery technology (genegun) for the experimental inoculation of PV DNA has also provided an advance in the experimental infection of PV DNA.

1.8.1 Differentiating organo-raft cultures

The productive life cycle of HPV was first re-created in vitro in a differentiating organoraf system by Meyers et al. (1992; Figure 1.4). Since then, it was possible to study the life cycle of different HPV types, such as HPV-11, HPV-16, and HPV-18, in vitro with this system (Meyers et al., 1997; Flores et al., 1999; Thomas et al., 2001). Oragnoraf cultures can be differentiated from primary human keratinocytes or immortalised cervical carcinoma cell lines, such as HPV-31b infected CIN-612 cells and HPV-16 infected W12 cells (Meyers et al., 1994). Keratinocytes harbouring the HPV DNA are induced to differentiate at an air-liquid interface, above a collagen matrix of feeder fibroblast cells, until a fully differentiated epithelium is synthesised approximately after two weeks (Meyers et al., 1994). This system is particularly useful for the study of the effects of genetic mutations in the virus genome, as well as the viral late events that are triggered only in the differentiated layers of the epithelium.

1.8.2 CRPV-rabbit model

CRPV infection is a naturally-occurring disease of wild CT rabbits, commonly found in states bordering the Mississippi River (Shope, 1933). The virus was first isolated from its natural host in the 1930s and shown to be responsible for the growth of large cutaneous warts on the
rabbit skin (Shope, 1933; Figure 1.4). Experimental infections of CRPV have shown that the virus can infect the cutaneous epidermis of numerous other rabbit species, including the jack, snowshoe and domestic rabbits (Kidd & Parsons, 1935). However, active viruses were only recovered from CT, jack and snowshoe rabbits, and very rarely if ever, from domestic rabbits (Shope, 1935; Selbie & Robinson, 1947; Beard & Rous, 1935). The course of CRPV infection in both naturally-occurring and experimentally-induced papillomas in CT and domestic rabbits was shown to be influenced by similar factors and involved both virus and host variations (Kidd et al., 1935; Syverton et al., 1950; Ginder, 1952). Besides the difference in virus-productivity in the CT and domestic rabbit infections, the rate of benign-to-malignancy progression also differed significantly between the two rabbit species. CRPV infections of domestic rabbits showed a 75% rate in carcinoma development, whereas only 25% of CRPV infections progressed into cancers in CT rabbits (Syverton, 1952). One likely explanation for the non-productivity of CRPV infections in domestic (NZW) rabbits maybe linked to the failure to activate the viral late promoter, and to express the L1 and L2 ORFs (Nasseri & Wettstein, 1984; Zeltner et al., 1994). It is not clear if the lack of late promoter activity may also contribute partly to the high rate of carcinoma development in the domestic rabbits. The early stage of the CRPV life cycle does not seem to be affected in the 'secondary' domestic rabbit host, as studies of the early viral transcript patterns showed similarity in the expression of viral transcripts in both the CT and NZW rabbit hosts (Phelps et al., 1985).

The CRPV-rabbit model has been widely used in different aspects of PV research. It has proved to be a powerful tool in studies of viral gene functions, in vaccine development, and for the identification of host factors that influence disease outcome (Brandsma et al., 1991 & 1992; Han et al., 1992; Wu et al., 1994; Breitburd et al., 1995 & 1997; Christensen et al., 1996; Donnelly et al., 1996; Selvakumar et al., 1997; Sundaram et al., 1997; Han et al., 2000; Lechman et al., 2000). In addition, this model also offers advantages in terms of the handling of small animals and the development of experimental protocol to study CRPV infections.
Recently, a new CRPV subtype (CRPVb) with significant differences in the E6 and E7 gene was described by Salmon et al. (1997). Unlike CRPV, CRPVb is able to complete its productive life cycle in domestic (NZW) rabbits (Salmon et al., 2000). The isolation of CRPVb has made CRPV-rabbit infections a convenient experimental model for the study of virus evolution and host genetics (Salmon et al., 1997 & 2000).

1.8.3 BPV-bovine model

PVs that cause epithelial infections in cattle belong to two different branches in the PV phylogeny (see Figure 1.1 & 1.4). BPV-1 is the most extensively used animal model for the study of PV replication and promoter usage. This model has enabled the identification of viral and host proteins that are required for virus replication and transformation (Yang et al., 1991; Seo et al., 1993; Muller et al., 1994; Melendy et al., 1995). Examination of BPV-1 genome organisation, as well as its transcription patterns also allowed the identification of multiple PV promoters and viral transcription patterns (Stenlund et al., 1985 & 1987; Ahola et al., 1987; Baker and Holwey, 1987).

BPV-2 and BPV-4 have also been used as models for vaccination studies (Jarrett et al., 1990a, 1990b & 1991; Campo et al., 1993). BPV-2 is a cutaneous fibropapillomavirus, whereas BPV-4 is a mucosal tropic PV that causes oral lesions, as well as alimentary carcinomas in cattle (Campo, 1997). Vaccinations using DNA-based, peptide-based or VLP vaccines were able to protect the animals from re-infection by the respective BPV type (Campo et al., 1997). However, the difficulties in handling cattle, as well as the cost of housing them in research facilities have severely restricted the widespread use of this animal model.

1.8.4 COPV-canine model

COPV causes mucosal papillomas in the oral cavity of dogs (Figure 1.4). COPV has become popular as a model for vaccine development, particularly for the development of
vaccines against mucosal PV infections (Chambers et al., 1960; Bell et al., 1994; Ghim et al., 1995; Suzich et al., 1995; Stanley et al., 2001). Although some similarities have been found between COPV infections and HPV-associated recurrent laryngeal papillomatosis (Nicholls et al., 1999), the COPV life cycle is unique in a number of ways, and does not generally resemble most mucosal HPV life cycles (Nicholls et al., 2001; Peh et al., in press). The COPV genome contains a large non-coding region (NCR-2) between its E2 and L2 ORFs which is not found in any other PV genome (Delius et al., 1994). In addition, the COPV genome lacks an E5 ORF (Delius et al., 1994). At this time, it is not clear how these distinct features of the COPV genome affect the virus life cycle. Based on protein and DNA sequence analyses, COPV seems to be most closely related to the cutaneous PVs, HPV-1 and HPV-63 (Delius et al., 1994). Despite these unusual characteristics, COPV remains useful as a model for the development of vaccines and the study of the PV life cycle.

1.8.5 ROPV-rabbit model

ROPV was isolated from mucosal papillomas which were found on the under surface of the tongue of domestic rabbits in the New York area (Parsons & Kidd, 1936; Figure 1.4). The complete sequence of a ROPV strain was recently published and found to be most closely related to the CRPV genome (Christensen et al., 2000). However, ROPV does not infect the cutaneous skin or the mucous membrane of the nose, conjunctiva and female genitals, and is rarely found elsewhere on the buccal mucous membrane (Parsons & Kidd, 1936). Experimental infections of the adult rabbit male genital tissues have found that ROPV-induced lesions can be produced at these sites, making this a suitable animal model for genital HPV infections (Harvey et al., 1998). Serological studies following ROPV infections as well as DNA vaccination have also been carried out using this model, proving its usefulness as an experimental model for immunological studies of PV infections (Christensen et al., 2000; Han et al., 2000). ROPV is likely to become a useful animal model for the study of mucosal HPV infections.
1.8.6 Murine models

The discovery of the athymic 'nude' (Flanagan, 1966; Pantelouris, 1968) and severe combined immunodeficiency (SCID) (Bosma et al., 1983) mice, has opened up new avenues in experimental transplantation and cancer research, including PV research (Reed & Manning, 1973; Giovanella & Fogh, 1985; Phillips et al., 1989; Hendrickson, 1993). In 1985, histological features of HPV infection and virus production were demonstrated following the transplantation of HPV-11 (Hershey strain) infected human cervical material under the renal capsule of nude mice (Kreider et al., 1985). Subsequently, virus production was shown using other human tissue types, such as human foreskin, in the same animal model (Kreider et al., 1986, 1987a & 1987b). This method has now been adopted for the propagation of other animal and human PVs such as BPV-1, BPV-4, CRPV, HPV-1 and HPV-16 (Gaukroger et al., 1989; Stanley et al., 1989; Christensen & Kreider, 1990; Kreider et al., 1990). A renal capsule transplantation study using HPV-11 infected tissue in nude and SCID mice revealed that the latter is a more efficient and versatile model than its predecessor (Bonnez et al., 1993). Transplantation approaches have also shown that the productive HPV-16 life cycle can also be completed in human skin xenografts inoculated with viral DNA then grown on the backs of SCID mice (Brandsma et al., 1991; Bonnez et al., 1998). The nude and SCID mice models are undoubtedly good non-human animal models for the study of HPV infections (reviewed in Bonnez, 1998).

1.8.7 Particle-mediated gene delivery technology (Genegun)

Particle bombardment of genetic material into cells, animals, human and plants using jet injectors (Figure 1.4) is designed primarily for the delivery of DNA vaccines in vivo (Wolff et al., 1990; Yang et al., 1990; Tang et al., 1992; Davis et al., 1994). The first successful application of genegun technology in PV research was demonstrated in the CRPV-rabbit model. Brandsma et al. (1991) showed that papilloma formation and growth can be efficiently induced following the inoculation of CRPV DNA onto the backs of domestic rabbits. The authors also showed that the genegun DNA delivery method is more efficient and economical to the conventional inoculation
methods by scarification, scratching or puncturing with razor blades or needles. This approach was soon extended to the investigation of viral gene function by delivering mutant forms of the CRPV genome into rabbit skin (Brandsma et al., 1991 & 1992; Wu et al., 1994). Genegun technology is now widely used in PV vaccination studies (Xiao & Brandsma, 1996; Sundaram et al., 1997; Han et al., 2000; Stanley et al., 2001) as well as for experimental inoculation of epithelial tissues fragments (Brandsma et al., 1995).
Figure 1.4. Experimental tools used in PV research and infections caused by animal PVs. a) Shows the process by which a differentiating organo-raft culture is grown in vitro. I. Epithelial cells are placed on a submerged collagen matrix containing fibroblasts. II. When the epithelial cells are confluent, the matrix is lifted onto a metal grid and fed from below. The epithelial cells stratify and differentiated typically after two or three weeks. b) A Helios gene gun (Bio-Rad Laboratories, USA) used for particle-mediated gene delivery. c) Cottontail rabbit with horn-like cutaneous papillomas caused by CRPV (holiday snapshot from the internet). d) Cutaneous papillomas developing around the head and neck regions of the cow as a result of BPV-2 infections (picture source - http://duke.usask.ca/~misra/virology/stud2002/warts/species.html). e) Oral papillomas in and around the oral cavity of a dog infected with COPV (Snickers and her warts). f) Oral lesions found under the tongue of a NZW rabbit caused by ROPV (from Harvey et al., 1998).
1.9 Project Objectives

PVs can infect a wide range of vertebrate hosts. Host and tissue specificity and the dependence on cell differentiation are conserved characteristics of all PVs. However, the general life cycle patterns of different PV types have not been extensively studied. More importantly, the role of the viral E4 protein in the PV life cycle has been elusive, even though several molecular mechanisms of protein function have recently been proposed (Davy, 2002). The primary aim of this thesis is to investigate the in vivo role of the PV E4 protein during the virus life cycle. The following are the main objectives of this investigation.

1. To compare the E4 protein sequences and study the E4 expression patterns of different PV types. E4 protein sequence analyses were carried out using alignment programmes such as Blast, Lalign and MegAlign. Specific E4 antibodies were made against COPV, CRPV, ROPV, BPV-1 and HPV-11. These were used to examine the E4 expression patterns in PV infected tissues samples. The intracellular expression patterns of E4 were studied and related to the E4 protein sequence analyses.

2. To relate E4 expression to other events in the virus life cycle. Viral DNA amplification and L1 expression in different PV infections were detected using in situ hybridisation and immuno-detection methods respectively. Double immuno-staining of E4 and viral DNA or L1 expression patterns showed that the late stage of the PV life cycle is conserved in all PV infections. Specific antibodies to surrogate markers of E7 expression were used to detect the presence of the early stage of the virus life cycle. Double immuno-staining of E4 and such surrogate markers showed an overlap between the early and late stage of the virus life cycle.

3. To design an E4 knockout mutant and investigate the role of E4 in the life cycle of the virus. A series of C-terminal truncation mutants were
constructed in the E4 ORF of HPV-16 and CRPV genome by site-directed mutagenesis. The effect of the CRPV E4 knockout mutant DNA was investigated in vivo following infections of CT and NZW rabbit hosts. The loss of full-length E4 expression showed a dramatic effect on the late stage of the CRPV life cycle in both the CT and NZW rabbit infections.
Chapter 2

CHAPTER 2 - GENERAL METHODS AND MATERIALS

DNA Methods

2.1 Bacterial transformation ................................................................. 50
  2.1.1 E. coli genotypes ................................................................. 50
  2.1.2 Preparation of electrocompetent E. coli cells ......................... 50
  2.1.3 Bacterial transformation by electroporation ......................... 51

2.2 DNA preparations ................................................................. 52
  2.2.1 Mini maxi DNA preparations .............................................. 52
  2.2.2 Caesium Chloride gradient .................................................. 52
  2.2.3 High quality midi DNA preparations ..................................... 54
  2.2.4 Phenol/Chloroform/Isomylalcohol extraction and ethanol precipitation ..... 55
  2.2.5 Visualising DNA on TBE (+EthBr) agarose gels by electrophoresis . 55

2.3 DNA cloning ............................................................................. 56
  2.3.1 Cloning of DNA fragments .................................................. 56
  2.3.2 DNA gel extraction by electro-elution ...................................... 57

2.4 Polymerase Chain Reaction .................................................... 57

2.5 Oligonucleotides ........................................................................ 58
  2.5.1 Oligonucleotide primer ............................................................ 58
  2.5.2 E4 mutant oligonucleotide primer preparation ....................... 58

2.6 Synthesis of the E4 mutant genomes ........................................ 59
  2.6.1 Kunkel site-directed mutagenesis (New England Biolabs Ltd, UK) . 59
    A) E4 mutant primers ................................................................. 59
    B) Cloning of the CRPV and HPV-16 E4 ORFs into pET-23b(+) .......... 59
    C) Verification of the CRE4-pET and W12E4-pET clones ................. 61
    D) Preparation of M13R408 helper phage .................................... 61
    E) M13R408 helper phage titration (top agar method) ..................... 62
    F) Transformation of CJ236 E. coli with CRE4-pET and W12-pET ...... 62
    G) Synthesis of single-stranded uracil-enriched DNA templates for mutagenesis ................................................................. 63
    H) Phosphorylation of mutant oligonucleotide primers .................... 64
    I) Annealing of mutant primers to the single-stranded uracil-enriched DNA template ................................................................. 64
    J) Synthesis of the complementary DNA strand ............................ 64
    K) Bacteria transformation .......................................................... 65
  2.6.2 Quikchange site-directed mutagenesis (Stratagene Europe, UK) .... 65
    A) Mutant E4 primers ................................................................. 65
    B) DNA mutagenesis ................................................................. 66
    C) Removal of the DNA template ............................................... 66
    D) Transformation of E. coli ...................................................... 66
    E) E4 mutation verification ........................................................ 67
  2.6.3 Cloning of mutated E4 ORFs into the parental genome plasmids .... 67
  2.6.4 Preparation of the re-circularised E4 mutant viral genome ........... 68

2.7 DNA Sequencing ....................................................................... 69
  2.7.1 Thermal cycle sequencing .................................................... 69
  2.7.2 Automated DNA sequencing ................................................ 70
  2.7.3 HPV-16 and CRPV sequencing primers .................................... 71
Chapter 2

Protein Methods

2.8 Glutathione spherose-transferase and maltose binding protein fusion proteins .................................................. 72
2.8.1 Primers ...................................................................................................................................................... 72
2.8.2 Preparation of the E1^E4 gene fragments ................................................................................................. 72
2.8.3 Transformation of E. coli ......................................................................................................................... 73
2.8.4 Verification of the DNA clones ................................................................................................................. 73
2.8.5 Expression of the GST- and MBP-E1^E4 fusion proteins ............................................................................ 74

2.9 Peptide conjugation ....................................................................................................................................... 75
2.9.1 Keyhole limpet hemocyanin (KLH)-peptide conjugation ............................................................................ 75
A) Activation of the KLH carriers with sulfo-MBS (quantities for making two peptide conjugation reactions) .... 75
B) KLH-peptide conjugation ............................................................................................................................ 76
C) Cysteine assay (described by Sigma-Aldrich Company Ltd, UK) ................................................................. 76
   I. Cysteine standard assay ............................................................................................................................. 76
   II. Cysteine assay to determine total amount of unreacted cysteine before conjugation ............................... 77
   III. Cysteine assay to determine peptide coupling efficiency ........................................................................ 77
2.9.2 Bovine serum albumin (BSA)-peptide conjugation .................................................................................... 78

2.10 Protein analysis .............................................................................................................................................. 78
2.10.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) .................................................. 78
2.10.2 Staining protein gels with comassie blue .................................................................................................... 79
2.10.3 Semi-dry protein-membrane transfer (Western blot) .................................................................................. 79
2.10.4 Detection of proteins on western blot ...................................................................................................... 79

2.11 Other protein methods ............................................................................................................................... 80
2.11.1 Determining protein concentration using the Bio-Rad protein assay dye reagent ........................................ 80
2.11.2 Concentrating protein solutions with buffer exchange ........................................................................... 80
2.11.3 Protein dialysis ....................................................................................................................................... 81

Antibodies and Tissue Processing Methods

2.12 Preparation of specific antibodies ............................................................................................................... 82
2.12.1 Immunisation of rats and rabbits ................................................................................................................ 82
A) Immunisation of rats ........................................................................................................................................ 82
B) Immunisation of rabbits .................................................................................................................................... 83
2.12.2 Preparation of the MBP-CRPV E1^E4 immunogen ..................................................................................... 83
2.12.3 Antibody titre and specificity tests ........................................................................................................... 83
A) Western blots ................................................................................................................................................ 83
B) Enzyme-linked immuno-sorbent assay (ELISA) ............................................................................................ 84

2.13 Tissue preparations and histology ............................................................................................................. 85
2.13.1 Sectioning of paraffin-embedded formalin-fixed tissue blocks ................................................................. 85
2.13.2 Hematoxylin and eosin staining ................................................................................................................ 85

2.14 Immunohistochemistry ............................................................................................................................... 86
2.14.1 Immunodetection on formalin-fixed paraffin-embedded tissue sections .................................................. 86
2.14.2 Antigen retrieval methods ........................................................................................................................ 87
A) Antigen retrieval buffers .................................................................................................................................. 87
B) Microwave denaturation .................................................................................................................................. 87
C) Pressure cooking .............................................................................................................................................. 87
D) Microwave denaturation and protease digestion ............................................................................................. 88
Chapter 2

2.14.3 Signal amplification system ................................................................. 88
   A) ABC signal amplification system ......................................................... 88
   B) FAST RED – substrate for alkaline phosphatase activity ..................... 89
   C) Tyramide signal amplification (TSA)-Direct – substrates for HRP activity 89

2.15 DNA fluorescent in situ hybridisation on tissue sections ......................... 90
   2.15.1 Generation of Digoxigenin (DIG)-labelled DNA probes .................... 90
   2.15.2 Pre-hybridisation protocol .............................................................. 90
   2.15.3 DNA-DNA hybridisation ................................................................. 90
   2.15.4 Post-hybridisation washes .............................................................. 91
   2.15.5 Detection of DIG-labelled probes ................................................... 91
DNA METHODS

2.1 Bacterial transformation

2.1.1 *E. coli* genotypes

DH5α - *deoR, endA1, gryA86, hsdR17(rKm1, streptomycin-resistant), recA1, relA1, supE44, thi-1*, Δ(lacZYA-argFV169), φ80dlacZΔM15, F', θ (Clontech, UK).  
CJ236 - a dut⁻ ung⁻ strain of *E. coli*. dut⁻ genotype denotes the deficiency in dUTPase, which is necessary for the conversion of dUTP to dUMP. The accumulation of the intracellular dUTP leads to its incorporation into DNA at sites normally occupied by thymine. The ung⁻ genotype denotes the expression of an inactivated form of uracil-N-glycosylase, an enzyme which removes uracil residues that have been incorporated into DNA. Therefore, uracil-enriched DNA will be synthesised in this strain of *E. coli* (New England BioLabs Ltd, UK).

2.1.2 Preparation of electrocompetent *E. coli* cells

<table>
<thead>
<tr>
<th>Luria-Bertani medium (LB)</th>
<th>10g Bacto tryptone, 5g Yeast extract, 10g NaCl in 1L deionised water. Sterilised by autoclaving.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>LB agar</em></td>
<td><em>LB</em> with 2% (w/v) Bacto agar.</td>
</tr>
<tr>
<td>Chloramphenicol (CAP) stock solution</td>
<td>34mg/mL (w/v) working stock solution.</td>
</tr>
<tr>
<td>10% Glycerol</td>
<td>10% Glycerol (v/v) in double distilled water. Sterilised by autoclaving.</td>
</tr>
</tbody>
</table>

A single bacterial colony picked from a LB agar plate was used to inoculate 5mls of sterile LB. The culture was incubated overnight at 37°C in a bacterial culture shaker. 1ml of the overnight culture was transferred into 100mls of fresh LB in a 2L flat bottom conical flask and incubated on a shaker at 37°C for about 2-4hrs, until the absorbance value at OD₆₀₀ was approximately 0.5 (± 0.03). The culture was chilled in a bucket of ice for 20mins, then transferred into two pre-chilled 250ml centrifugal pots (Corning Incorporated, USA), and spun at 2,500 x g
for 15mins in the pre-chilled Beckman J6-HC centrifuge (Beckman Instruments Incorporated, USA). The bacteria pellets were gently resuspended in 100mls of pre-chilled 10% glycerol. The bacteria suspension was transferred into one new 250ml centrifuge tube and chilled on ice for 10mins, then spun at 2,500 x g for 15mins (Beckman J6-HC centrifuge). The pellet was gently resuspended in 10mls of ice-cold 10% glycerol, then chilled on ice for 20mins. The bacteria suspension was divided into 10x1ml aliquots in sterile 1.5ml eppendorf tubes, and spun at 65g for 10mins in a microcentrifuge. The bacterial pellets were resuspended in 50μL of ice-cold 10% glycerol and used immediately, or stored at -70°C. *E. coli* strain CJ236 was grown in LB containing 15μg/mL (w/v) of CAP.

### 2.1.3 Bacteria transformation by electroporation

<table>
<thead>
<tr>
<th>LB</th>
<th>10g Bacto tryptone, 5g Yeast extract, 10g NaCl in 1L deionised water. Sterilised by autoclaving.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB agar</td>
<td>LB with 2% (w/v) Bacto agar.</td>
</tr>
<tr>
<td>LB agar (+amp)</td>
<td>LB agar containing 100μg/mL (w/v) Ampicillin.</td>
</tr>
<tr>
<td>Ampicillin (amp)</td>
<td>100mg/mL (w/v) working stock solution.</td>
</tr>
</tbody>
</table>

One aliquot of electrocompetent *E. coli* was thawed slowly at 4°C, then transferred into a pre-chilled 0.2cm (width) electroporation cuvette (Bio-Rad Laboratories, USA). DNA (0.5-1μL) was added to the cells. The cuvette was cooled on ice for 5mins. The cells were electrophorated at 2.5V, then quickly resuspended in 1ml of LB, then incubated at room temperature for 30-45mins in a sterile bijou, with gentle shaking. 100μL of the bacteria suspension was spread onto LB agar (+amp), and incubated overnight at 37°C. Transformed colonies were selected by their resistance to ampicillin (100μg/mL (w/v)).

To measure the transformation efficiency, the cells were electroporated with 1μg, 100ng, 10ng and 1ng of DNA and plated as described below. The transformation efficiency was calculated from the equation \[ \frac{(1 \times 10^{n+1}) \times (1/ \text{number of colony})}{n} \], where n represents the concentration of DNA used.
2.2 DNA preparations

2.2.1 Mini and maxi DNA preparations

**LB (+amp)** - LB containing 100μg/mL (w/v) Ampicillin.

A single bacterial colony was inoculated into 5mls of LB (+amp) and incubated in a bacteria shaker overnight at 37°C. For mini DNA preparations, 3mls of the overnight culture was used. DNA was extracted using the Wizard Plus Minipreps DNA Purification System (Promega Ltd, UK) and a vacuum manifold (according to the manufacturer's protocol), and eluted in 50μL of deionised water. For maxi DNA preparations, 500mls of LB (+amp) was inoculated with 500μL of a saturated bacteria culture and incubated overnight on a shaker at 37°C. DNA was extracted using the Qiagen Plasmid Maxi Kit (Qiagen Ltd, UK; according to the manufacturer's protocol) and resuspended in deionised water.

2.2.2 Caesium Chloride gradient

**LB or LB agar (+amp)** - LB or LB agar containing 100μg/mL (w/v) Ampicillin.

**10M Lithium chloride** - 42.4g in 100mls of deionised water.

**Solution 1** - 50mM Glucose, 25mM Tris-HCl solution (pH8.0), 10mM EDTA.

**Solution 2** - 0.2M NaOH, 1% (w/v) SDS.

**Solution 3 (pH4.8)** - 24.9g Potassium acetate, 11.5mls Glacial acetic acid in 100mls deionised water.

**TE solution** - 10mM Tris-HCl solution (pH8.0), 0.1mM EDTA.

**Caesium chloride/TE** - 1mg/1ml Caesium chloride in TE solution.

**Ethidium bromide (EthBr)** - 10mg/mL (w/v) working stock solution.

A 500ml bacteria culture was set up in LB (+amp) and incubated overnight at 37°C. The overnight culture was spun at 4500 x g (Beckman J6-HC centrifuge) in two 250ml centrifuge pots (Corning Incorporated, USA). The pellets were resuspended in 18mls of sterile Solution 1, and left to stand room temperature for 10mins. 40mls of freshly prepared Solution 2 was added and mixed well, and the mixture was incubated on ice for 5mins. 20mls of pre-chilled Solution 3 was added to the lysate, mixed well, and the mixture was incubated on ice for at least 15mins.
10mls of deionised water was added, and the mixture was spun at 4500 x g (Beckman J6-HC centrifuge) for 15mins. The supernatant was filtered through a piece of Miracloth (CN Biosciences Ltd, UK), and the filtrate was divided into two pre-chilled 50ml falcon tubes (Corning Incorporated, USA). An equal volume of iso-propanol was added and the tubes were incubated at -20°C for 30mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 20mins. The supernatant was discarded. The pellets were left to dry at room temperature by inverting the tubes on the bench, then resuspended in a total of 6mls of TE solution and transferred into a new 50ml falcon tube. 2mls of 10M lithium chloride was added, the tube was left on ice for at least 10mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 20mins. 20mls of ice-cold ethanol was added to the supernatant in a new 50ml falcon tube, the mixture was incubated at -20°C for 15mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 20mins. The pellet was air dried then resuspended in 5.5mls of caesium chloride/TE, and 1μg/mL (w/v) EthBr. With a syringe and needle, the caesium chloride mixture was loaded into a polypropylene Quickseal Centrifuge Tube (Beckman Instruments Incorporated, USA). The tube was sealed using the Beckman Tube Sealer and spun in the Beckman XL-90 Ultracentrifuge (TV65.2 rotor) for 15hrs, at 300,000 x g and at 15°C.

The DNA band was carefully collected from the caesium chloride gradient using a clean syringe and needle. The EthBr was removed by repeated extraction with water saturated butanol, until the DNA solution was colourless. The DNA solution was purified with phenol/chloroform/isoamylalcohol extractions, and the DNA was precipitated with ethanol (see Section 2.2.4). The DNA pellet was washed with ice-cold 70% (v/v) ethanol in water, then resuspended in 300-500μL of deionised water.
2.2.3 High quality midi DNA preparations

<table>
<thead>
<tr>
<th>LB (+amp)</th>
<th>LB containing 100μg/mL (w/v) Ampicillin.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10M Lithium chloride</td>
<td>42.4g in 100mls of deionised water.</td>
</tr>
<tr>
<td>Solution 1</td>
<td>50mM Glucose, 25mM Tris-HCl solution (pH8.0), 10mM EDTA.</td>
</tr>
<tr>
<td>Solution 2</td>
<td>0.2M NaOH, 1% (w/v) SDS.</td>
</tr>
<tr>
<td>Solution 3 pH4.8</td>
<td>24.9g Potassium acetate, 11.5mls Glacial acetic acid in 100mls deionised water</td>
</tr>
<tr>
<td>TE solution</td>
<td>10mM Tris-HCl solution (pH8.0), 0.1mM EDTA in deionised water.</td>
</tr>
<tr>
<td>EthBr</td>
<td>10mg/mL (w/v) working stock solution.</td>
</tr>
</tbody>
</table>

100μL of an overnight culture was used to inoculate 100mls of LB (+amp). The culture was incubated overnight at 37°C on a shaker. The overnight culture was spun at 3250 x g (Beckman J6-HC centrifuge) in a 250ml centrifuge pot (Corning Incorporated, USA). The pellet was chilled on ice, then resuspended in 6mls of pre-chilled Solution 1. 12mls of freshly prepared Solution 2 was added and the bacterial lysate was incubated on ice for 5mins. 9mls of pre-chilled Solution 3 was added, and the mixture was incubated on ice for 10mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 15mins. The supernatant was filtered through a piece of Miracloth (CN Biosciences Ltd, UK) and transferred into a pre-chilled 50ml falcon tube (Corning Incorporated, USA). An equal volume of iso-propanol was added and the mixture was incubated at -70°C for at least 30mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 20mins. The pellet was air dried at room temperature by inverting the tube on the bench, then resuspended in 3mls of TE solution. 1ml of 10M lithium chloride was added, and the mixture was incubated on ice for at least 10mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 20mins. The supernatant was transferred into a new pre-chilled 50ml falcon tube. The DNA was precipitated with 10mls of ice-cold ethanol at -20°C for 15mins, then spun at 4500 x g (Beckman J6-HC centrifuge) for 20mins. The pellet was washed with ice-cold 70% (v/v) ethanol in water and allowed to air dry. DNA was resuspended in 600μL of TE solution, then transferred to a sterile 1.5ml eppendorf tube and incubated with 3μL of DNase-free-RNase (Roche Diagnostics Ltd, UK) in a 37°C waterbath for 15mins. The DNA solution was purified with phenol/chloroform:isoamyl alcohol extractions, and the
DNA ethanol precipitated (see Section 2.2.4). The DNA pellet was washed twice with ice-cold 70% (v/v) ethanol in water, then resuspended in 300-500μL of deionised water.

2.2.4 Phenol/Chloroform/Isoamylalcohol extraction and ethanol precipitation

Equal volume of phenol (Invitrogen Ltd, UK) was mixed with the DNA solution in an eppendorf, and vortexed vigorously for 3x1 min. The mixture was spun in a microcentrifuge for 1 min at 12,000 x g to separate the phenol and DNA liquid phases. The top liquid phase (the DNA solution) was carefully removed and transferred into a new eppendorf. This phenol step was repeated. Equal volume of chloroform/i soamylalcohol (BDH Laboratory Supplies, UK) was added to the DNA solution and mixed well by repeatedly inverting the eppendorf. The mixture was spun in a microcentrifuge at 12,000 x g for 1 min. The top liquid phase (the DNA solution) was carefully removed and transferred into a new eppendorf. This chloroform/i soamylalcohol step was repeated twice.

Following the phenol/chloroform/i soamylalcohol extraction, the DNA was precipitated by adding 1/10th volume of 3M sodium acetate (pH 5.2), and 2.5x volume of ice-cold ethanol to the DNA solution, followed by an incubation at -70°C for at least 30 mins or on dry ice for 10 mins. The DNA was spun down at 12,000 x g for 10 mins in a microcentrifuge.

2.2.5 Visualising DNA on TBE (+EthBr) agarose gels by electrophoresis

| 10x Tris borate EDTA (TBE) | 5.4g Trizma base, 27.5g Boric acid, 20ml 0.5M EDTA (pH 8) deionised water. |
| EtBr | 10mg/mL (w/v) working stock solution. |
| 6x DNA Gel Loading Buffer | 0.25% (w/v) Bromophenol blue, 0.25% (w/v) Xylene cyanol, 30% (v/v) Glycerol deionised water. |
Agarose gels were made by dissolving agarose powder (Roche Diagnostics Ltd, UK) in 1xTBE solution. For example, 0.5g of agarose was dissolved in 50mls 1xTBE to make a 1% mini DNA gel (dimensions 8x8x0.8cm). 0.5µL of the EthBr working stock was added to the gel solution just before pouring. The agarose (+EthBr) gel was immersed in an electrophoresis tank in TBE buffer. Each DNA sample was prepared with an appropriate amount of the 6xDNA gel loading buffer. The DNA sample was ran at a constant voltage of 70V, and DNA bands were visualised on a uv wavelength emission table.

2.3 DNA Cloning

2.3.1 Cloning of DNA fragments

Specific restriction enzyme digest sites were chosen for the construction of each DNA clone. The insert and vector DNA were linearised by incubations with the specific DNA restriction enzymes, in the supplier's recommended buffers and temperature. To prevent vector DNA from self-ligation, the digested DNA was treated with calf intestine alkaline phosphatase (Roche Diagnostics Ltd, UK) at 37°C for 1hr before gel-purification. Enzyme digested insert and vector DNA were run in an agarose (+EthBr) gel, cut and gel-purified using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK; according to the manufacturer's protocol). Typically, the ligation mix contained insert:vector DNA in the ratio of 2:1. The ligation reaction was carried out overnight in a 14°C waterbath with T4 ligase in the manufacturer's buffer (New England BioLabs Ltd, UK). The ligation reaction was purified with phenol and chloroform/isoamyl alcohol extractions, and the DNA was ethanol precipitated (see Section 2.2.4). The DNA pellet was washed twice with ice-cold 70% (v/v) ethanol in water and resuspended in deionised water. This DNA was used for E. coli transformation (see Section 2.1.3).
2.3.2 DNA gel extraction by electro-elution

DNA bands were excised from the agarose (+Eth Br) gels using new, clean scalpels, and the DNA/gel pieces were stored in sterile eppendorfs. The DNA gel pieces were placed in pretreated dialysis tubings (see Section 2.11.3) with 1ml of TBE. Air bubbles were removed from the dialysis tube before the tube was clipped on both ends. The dialysis tube was submerged in an electrophoresis tank in TBE buffer. DNA was eluted from the gel under a constant current (140mA) for about 20mins. Once the DNA had eluded into the TBE buffer in the tube, the solution was purified with phenol and chloroform/isoamylalcohol extraction, then ethanol precipitated on dry ice for 10mins.

2.4 Polymerase Chain Reaction (PCR)

### Primer stock solution - 10pmol/μL in deionised water.

### dNTP mix stock solution - 5mM of each dNTP in deionised water.

Each PCR reaction contained 2pmol of linearised DNA template, 50pmol of each primer (forward and reverse), 4μL of dNTP stock solution, 10μL of 10x PCR buffer (with 15mM MgCl₂; Perkin Elmer, USA), and 1μL of *Taq* polymerase (Perkin Elmer, USA). The volume of each PCR reaction was made up to 100μL with deionised water. The PCR program consisted of thirty cycles of DNA denaturation (94°C for 1min), DNA annealing (60°C for 1min), and DNA elongation (72°C for 1.5mins). This was followed by a final DNA elongation step at 72°C for 7mins. The PCR fragments were purified using the Wizard PCR Preps DNA Purification System (Promega Ltd, UK) and a vacuum manifold, according to the manufacturer’s protocol.
2.5 Oligonucleotides

2.5.1 Oligonucleotide primer

All oligonucleotide primers were obtained commercially from Oswel Research Products Ltd (UK), unless otherwise stated. All Oswel synthesised oligonucleotides were delivered in solution with a pre-determined concentration.

2.5.2 E4 mutant oligonucleotide primer preparation

The forward primers used in the construction of the HPV-16 and CRPV E4 mutants were synthesised in the Institute of Virology, University of Glasgow. These were delivered in a column matrix. The oligonucleotides were extracted from the column matrix using concentrated ammonia solution, in a fumehood. A disposable 5ml syringe (syringe A) was used to take up 1.5mls of concentrated ammonia solution (BDH Laboratory Supplies, UK), then inserted tightly onto one end of the column. Approximately 0.2ml of ammonia was pushed through the column, until the liquid meniscus was seen at the other end of the column matrix. A second empty disposable 5ml syringe (syringe B) was inserted tightly at this end of the column matrix. The column was incubated in ammonia at room temperature for 20mins before another 0.2ml of ammonia was pushed from syringe A into the column, and incubated for 20mins at room temperature. This extraction step was repeated until all the ammonia solution was passed from syringe A, through the column into syringe B. After the final incubation, the ammonia solution was pushed back and forth through the column several times. The oligonucleotide solution was pulled all into one syringe and immediately transferred into a 1.5ml screw-cap cryo tube, which was sealed with nesofilms around the neck. The oligonucleotide solution was incubated overnight in a 55°C waterbath. The de-protected oligonucleotide solution was divided equally into two clean eppendorf tubes with punctured caps, and dried overnight in the Centrifugugal Evaporator RC 10-22 (Jouan S. A., France). The DNA pellet was resuspended in 100µL of...
deionised water and the concentration of the oligonucleotide solution was determined at OD_{260} (absorbance value of 1.0 at OD_{260} = 30\mu g/mL).

2.6 Synthesis of the E4 mutant genome

2.6.1 Kunkel site-direct mutagenesis (New England BioLabs Ltd, UK)

A) E4 mutant primers

One oligonucleotide primer containing the mutation(s) was required for each mutagenesis reaction. The following E4 mutant primers were synthesised in the in-house facility in the Institute of Virology, Glasgow. The oligonucleotides were extracted from their column matrix as described in section 2.5.2.

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Oligonucleotide sequence (5' -&gt; 3')</th>
<th>Primer length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HPV-16</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16st15/6/0</td>
<td>CTC CTG AAA TAA TAA GGC AGC ACT TAG CCA ACC</td>
<td>33</td>
</tr>
<tr>
<td>16st20</td>
<td>GGC AGC ACT TAG CCA ACC ACC CCG</td>
<td>24</td>
</tr>
<tr>
<td>16st32/4</td>
<td>CCA AAG CCG TAG CCT TAG GCA CCG</td>
<td>24</td>
</tr>
<tr>
<td>16st67/3</td>
<td>ACT CAG TAG ACA GTG CTC CAA TCC TGA CTG CAT</td>
<td>33</td>
</tr>
<tr>
<td>16st84</td>
<td>AAG GAC GGA TCC ACT GTA ATA GTA</td>
<td>24</td>
</tr>
<tr>
<td><strong>CRPV</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRst9/0</td>
<td>CAG GGG CAC TCA CTA GCG GCT GGG</td>
<td>24</td>
</tr>
<tr>
<td>CRst41</td>
<td>ACC TTC GTC TGC CTA GTG TAC GCG</td>
<td>24</td>
</tr>
<tr>
<td>CRst74</td>
<td>GAG CAG GCG GTC TTA TGT CCC CCG</td>
<td>24</td>
</tr>
<tr>
<td>CRst82</td>
<td>TAG TTC TTC GTC CTA CGT CCT CTG</td>
<td>24</td>
</tr>
<tr>
<td>CRst89</td>
<td>GAG GTG TTC TTC CTA CCG TCG TAG</td>
<td>24</td>
</tr>
</tbody>
</table>

B) Cloning of the CRPV and HPV-16 E4 ORFs into pET-23b(+)

E4 ORFs of CRPV (Washington B strain; Nasseri et al., 1989) and HPV-16 (W12 clone; Stanley et al., 1989) were cloned into pET-23b(+) vector plasmid (CN Biosciences Ltd, UK; Appendix 2) for the synthesis of the single-stranded DNA templates. The CRPV E4 ORF was excised from pLA-CRPV (CRPV genome cloned into pLAII vector at the Sal I
site) following a restriction enzyme digest with EcoR I and Sal I, which cut the CRPV genome at positions 1063 and 4571 respectively. The HPV-16 E4 ORF was excised from pSp-W12E (W12 clone of the HPV-16 genome cloned into pSp64 vector at the BamH I site) at the Hind II (position 3210) and BamH I (position 6150) restriction enzyme sites. pET-23b(+) DNA was digested at the multiple cloning site region with Sal I and EcoR I, or with Hind II and BamH I.

50µg of pLA-CRPV was digested with 50units of Sal I and 50units of EcoR I in SURECUT™ buffer H (Roche Diagnostics Ltd, UK) at 37°C for 2hrs. Similarly, 50µg of pSp-W12 was digested with 50units each of Hind II and BamH I in SURECUT™ buffer B (Roche Diagnostics Ltd, UK) at 37°C for 2hrs. 30µg of pET-23b(+) was digested with 30units each of Sal I and EcoR I in SURECUT™ buffer H (for CRPV), or Hind II and BamH I in SURECUT™ buffer B (for HPV-16), at 37°C for 2hrs, followed by a 3hr incubation at 37°C with 10units of calf intestine alkaline phosphatase (Roche Diagnostics Ltd, UK). The digests were ran in a 0.8% agarose (+EthBr) gel (see Section 2.2.5) and the digested DNA were extracted using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK), according to the manufacturer's protocol.

Approximately 6µg of the Sal I/EcoR I digested CRPV E4 ORF (or the Hind II/BamH I digested HPV-16 E4 ORF) and 3.75µg of the Sal I/EcoR I (or Hind II/BamH I) digested pET-23b(+) DNA was used in a ligation reaction containing 20units of T4 ligase (New England BioLabs Ltd, UK). The ligation was carried out overnight in a 14°C waterbath. The ligation reaction was phenol, chloroform/isoamyl-alcohol extracted and ethanol precipitated (see Section 2.2.4). The DNA pellet was washed with ice-cold 70% (v/v) ethanol in water and resuspended in 50µL of deionised water. 3µL of each ligation mix was used to transform one aliquot of electrocompetent DH5α (see Section 2.1.3).
C) Verification of the CRE4-pET and W12E4-pET clones

Twelve single colonies were picked from the CRE4-pET and W12E4-pET LB agar (+amp) plates, and grown overnight in 5mls of LB (+amp). The DNA from each clone was extracted using the Wizard Plus Minipreps DNA Purification System (Promega Ltd, UK; see Section 2.2.1). The clones were checked with restriction enzyme digests. 6μL of DNA and 5units of restriction enzyme were used in each reaction. The CRE4-pET clones were checked with a Xho I digest (in SURECUT™ buffer H; Roche Diagnostics Ltd, UK) and a Sal I/Bsm I double digest. The W12E4-pET clones were checked with a Xho I digest (in SURECUT™ buffer H; Roche Diagnostics Ltd, UK), and a Hind II/Bam H I double digest at 37°C for 1hr. The restriction digests were ran in 0.8% agarose (+EthBr) gels (see Section 2.2.5).

D) Preparation of M13R408 helper phage

100mls of LB was inoculated with 1ml of saturated CJ236 culture in a 1L flask, and incubated at 37°C on a shaker until the absorbance value at OD600 was approximately 0.2. 600μL of a M13R408 helper phage stock was added and the culture was incubated overnight at 37°C. The next day, the phage culture was spun at 8,500 x g in a Sorvall centrifuge (GS/3 rotor) at 4°C for 15mins. The clear supernatant was transferred into two clean 50ml falcon tubes (Corning Incorporated, USA) and heat inactivated at 65°C for 10mins. The phage supernatant was stored at 4°C until ready to use.
Chapter 2 – General Methods and Materials: DNA Methods

E) M13R408 helper phage titration (top agar method)

<table>
<thead>
<tr>
<th>LB or LB agar (+amp)</th>
<th>LB or LB agar containing 100µg/mL (w/v) Ampicillin.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top agar</td>
<td>10g Bacto tryptone, 5g Yeast extract, 10g NaCl, 10g Bacto agar in 1L deionised water.</td>
</tr>
</tbody>
</table>

Seven LB agar plates were pre-warmed at 37°C. A 100mL bottle of set top agar was melted and kept warm in a 42°C waterbath. Ten-fold serial dilution of the helper phage supernatant was prepared in LB to concentrations ranging from 10^{-4} to 10^{8}. 50µL of each dilution was mixed into 3mLs of melted top agar containing 300µL of overnight CJ236 culture. The top agar mixtures were poured onto the pre-warmed LB agar plates and allowed to set. Plates were incubated overnight at 37°C. Controls included top agar mixtures prepared without phage supernatant or CJ236 culture. The number of plaques were counted the next day and used to calculate the amount of plaque forming units (p.f.u.) in the supernatant (p.f.u./mL = no. of plaques x 20).

F) Transformation of CJ236 E. coli with CRE4-pET and W12-pET DNA

| LB/LB agar (+amp, +CAP) | LB/LB agar containing 100µg/mL (w/v) Ampicillin and 15µg/mL (w/v) Chloramphenicol |

50µL of electrocompetent CJ236 was transformed with 3µL of CRE4-pET and W12-pET DNA. Transformed bacteria clones were selected for amp and CAP resistance on LB agar (+amp, +CAP) plates. Five colonies from each plate were picked and used to inoculate 5mLs of LB (+amp +CAP). The cultures were grown overnight at 37°C. DNA from 3mLs of the overnight culture was prepared (see Section 2.2.1). CRE4-pET clones were digested with Pst I and Sal I in SURECUT™ buffer H (Roche Diagnostics Ltd, UK), while W12E4-pET clones were digested with BamH I and Hind II in SURECUT™ buffer B (Roche Diagnostics Ltd, UK).
G) Synthesis of single-stranded uracil-enriched DNA templates for mutagenesis

| **LB** | 10g Bacto tryptone, 5g Yeast extract, 10g NaCl in 1L deionised water. Sterilised by autoclaving. |
| **LB (+Urd)** | LB containing 100μg/mL (w/v) of Urd. |
| **LB (+amp +Urd)** | LB containing 50μg/mL (w/v) amp and 100μg/mL (w/v) Urd. |
| **Uridine (Urd) stock solution** | 25mg/L (w/v) in deionised water. |
| **40% PEG₆₀₀₀ stock solution** | 200g in 500mLs deionised water. |
| **Phage precipitation solution** | mix equal volumes of 40% (v/v) PEG₆₀₀₀ and 5M NaCl. |
| **TE solution** | 10mM Tris-HCl solution (pH8.0), 0.1mM EDTA. |

CJ236 E. coli containing the CRE4-pET or W12E4-pET plasmids were grown overnight in 9mLs of LB (+amp +Urd). 4mLs of overnight culture was used to inoculate 200mLs of LB (+ Urd) in a 2L baffle flask. The culture was incubated at 37°C on a shaker for 30mins before 10¹¹ p.f.u. of M13R408 was added. The infected culture was incubated at 37°C for approximately 9hrs, then spun at 14,000 x g in a Sorvell centrifuge (GSA rotor) for 10mins twice. The clear supernatant was transferred into fresh centrifuge tubes. 0.25x volume of the phage precipitation solution was added and the supernatant was incubated on ice for 30mins, then spun at 21,000 x g in a Sorvell centrifuge (GSA rotor) for 15mins. The phage pellets were resuspended in 5mLs of TE solution. The phage solution was transferred into a 30mL Oakridge centrifuge tube and spun at 6,000 x g in a Sorvell centrifuge (SS34 rotor) for 15mins. The phage pellet was resuspended in 2mLs of TE solutions and incubated with 2mLs of phenol at room temperature for 1hr on a rotating wheel, interspersed by 3x1min periods of vigorous vortexing. The contents were transferred into two eppendorfs and spin at 12,000 x g in a microcentrifuge for 15mins. The top liquid phase (containing the single-stranded phage DNA) was purified with phenol/chloroform/isoamylalcohol extractions (see Section 2.2.4) and the DNA was precipitated twice with 80μL/mL of (4MNaCl, 50mM EDTA) and 2.5x volume of ethanol for 2hrs at -20°C. After the last precipitation and spin, the DNA were washed with ice-cold 70% (v/v) ethanol in water and resuspended in 150μL of deionised water. The
Chapter 2 – General Methods and Materials: DNA Methods

concentration of the single-stranded DNA solution was calculated (absorbance value of 1.0 at OD_{260} = 40\mu g/mL).

H) Phosphorylation of mutant oligonucleotide primers
The mutant oligonucleotide primers were phosphorylated in a 30\mu L reaction mixture containing 360ng of oligonucleotide, 3\mu L of 10x kinase buffer, 1.5\mu L of 20mM ATP, and 1.5\mu L T4 kinase (New England BioLabs Ltd, UK). The reactions were carried out at 37°C for 40mins and stopped by a 10min incubation at 70°C.

I) Annealing of mutant primers to the single-stranded uracil-enriched DNA template

\[
\text{TM buffer} \quad - \quad 200mM \text{Tris-Cl solution (pH8.0), 50mM MgCl}_2
\]

The annealing reactions (total volume of 20\mu L) were carried out in an eppendorf containing 500ng of single-stranded uracil-enriched DNA, 36ng of phosphorylated oligonucleotide, and 2\mu L of TM buffer. The eppendorf was heated at 95°C for 10mins in a 500mL beaker containing 250mLs of water. The waterbath was allowed to cool to about 35°C, before the eppendorf was spun briefly in a microcentrifuge.

J) Synthesis of the complementary DNA strand

\[
\text{5x pol/ligase buffer} \quad - \quad 20mM \text{Tris-Cl solution (pH8.0), 2mM DTT, 10mM MgCl}_2, 0.5mM \text{of each dNTP, 1mM ATP.}
\]

The complementary DNA strand was synthesised in a 50\mu L reaction mixture containing 20\mu L of the annealed reaction mix, 10\mu L of 5x pol/ligase buffer, 21units of T7 polymerase (New England BioLabs Ltd, UK) and 15units of T4 ligase (New England BioLabs Ltd, UK). The mixture was incubated at 22°C for 10mins, then at 37°C for 90mins. DNA was purified with phenol/chloroform/isoamylalcohol extractions and
ethanol precipitated (see Section 2.2.4). The DNA pellet was washed with ice-cold 70% (v/v) ethanol in water, and resuspended in 20μL of deionised water.

**K) Bacteria transformation**

| LB or LB agar (+amp) | LB or LB agar containing 100μg/mL (w/v) Ampicillin. |

5μL of the double-stranded E4 mutant DNA was used to transform an aliquot of electrocompetent DH5α (see Section 2.1.3). The transformed bacteria were selected for ampicillin resistance on LB agar (+amp). Five transformed bacteria colonies were picked for each mutant clone and grown overnight in 5mLs of LB (+amp). 3mLs of the overnight culture was used for DNA extraction (see Section 2.2.1), and the DNA was used for sequencing (see Section 2.7).

### 2.6.2 QuikChange site-directed mutagenesis (Stratagene Europe, UK)

**A) Mutant E4 primers**

Two complementary mutant oligonucleotide primers were required for each mutagenesis reaction. In addition to the set of primers described in Section 2.6.1 A, the following primers were obtained from Oswel Research Products Ltd (UK).

<table>
<thead>
<tr>
<th>Oligonucleotide ID</th>
<th>Oligonucleotide sequence (5' -&gt; 3')</th>
<th>Primer length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV-16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16st15/6/0 P2</td>
<td>GGT GGT TGG CTA AGT GCT GCC TTA TTA TTT CAG GAG</td>
<td>36</td>
</tr>
<tr>
<td>16st20 P2</td>
<td>CGG GGT GGT TGG CTA AGT GCT GCC</td>
<td>24</td>
</tr>
<tr>
<td>16st20m P2</td>
<td>CGG 5GT 5GT TGG CTA AGT GCT GCC</td>
<td>24</td>
</tr>
<tr>
<td>16st20A P2</td>
<td>CGG GGT GGT TGG CTA AGT GCC</td>
<td>23</td>
</tr>
<tr>
<td>16st32/4 P2</td>
<td>CGG TGC CTA AGG CTA CGG CTT TGG</td>
<td>24</td>
</tr>
<tr>
<td>16st67/3 P2</td>
<td>ATG CAG TCA GGA TTG GAG CAC TGT CTA CTG AGT</td>
<td>33</td>
</tr>
<tr>
<td>16st84 P2</td>
<td>CTA TTA CAG TTT ATC CGT CCT TTG</td>
<td>24</td>
</tr>
<tr>
<td>CRPV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRst9/0 P2</td>
<td>CCC AGC CGC TAG TGA GTG CCC CTG</td>
<td>24</td>
</tr>
<tr>
<td>CRst41 P2</td>
<td>CCA CGC GTA CAC TAG GCA GAC GAA GG</td>
<td>26</td>
</tr>
<tr>
<td>CRst74 P2</td>
<td>CCG GGG ACA TAA GAC CGC CTG CTC</td>
<td>24</td>
</tr>
<tr>
<td>CRst82 P2</td>
<td>CAG AGG ACG TAG GAC GAA GAA CTA</td>
<td>24</td>
</tr>
<tr>
<td>CRst89 P2</td>
<td>CTA CGA CGG TAG GAA GAA CAC CTC</td>
<td>24</td>
</tr>
</tbody>
</table>
Chapter 2 – General Methods and Materials: DNA Methods

B) DNA mutagenesis

| 10x Reaction buffer | 100mM KCl, 100mM (NH₄)₂SO₄, 200mM Tris-HCl (pH 8.8), 20mM MgSO₄, 1% (v/v) Triton X-100, 1mg/mL (w/v) Nuclease-free bovine serum albumin. |

A 50μL reaction mix was prepared for each E4 mutagenesis. Each reaction mix contained 5μL of 10x reaction buffer, 0.5μL (5-50ng) of double-stranded DNA template (CRE4-pET or W12E4-pET), 125ng of each primer, 1μL of dNTP mix (supplied with kit), and 2.5units native Pfu DNA polymerase (Stratagene Europe, UK). Reactions were overlaid with 1 drop of mineral oil (Sigma-Aldrich Company Ltd, UK) and subjected to a thermal cycle program which consisted of 30secs at 95°C, followed by sixteen cycles at 95°C (30secs), 55°C (1min) and 68°C (14mins), and a final cooling step at 4°C.

C) Removal of the DNA template

12units of the restriction enzyme Dpn I (Stratagene Europe, UK) was added into each mutagenesis reaction and mixed thoroughly by gentle pipetting. The digest was incubated at 37°C for 1hr.

D) Transformation of E. coli

| LB or LB agar (+amp) | LB or LB agar containing 100μg/mL (w/v) Ampicillin. |

The Dpn I digested reaction was transferred into a clean eppendorf and made up to 250μL with deionised water. The DNA solution was purified with phenol/chloroform/isoamylalcohol extractions, and the DNA ethanol precipitated (see Section 2.2.4). The DNA pellets were washed with ice-cold 70% (v/v) ethanol in water, and resuspended in 20μL of deionised water. 2μL of each mutant DNA was used to transform one aliquot of electrocompetent DH5α (see Section 2.1.3). Transformed bacteria were selected for their resistance to ampicillin on LB agar (+amp).
E) E4 mutation verification

| LB or LB agar (+amp) | LB or LB agar containing 100μg/mL (w/v) Ampicillin. |

Five single transformed colonies were picked from each LB agar (+amp) plate and grown overnight in 5mLs of LB (+amp). DNA was extracted from 3mLs of overnight culture using the Wizard Plus Minipreps DNA Purification System (Promega Ltd, UK; according to the manufacturer’s protocol) and eluted in 45μL of deionised water. 4μL of DNA was run in a 0.8% agarose (+EthBr) gel (see Section 2.2.5) along with the W12E4-pET or CRE4-pET parental plasmids. The E4 mutant clones which showed a similar molecular weight as the parental plasmids were selected. 8μL of the selected clones were sequenced (section 2.7) in both direction along the entire E4 ORF. Sequencing primers are described in section 2.7.3.

2.6.3 Cloning of mutated E4 ORFs into the parental genome plasmids

HPV-16

15μg of the mutant W12E4-pET DNA clones was digested with 20units of Apa I (in SURECUT™ buffer A; Roche Diagnostics Ltd, UK) at 25°C for 1hr in a 40μL digest. Following this incubation, 2μL of SURECUT™ buffer A, 14μL of deionised water and 20units of Hind I (Roche Diagnostics Ltd, UK) were added, and the digest was incubated at 37°C for 1hr. The digest was run in a 0.8% agarose (+EtBr) gel (see Section 2.2.5) and the 1.4kb DNA fragment was purified using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK; according to the manufacturer’s protocol), and eluted in 40μL of deionised water. 15μg of pTZhW12 DNA were digested with Apa I and Hind II, and the 9.5kb DNA fragment was gel purified using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK; according to the manufacturer’s protocol), and resuspended in 90μL of deionised water. Similarly, 1.5μg of 9.5kb DNA fragment (pTZhW12 without the E4 ORF) and 7μg of the 1.4kb DNA fragment (containing the mutated E4 ORF) were used in a ligation reaction (see Section 2.3.1).
CRPV

Approximately 30 μg of each CRst-pET mutant DNA was digested with 50 units of Sac I and 50 units of Sma I in a total of 150 μL of SURECUT™ buffer A (Roche Diagnostics Ltd, UK) at 37°C for 1 hr. The DNA fragment containing the mutated E4 ORF (2.1 kb) was purified using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK; according to the manufacturer’s protocol), and eluted in 40 μL of deionised water. Similarly, 15 μg of pLACPVP DNA was digested with Sac I and Sma I and the 8.9 kb DNA fragment (pLACPVP without the E4 ORF) was gel purified and resuspended in 90 μL of deionised water. Approximately 1.5 μg of the 8.9 kb DNA fragment and 7 μg of the 2.1 kb DNA fragment were ligated with 20 units of T4 ligase in a 40 μL reaction volume (see Section 2.3.1). 1 μL of the ligation was used to transform 50 μL of electrocompetent DH5α (see Section 2.1.3).

2.6.4 Preparation of the re-circularised E4 mutant viral genome

HPV-16

50 μg of the wildtype and each E4 mutant HPV-16 genome plasmid (pTZhW12 or pTZhW12st respectively) was digested in a 100 μL reaction mixture containing 50 units of BamH I (in SURECUT™ buffer B; Roche Diagnostics Ltd, UK) at 37°C for 1 hr. The digest was ran in a 0.8% agarose (+EthBr) gel, and the linearised HPV-16 genome DNA fragment was purified using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK; according to the manufacturer’s protocol), and eluted in 50 μL of deionised water. 45 μL of the DNA was in a ligation mix containing 40 units of T4 ligase (New England BioLabs Ltd, UK) in a total volume of 200 μL. The ligation was carried out at 14°C overnight.
200μg of pLACRPV and pLACRst plasmid DNA was digested in a 200μL reaction mixture containing 80 units of Sal I (in SURECUT™ buffer H; Roche Diagnostics Ltd, UK) at 37°C for 2hrs. The linearised CRPV genome was excised from a 0.8% agarose (+EthBr) gel, and electroeluted from the gel fragment as described in section 2.3.2. The precipitated DNA was washed with ice-cold 70% (v/v) ethanol in water and resuspended in 150μL of deionised water. Ligation of the CRPV genome was carried out overnight in a 14°C waterbath with 20 units of T4 ligase (New England BioLabs Ltd, UK). The ligation reaction was purified by phenol/chloroform/isoamylalcohol extraction, and ethanol precipitated (see Section 2.2.4). The DNA pellet was washed with ice-cold 70% (v/v) ethanol in water, resuspended in 100μL of deionised water.

2.7 DNA Sequencing

2.7.1 Thermal cycle sequencing

| 3M Sodium acetate | 40.81 of Sodium acetate,3H₂O in 100mL deionised water. Adjust pH to 5.2 with Glacial acetic acid. |
| DNA Loading Buffer | 50mg/mL (w/v) Blue dextran, 25mM EDTA, 5:1 Deionised formamide; (Blue dextran+EDTA) in deionised water. |

Sequencing reactions were prepared according to the Taq DyeDeoxy Terminator Cycle Sequencing Kit (Perkin Elmer, USA) protocol. Each reaction contained approximately 0.5μg of double stranded DNA template or 30ng of PCR product, 10pmole of primer, 8μL of terminator ready reaction mix and made up to 20μL with deionised water. Cycle sequencing involved thirty cycles of DNA denaturation at 96°C for 30secs, annealing at 50°C for 15secs and elongation at 60°C for 4mins. The extension product was transferred into a 1.5mL eppendorf and ethanol precipitated with 2μL of 3M sodium acetate pH 5.2 and 50μL of 95% (v/v) ethanol at -20°C for 15mins. The DNA was spun in a microcentrifuge at 12,000 x g for 10mins, and the pellet was
washed with ice-cold 70% (v/v) ice-cold ethanol, air dried and resuspended in 4μL of DNA loading buffer.

2.7.2 Automated DNA sequencing

<table>
<thead>
<tr>
<th>Ammonium persulphate (APS)</th>
<th>10% (w/v) APS in deionised water.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10xTBE</td>
<td>5.4g Trizma base, 27.5g Boric acid, 20mL 0.5M EDTA (pH 8) in deionised water.</td>
</tr>
</tbody>
</table>

The sequencing gel plates were cleaned extensively with double distilled water and air dried on racks. 50mLs of sequencing gel solution was made with 18g of AlanaR grade Urea (BDH Laboratory Supplies, UK), 5mLs of 10x TBE, 5mLs of Long Range Gel Solution (BioWhittaker Molecular Applications, USA), 250μL ammonium persulphate and approximately 25mLs of double distilled water. The gel solution was filtered through a 0.2μm cellulose nitrate filter membrane (Whatman International Ltd, UK) and degased using a vacuum pump. The sequencing gel plates and apparatus were assembled. 25μL of Temed (N,N,N',N'-Tetramethylethylenediamine; Sigma-Aldrich Company Ltd, UK) was added into the gel solution which was immediately poured between the plates using a 50mL syringe. Once the gel was set, the plates were loaded onto the ABI Prism 377 DNA Sequencer (Perkin Elmer, USA). The gel tanks were filled with TBE buffer, the 36-well comb was inserted at the top of the gel, and the wells washed with TBE buffer. 2μL of the sequenced product was loaded into each well and the gel was run on a pre-set program for 7hrs. The sequencing data were tracked and generated by the ABI Prism AutoAssembler-DNA Sequence Assembly Software (PE Applied Biosystems, Perkin Elmer, USA). Sequence data were analysed using the LaserGene Software for Macintosh (DNASTAR Incorporated, USA) or the ABI Prism DNA Sequencing Analysis Software (PE Applied Biosystems, Perkin Elmer, USA).
### 2.7.3 HPV-16 and CRPV sequencing primers

<table>
<thead>
<tr>
<th>Oligonucleotide ID</th>
<th>Oligonucleotide sequence (5' -&gt; 3')</th>
<th>Primer length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV-16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16E4C1</td>
<td>GTT CAT GAA GGA ATA CGA AC</td>
<td>20</td>
</tr>
<tr>
<td>16E4C2</td>
<td>CGC GAC CCA TAC CAA AGC CG</td>
<td>20</td>
</tr>
<tr>
<td>16E4C3</td>
<td>GAC GTT AGC CTT GAA GTG TAT</td>
<td>21</td>
</tr>
<tr>
<td>16E4C4</td>
<td>GGC GTG CTT TTT GCT TTT C</td>
<td>19</td>
</tr>
<tr>
<td>16E4C5</td>
<td>TTA CAA TAT GGA AGT ATG GGT G</td>
<td>22</td>
</tr>
<tr>
<td>16E4NC1</td>
<td>CAA CAA CTT AGC TGT GTG GC</td>
<td>20</td>
</tr>
<tr>
<td>16E4NC2</td>
<td>CCA TGT AGA CGA CAC TGC AG</td>
<td>20</td>
</tr>
<tr>
<td>16E4NC3</td>
<td>TTG CAG AAC GTG TGT GTC GC</td>
<td>20</td>
</tr>
<tr>
<td>16E4NC4</td>
<td>GGT CAG TGA AAG TGG GAT TAT T</td>
<td>22</td>
</tr>
<tr>
<td>CRPV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRE4C1</td>
<td>GTG CAA AGC AAG CCA TAG A</td>
<td>19</td>
</tr>
<tr>
<td>CRE4C2</td>
<td>TGC TGC TAA TGG ACA CTA TG</td>
<td>20</td>
</tr>
<tr>
<td>CRE4C3</td>
<td>AAG CCG TTT CAG TTG TCT GTC</td>
<td>21</td>
</tr>
<tr>
<td>CRE4C4</td>
<td>GCG CTC GTG GAT GAT GCT ACA</td>
<td>21</td>
</tr>
<tr>
<td>CRE4C5</td>
<td>TAC GAC GTG TGG AAG AAC AC</td>
<td>20</td>
</tr>
<tr>
<td>CRE4C6</td>
<td>TTC CCA TCA GCA GCT TTT GTA</td>
<td>21</td>
</tr>
<tr>
<td>CRE4NC1</td>
<td>CAG TGT AGC CGT GTG TTT G</td>
<td>19</td>
</tr>
<tr>
<td>CRE4NC2</td>
<td>TGC AGT CGA ATA GTG AGG AG</td>
<td>20</td>
</tr>
<tr>
<td>CRE4NC3</td>
<td>AGG CCT TCA GCT TCA CCA</td>
<td>18</td>
</tr>
<tr>
<td>CRE4NC4</td>
<td>CTC GTC TTC CGG GTC GCT TAG</td>
<td>21</td>
</tr>
<tr>
<td>CRE4NC5</td>
<td>GTT GTT CCC TCT GTC ACC ATC</td>
<td>21</td>
</tr>
<tr>
<td>CRE4NC6</td>
<td>ACA CCA AGG CTG CCA TAC TG</td>
<td>20</td>
</tr>
<tr>
<td>CRE4NC7</td>
<td>CCT GTA ACA TTT GTG CTT CAT T</td>
<td>22</td>
</tr>
</tbody>
</table>
2.8 Glutathione Sepharose-Transferase (GST) and Maltose Binding Protein (MBP) E1^E4 fusion proteins

Prokaryotic gene fusion vectors pGex-4T-1 (Amersham Pharmacia Biotech, UK) and pMAL-c2 (New England BioLabs Ltd, UK) were used for the expression of GST and MBP fusion proteins respectively.

2.8.1 Primers

A set of forward (sense) and reverse (anti-sense) oligonucleotide primers were prepared for the cloning of each protein gene sequence. Additional nucleotide sequences were added to the 5' end of the forward primer to create a BamH I (GAATTC) and an EcoR I (GAATTC) recognition sites (see Figure 4.1). Similarly, additional nucleotide sequences were added to the 5' end of the reverse primer to create a Sma I (ACCGG) and a Sal I (GTCGAC) recognition sites. The reverse primer also contained an additional termination codon (CTA; antisense sequence) between the Sal I recognition site and the end of the E4 gene sequence (see Figure 4.1). All primers were obtained from Oswel Research Products Ltd (UK).

2.8.2 Preparation of the E1^E4 gene fragments

The E1^E4 ORF was amplified by PCR (see Section 2.4), then purified using the Wizard PCR Preps DNA Purification System (Promega Ltd, UK), according to the manufacturer's protocol. The DNA was eluted in 50μL of deionised water. 30μL of the PCR product was digested with 15units each of EcoR I and Sal I in SURECUT™ buffer H (Roche Diagnostics Ltd, UK) at 37°C for 2hrs, then purified from a 2% agarose (+EthBr) gel using the QIAquick Gel Extraction Kit (Qiagen Ltd, UK) into 50μL of deionised water.

25μg of the fusion protein expression vector (pGex4T-1 or pMAL-c2) was digested with 25units each of EcoR I and Sal I in SURECUT™ buffer H at 37°C for 2hrs, followed by a 2hr
incubation with 2.5 units of calf intestine alkaline phosphatase (Roche Diagnostics Ltd, UK) at 37°C. The linearised vector was purified from a 1% agarose (+EthBr) gel using the QiAquick Gel Extraction Kit (Qiagen Ltd, UK), and eluted in 50μL of deionised water.

10μL of the EcoR I/Sal I digested, phosphatase-treated vector DNA was incubated with 20μL of the EcoR I/Sal I digested E1^E4 PCR product and 0.5μL of T4 ligase (New England BioLabs Ltd, UK) overnight in a 14°C waterbath. The ligation was purified with phenol/chloroform/isoamyl alcohol extraction, then ethanol precipitated (see Section 2.2.4). The DNA pellet was washed with ice cold 70% (v/v) ethanol in water, air-dried and resuspended in 10μL deionised water.

2.8.3 Transformation of E.coli

**LB or LB agar (+amp)** - LB or LB agar containing 100μg/mL (w/v) Ampicillin.

2μL of the ligated DNA was used to transform an aliquot of DH5α by electroporation (see 2.1.3). Transformed DH5α cells were selected for ampicillin resistance on LB agar (+amp) plates overnight at 37°C. Single colonies were picked from the agar plate the next day and grown overnight in 5mLs of LB (+amp) at 37°C.

2.8.4 Verification of the DNA clones

DNA was extracted from 3mLs of the overnight culture using the Wizard Plus Minipreps DNA Purification System (Promega Ltd, UK) and eluted in 50μL of deionised water. 5μL of each DNA clone was digested with 5units each of EcoR I and Sal I in SURECUT™ buffer H (Roche Diagnostics Ltd, UK) at 37°C for 1hr. The digest was checked on a 1.2% agarose (+EthBr) gel for the presence of the insert fragment. 4μL of each positive clone was used for sequencing (see Section 2.7). Each clone was sequenced in both directions. Primers pGexP1 (CGTATTGAAGCTATCCCACA) and pMalP1 (GATGTCCGCTTTCTGGTATGC) were used to sequence the coding strand of the pGex4T-1 and pMalc clones respectively. The same reverse
primers (see Figure 4.1) were used to sequence the non-coding strand of the respective pGex4T-1 and pMalc-2 clones.

2.8.5 Expression of the GST- and MBP-E1^E4 fusion proteins

| Isopropyl-1-thio-β-D-galactopyranoside (IPTG) | 100mM IPTG working stock solution. |
| NET-N     | 100mM NaCl, 1mM EDTA, 20mM Tris-HCl solution (pH 8), 0.5% (v/v) NP40. |
| GST elution buffer | 30mM Reduced glutathione, 50mM Tris-HCl solution (pH8.0), 100mM NaCl. |
| MBP elution buffer | 10mM Maltose, 50mM Tris-HCl solution (pH8.0). |

A 5mL LB (+amp) culture was set up overnight, and used to inoculate 500mLs of LB (+amp) the next day. The large batch of culture was incubated on a shaker at 37°C for several hours until the OD₆₀₀ was approximately 0.4 (about 3hrs). 0.25mLs of IPTG was added and the culture was incubated for another 3hrs. The culture were transferred into two 250mL centrifuge tubes (Corning Incorporated, USA) and spun at 3200 x g for 10mins in a pre-chilled Beckman J6 centrifuge. The bacteria pellet was chilled on ice for 5mins, resuspended in 10mLs of ice-cold NET-N, then transferred into a pre-chilled 30mL glass beaker. The bacteria suspension was sonicated for 3x1min with 1min cooling intervals (on ice). The lysate was transferred into a chilled 50mL falcon tube and centrifuged at 4500 x g (Beckman J6-HC centrifuge) for 15mins. The clear supernatant was transferred into a pre-chilled 50mL falcon tube.

To purify GST fusion proteins, 400µL of chilled Glutathione Sephrose 4B beads (Sigma-Aldrich Company Ltd, UK) was washed three times with fresh cold NET-N prior to addition to the lysate solution. The beads were mixed on a rotator at 4°C for 25mins. The mixture was spun in a bench top centrifuge at 500 x g for 3mins, and the supernatant carefully removed. The beads were washed once with 20mLs NET-N, transferred into a clean chilled eppendorf tube and washed three times with 1.5mLs of cold NET-N. 1mL of chilled GST elution buffer was added to the beads, mixed at room temperature for 5mins, then spun briefly in a microcentrifuge at 1,500 x g. The supernatant was put through a poly-prep chromatography column (Bio-Rad Laboratories, USA), and collected in a chilled eppendorf. A total of 10x1mL of GST elution
fractions were collected in separate eppendorfs. The protein concentration of each fraction was measured at OD_{280}. Fractions with an OD_{280} value of >0.2 were pooled together.

Similarly, Amylose Resin (New England BioLabs Ltd, UK) and MBP elution buffer were used in replacement of the glutathione sephrose beads and GST elution buffer for the purification of MBP-fusion proteins.

2.9 Peptide conjugation

2.9.1 Keyhole Limpet Hemocyanin (KLH)- peptide conjugation

A) Activation of the KLH carrier with sulfo-MBS (quantities for making two peptide conjugation reactions)

<table>
<thead>
<tr>
<th>Keyhole limpet hemocyanin (KLH)</th>
<th>Reconstitute with 2mL degased deionised water to make a 10mg/mL (w/v) solution containing 0.083M sodium phosphate, 0.9M NaCl, pH 7.2 with stabiliser.</th>
</tr>
</thead>
<tbody>
<tr>
<td>m-Maleimidobenzoyl-N-hydroxysuccinimide ester (Sulfo-MBS)</td>
<td>Made up as a 2mg/mL (w/v) solution in conjugation buffer. Used immediately.</td>
</tr>
<tr>
<td>Conjugation buffer</td>
<td>0.083M Sodium phosphate, 0.9M NaCl, 0.1M EDTA, pH 7.2.</td>
</tr>
<tr>
<td>Degassed deionised water</td>
<td>10mM Maltose, 50mM Tris-HCl solution (pH8.0).</td>
</tr>
</tbody>
</table>

20mLs of deionised water was degassed by running a stream nitrogen gas through it for about 5mins. A 20mg/mL (w/v) KLH solution (in conjugation buffer) was made by adding 2mLs of degased deionised water to a vial of lyophilised Imject KLH powder (Perbio Science UK Ltd, UK). 2mg of sulfo-MBS (Perbio Science UK Ltd, UK) was measured on a piece of foil and dissolved in 1mL of conjugation buffer in an eppendorf. The sulfo-MBS solution was added immediately to the KLH solution (murky solution), and the mixture was stirred with a metallic stirrer at room temperature for 1hr. In the meantime, a PD-10 column (Amersham Pharmacia Biotech Ltd, UK) was equilibrated with 25mLs of conjugation buffer. The KLH-MBS solution was put through the equilibrated PD-10 column and the elute was collected in a sterile eppendorf. The first fraction (of
approximately 1mL) was used as a blank for plotting a protein elution profile. The protein was eluted in 10x1mL fractions of conjugation buffer into fresh eppendorfs. The protein concentration of each fraction was measured at OD$_{280}$. These were used to plot a protein elution curve. The fractions that made the first peak of protein elution curve were pooled together, as they contained the MBS activated KLH molecules.

B) KLH-peptide conjugation

8mg of each peptide was weighed out on a piece of foil and resuspended in 1mL of conjugation buffer in a glass bijou. A small amount of Dimethylformamide (DMF) was used to dissolve peptides with a high percentage of hydrophobic residues, prior to the addition of the conjugation buffer. It was ensured that the volume of DMF did not exceed 10% (v/v) of the final conjugation volume. 50μL of the unreacted peptide solution was stored at 4°C (used for the Cysteine assay). The remaining peptide solution was added to the activated KLH. The conjugation reaction was stirred overnight at 4°C, and the percentage of coupling was determined by the Cysteine assay the next day.

C) Cysteine assay (described by Sigma-Aldrich Company Ltd, UK)

| **5,5-Dithio-bis-2-nitrobenzoic acid (DTNB, Ellman’s reagent)** | Dissolve 5mg in 5mLs of DTNB buffer to make 1mg/mL (w/v) Ellman’s reagent. |
| **DTNB buffer** | 0.1M Sodium phosphate buffer, pH 8.0. |
| **L-Cysteine hydrochloride solution** | Made up in deionised water to make up the desired standard concentrations. |

I. Cysteine standard assay

Fresh cysteine solutions were made up for each cysteine standard assay. Serial dilutions of cysteine solutions ranging from 0.4-0.04mg/mL (w/v) were prepared in deionised water. 50μL of each prediluted standard was placed in individual glass test tubes. 50μL of deionised water was used as blank. 0.1mL of deionised water, 0.75mL
of DTNB buffer and 0.1 mL of Ellman's reagent were added immediately to each test tube. The contents were mixed and their absorbances measured at OD\textsubscript{412}. The absorbance values were plotted against the final cysteine concentration (2-20\(\mu\)g/mL (w/v)) on graph paper.

**II. Cysteine assay to determine total amount of unreacted cysteine before conjugation**

This was done on the same day when the conjugation reaction was carried out. A solution containing the same peptide concentration used in the conjugation reaction was made up by diluting the unreacted peptide solution in deionised water. 50\(\mu\)L of this diluted solution was added to a clean test tube. 50\(\mu\)L of DTNB buffer was used as blank. 0.1 mL of deionised water, 0.75 mL of DTNB buffer and 0.1 mL of Ellman's reagent were added immediately to each test tube. The assay tubes were mixed and their absorbance values were measured at OD\textsubscript{412}.

**III. Cysteine assay to determine peptide coupling efficiency**

This was done the following day. 50\(\mu\)L of the conjugated reaction was used. 50\(\mu\)L of DTNB was used as blank. 0.1 mL of deionised water, 0.75 mL of DTNB buffer and 0.1 mL of Ellman's reagent were added immediately to each test tube. The contents were mixed and their absorbance values measured at OD\textsubscript{412}. Using the cysteine standard curve, the absorbance values of the diluted unreacted peptide (cys total) and conjugated peptide mixture (cys free) were used to determine the amount of cysteine (\(\mu\)g/mL) present in the peptide solution before and after the conjugation reaction. The following equations were used to calculate the percentage of peptide coupling.

\[
\begin{align*}
\text{Cys total (}\mu\text{g/mL)} &= \mu\text{g/mL} \times \text{(dilution factor)} \\
\text{Cys free (}\mu\text{g/mL)} &= \mu\text{g/mL} \times \text{(dilution factor)} \\
\text{Percentage of coupling} &= \frac{(\text{cys total}) - (\text{cys free})}{\text{cys total}} \times 100
\end{align*}
\]
2.9.2 Bovine Serum Albumin (BSA)- peptide conjugation

The same buffers and equipment were used here as KLH-peptide conjugation described above. 1mg of peptide was dissolved in 1mL of conjugation buffer, of which 50μL was taken out to determine the total amount of cysteine before the conjugation reaction. 200μL of degassed deionised water was added to dissolve the lyophilised maleimide activated BSA (Perbio Science UK Ltd, UK). The peptide solution was added immediately into the BSA solution and the conjugation reaction was stirred overnight at 4°C. The total cysteine concentration was determined on the same day as described before. The percentage of coupling was calculated as described in section 2.9.1.C.

2.10 Protein analysis

2.10.1 Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

| Sodium dodecyl sulphate (SDS) solution | 10% (w/v) SDS in deionised water. |
| 1M Tris-HCl solution (pH6.8) | 12.1g of Trizma base in 100mLs. Adjust pH to 6.8 with concentrated HCl. |
| 1.5M Tris-HCl solution (pH8.8) | 18.15g of Trizma base in 100mLs. Adjust pH to 8.8 with concentrated HCl. |
| APS | 10% (w/v) APS in deionised water. |
| 5xSample loading buffer | 9M Urea, 4% (v/v) NP-40, 2% (v/v) β-Mercaptoethanol. |
| Protein gel running buffer | 3.028g Trizma base, 14.413g Glycine, 1g SDS in 1L. |

Mini SDS-polyacrylamide running and stacking gel solutions were prepared as shown in Appendix 3. Mini protein gels were cast in Protean II mini-gel apparatus (Bio-Rad Laboratories, USA). Temed (Sigma-Aldrich Company Ltd, UK) was added just before each gel solution was used. Protein samples were prepared in sample loading buffer, denatured at 95°C for 5mins before they were loaded into the sample wells. The samples were separated by electrophoresis in protein gel running buffer at a constant current of 60mA for about 1hr. Rainbow Molecular Markers (Amersham Pharmacia Biotech Ltd, UK) were used to determine the sizes of the separated protein bands.
2.10.2 Staining protein gels with coomassie blue

| **Coomassie blue stain** | 0.25% (w/v) Coomassie brilliant blue R-250, 50% (v/v) Methanol, 10% (v/v) Acetic acid. |
| **Destain solution** | 5% (v/v) Methanol, 7.5% (v/v) Acetic acid. |

SDS-polyacrylamide gels were soaked in the commassie quick stain for 5-30mins with shaking. The gel was soaked in destain solution overnight on a shaker. Gels were dried onto a piece of absorbent 3MM paper (Whatman International Ltd, UK) on a heated vacuum gel dryer.

2.10.3 Semi-dry protein-membrane transfer (Western blot)

| **Bjerrum & Schafer-Nielsen transfer buffer (pH 9.2)** | 5.82g Trizma base, 2.93g Glycine, 20% (v/v) Methanol, 3.75mL 10% (w/v) SDS solution in 1L (do not add acid/base to adjust pH). |

Gels were equilibrated in the transfer buffer for 10mins. A sheet of the Immobolin-P membrane (Millipore Corporation, USA) was cut to the size of each gel and soaked in the transfer buffer for about 10mins. Six sheets of thin 3MM filter paper (Whatman International Ltd, UK) of approximately the same dimensions as the gel were soaked in the transfer buffer. Three sheets of filter paper were placed onto the platinum anode of the semi-dry electrophoretic transfer cell (Bio-Rad Laboratories, USA). The nitrocellulose membrane was placed on top of the filter paper, followed by the equilibrated protein gel, and lastly the stack was topped with three pieces of filter paper. Air bubbles were removed by rolling a 10mL pipette across the stack. The cathode and safety cover were assembled and the protein transfer was carried out at 15V for 15-30mins.

2.10.4 Detection of proteins on western blot

| **Phosphate Buffered Saline (PBS)** | 10g NaCl, 0.25g KCl, 1.437 NaH₂PO₄, 0.25g KH₂PO₄ in 1L. |
| **5% MARVEL™/PBS** | 5% (w/v) MARVEL™ in PBS. |
| **Wash buffer** | 0.1% (v/v) TWEEN20™ in PBS. |
Chapter 2 – General Methods and Materials: Protein Methods

The protein membrane was blocked in 10% MARVEL™/PBS either overnight at 4°C or for 1 hr at room temperature, then incubated with the primary antibody (prepared in 5% MARVEL™/PBS) in a sealed plastic bag at 37°C for 1 hr with shaking. The protein membrane was washed in wash buffer for 3x5 mins on a shaker. Horse-radish peroxidase (HRP) linked species specific secondary antibody was prepared in 5% MARVEL™/PBS. The protein membrane was incubated with the secondary antibody and washed as described in the previous steps. HRP activity was detected using the Enhanced Chemiluminescence (ECL) western blotting detection reagents kit (Amersham Pharmacia Biotech Ltd, UK; according to the manufacturer’s protocol) and developed on the Kodak X-OMAT AR autoradiography film (Eastman Kodak Company, USA) in a dark room. Alternatively, fast DAB tablets (Sigma-Aldrich Company Ltd, UK) were used to detect HRP enzyme activity, according to the manufacturer’s protocol.

2.11 Other protein methods

2.11.1 Determining protein concentration using the Bio-Rad protein assay dye reagent

Protein concentrations were measured using the Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad Laboratories, USA) according to the manufacturer’s protocol. The Microassay procedure was used to measure protein concentrations ranging from 1-20 μg or ≤25 μg/mL (w/v). BSA standards ranging from 1-25 μg/mL (w/v) were used.

2.11.2 Concentrating protein solutions with buffer exchange

The Ultrafree-15 Centrifugal Filter Device (Millipore Corporation, USA) was used to concentrate as well as to change the buffering solution in protein samples. The filter device was used according to the manufacturer’s recommendations.
2.11.3 Protein dialysis

<table>
<thead>
<tr>
<th>Treatment buffer 1</th>
<th>2% (w/v) Sodium carbonate, 1mM EDTA solution.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment buffer 2</td>
<td>1mM EDTA solution.</td>
</tr>
</tbody>
</table>

Dialysis tubings were cut into lengths of about 10cm, boiled for 10mins in treatment buffer 1, then washed once in deionised water. They were boiled in treatment buffer 2 for 10mins, then left to cool to room temperature, before they were stored at 4°C.
ANTIBODIES AND TISSUE PROCESSING METHODS

2.12 Preparation of specific antibodies

2.12.1 Immunisation of rats and rabbits

Rats and rabbits were used for the generation of E1^E4 specific antibodies (see Chapter 4). The immunogens used were GST and MBP E1^E4 fusion proteins, as well as KLH conjugated peptides. All immunogens were prepared in PBS.

A) Immunisation of rats

Six adult rats were used per immunogen (see Chapter 4.3). Each immunogen was prepared in 100μL aliquots containing 50μg of protein. Adjuvants were used during the inoculation at a ratio of 1:1, i.e. 100μL of adjuvant was added to each 100μL immunogen aliquot. Freund's complete adjuvant was used in the first inoculation, and Freund's Incomplete adjuvant was used in subsequent inoculations. Each rat was inoculated subcutaneously at four sites with 50μL of the immunogen:adjuvant mix. Following the first inoculation, boosters were given every fourteen days afterwards. Pre-immune sera sample was collected from each rat before the first inoculation. Test bleed samples were taken every seven days after the third and subsequent inoculations.

An alternative immunisation protocol was used for the MBP-CRE1^E4 immunogen (see below for preparation of this immunogen). The same number of rats and amount of immunogen:adjuvant were used as described above. Instead of a fourteen-day incubation period between boosters, inoculations were delivered every twenty-one days after the first inoculation. Test bleeds were collected every ten days after the third and subsequent inoculations.
B) Immunisation of rabbits

Two rabbits (2.5-3.5kg) were used for per immunogen (see Chapter 4.3). The immunoogen was prepared in 0.5mL aliquots containing 200μg of protein. Adjuvants were used during the inoculation at a ratio of 1:1, i.e. 0.5mL of adjuvant was added to each 0.5mL immunogen aliquot. Freund's complete adjuvant was used in the first inoculation, and Freund's Incomplete adjuvant was used in subsequent inoculations. Each rabbit was injected subcutaneously at four sites with 0.25mL of the immunogen:adjuvant mix. Following the first inoculation, boosters were given every fourteen days afterwards. Pre-immune sera sample was collected from each rabbit before the first inoculation and test bleed samples were taken every seven days after the third and subsequent inoculations.

2.12.2 Preparation of the MBP-CRPV E1^E4 immunogen

The protein sample was diluted to a concentration of 0.5-2mg/mL (w/v) in PBS and divided into two halves. One half was heated at 80°C in a heating block for 10mins, then cooled to room temperature. This was mixed with the other half of the protein sample containing native MBP-CRPV E1^E4 protein. This mixture was used to immunise rats (as described above).

2.12.3 Antibody titre and specificity tests

Specificity of the polyclonal E1^E4 antisera were tested on western blots and in Enzyme-Linked Immuno-Sorbent Assays (ELISAs).

A) Western blots

<table>
<thead>
<tr>
<th>Wash buffer</th>
<th>0.1% (v/v) TWEEN20™ in PBS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5% MARVEL™/PBS</td>
<td>5% (w/v) MARVEL™ in PBS.</td>
</tr>
</tbody>
</table>

100ng of protein was loaded in each lane of a mini SDS-PAGE gel and separated by electrophoresis (see Section 2.10.1). The proteins were transferred onto nitrocellulose
membranes by Semi-dry western blotting (see Section 2.10.3). The protein blots were blocked with 10% MARVEL™/PBS at 37°C for 1hr, then incubated with 1:250 dilution of the pre-immune sera and test bleed antisera in 5% MARVEL™/PBS at 37°C for 1hr. The blots were washed 3x5mins in wash buffer on a shaker at room temperature, and incubated with HRP labelled rat (Sigma-Aldrich Company Ltd, UK) or rabbit (Sigma-Aldrich Company Ltd, UK) specific secondary antibodies (1:1000 dilution) in 5% MARVEL™/PBS at room temperature for 1hr. Unbound antibodies were washed as described before and HRP activity was detected with Fast-DAB tablets (Sigma-Aldrich Company Ltd, UK) according to the manufacturer’s protocol.

B) Enzyme-Linked Immuno-Sorbent Assay (ELISA)

**ELISA coating buffer** - 0.05M Carbonate-bocarbonate buffer (pH9.6). 1.59g Sodium carbonate, 2.93g Sodium hydrogen carbonate per litre of deionised water.

**PBS** - 10g NaCl, 0.25g KCl, 1.437 NaH₂PO₄, 0.25g KH₂PO₄ in 1L.

**2% or 4% MARVEL™/PBS** - 2% or 4% (w/v) MARVEL™ in PBS.

**Wash buffer** - 0.1% (v/v) TWEEN20™ in PBS.

ELISAs were also used to determine the specificity and titre of the polyclonal antisera. Protein samples were prepared in ELISA coating buffer. ELISA plates were coated at 4°C overnight with 500ng of protein/well. Unbound proteins were rinsed away with 3x PBS washes. Each well was blocked with 2% MARVEL™/PBS for 1hr at 37°C. The blocking buffer was poured away and 50μL of 4% MARVEL™/PBS was added to each well. 50μL of the pre-diluted (1:10 dilution in PBS) antisera was added into the first column and mixed. 50μL of antisera from the first column was transferred into the second well, and mixed. Similar doubling dilutions were carried out across the plate until the second column from the end of the row and the last 50μL of antisera was discarded. 50μL of PBS was added into the last column as a negative control. The plate was incubated at 37°C for 1hr, then washed 3x with wash buffer. 100μL of the HRP labelled rat and rabbit specific antibody (1:2000 dilution in 2% MARVEL™/PBS) was added and
left at 37°C for 1hr. The plate was washed 3x with wash buffer, followed by 2x with PBS. Tetramethylbenzidine Dihydrochloride (TMB) substrate solution (Sigma-Aldrich Company Ltd, UK) was prepared according to the manufacturer's protocol. 100µL of the substrate solution was added into each well. The enzyme reaction was stopped shortly with the addition of 50µL of 1M sulphuric acid. The wells were read in an ELISA plate spectrophotometer at an OD$_{450}$.

### 2.13 Tissue preparations and histology

#### 2.13.1 Sectioning of paraffin-embedded formalin-fixed tissue blocks

All paraffin-embedded formalin-fixed tissue blocks were sectioned in the Histology Department, at the National Institute for Medical Research. Tissue blocks were sectioned to 5µm thickness and put onto electrostatically charged microscopic slides (BDH Laboratory Supplies, UK).

#### 2.13.2 Hemotoxylin and eosin staining

The staining of tissue sections with hemotoxylin and eosin counterstains were carried out by the Histology Department, at the National Institute for Medical Research.
2.14 Immuno-histochemistry

2.14.1 Immuno-detection on formalin-fixed paraffin-embedded tissue sections

<table>
<thead>
<tr>
<th>Blocking buffer</th>
<th>10% (v/v) normal goat serum in PBS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wash buffer</td>
<td>0.05% (v/v) TWEEN20&lt;sup&gt;TM&lt;/sup&gt; in PBS.</td>
</tr>
<tr>
<td>4',6-Diamidino-2-phenylindole dihydrochloride (DAPI)</td>
<td>Working concentration 100μg/mL (w/v), use 1/200 (Sigma-Aldrich Company Ltd, UK).</td>
</tr>
<tr>
<td>Popidium iodide (PI)</td>
<td>Working solution 100μg/mL (w/v), use 1/200 (Sigma-Aldrich Company Ltd, UK).</td>
</tr>
<tr>
<td>Sytox green (SG)</td>
<td>Working concentration 25μM, use 1/100 (Molecular Probes Europe BV, Netherlands).</td>
</tr>
</tbody>
</table>

The tissue section was dewaxed in Xylene (1x10mins, 1x5mins; BDH Laboratory Supplies, UK), then hydrated through graded alcohol solutions (100% ethanol for 2x3mins, then 2mins each in 90%, 80%, 50%, 30% (v/v) ethanol in deionised water) and PBS for 2x5mins. The perimeter around the tissue section was circled with an ImmEdge pen (Vector Laboratories Incorporated, USA). The tissue section was incubated with blocking buffer for 1hr at room temperature, then primary antibodies at room temperature for at least 1hr. Unbound antibodies were washed with wash buffer for 3x5mins on a shaker. The secondary antibody mixture containing fluorophore-labelled (fluorescein or texas red; Sigma-Aldrich Company Ltd, UK) or Alexa-labelled (Molecular Probes Europe BV, Netherlands) species specific antibodies and a nucleic acid dye (DAPI, PI or SG) was prepared in blocking buffer, and used at room temperature for 1hr. The tissue section was washed 3x5mins in wash buffer on a shaker at room temperature. The section was rinsed with PBS, then mounted with one drop of citifluo (Agar Scientific, UK), and viewed by fluorescent microscopy. Alternatively, HRP-labelled species specific secondary antibodies were used. The HRP substrate 3',3'-diaminobenzidine tetrahydrochloride (DAB; Sigma-Aldrich Company Ltd, UK; prepared according to the manufacturer’s instructions) was used to detect enzyme activity. The brown reaction product was viewed by light microscopy. All antibody incubations were carried out in a dark humid box.
2.14.2 Antigen retrieval methods

To detect masked or hidden antigen epitopes in formalin-fixed tissue sections, it was sometimes necessary to treat the sections with one of the following antigen retrieval methods. This step was usually carried out before the addition of the blocking buffer in the immuno-detection protocol (section 2.14.1).

A) Antigen retrieval buffers

<table>
<thead>
<tr>
<th>10mM sodium citrate buffer (pH6.0)</th>
<th>approximately 43mLs of 0.1M Sodium citrate, 7mLs of 0.1M Citric acid in 500mLs buffer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>10mM sodium citrate buffer with EDTA (pH8.0)</td>
<td>add approximately 45mLs of 0.1M Sodium citrate and 1mL of 0.5M EDTA (pH8.0) in 450mLs of deionised water. Adjust to pH8.0 with 0.1M Citric acid.</td>
</tr>
</tbody>
</table>

B) Microwave denaturation

Slides were placed in a metal rack and pre-soaked in 500mLs of antigen retrieval buffer in a plastic beaker at room temperature for 5mins. The beaker was covered with two sheets of paper towels, and the towels were punctured several times to allow buffer evaporation. The level of buffer was marked on the beaker and the slides were cooked in a microwave (800W) for 5mins at high power, then 2x5mins at mid-high power. The buffer level was topped with deionised water to the original mark. The slides were left to cool slowly to room temperature before they were washed in PBS.

C) Pressure cooking

Slides were placed in a metal rack and pre-soaked in the antigen retrieval buffer for 5mins at room temperature. In the meantime, 1.5L of buffer was heated in a pressure cooker on a portable gas stove. Once the buffer was boiled, the metal rack was placed in the cooker, which was sealed and allowed to reach full pressure. At full pressure, the heat was reduced and the sections were cooked for 3-8mins (at full pressure), after which the pressure cooker was placed immediately into a sink of cold water. The
pressure from the cooker was released and the slides were left to cool for at least 15mins, then washed in PBS for 5mins.

D) Microwave denaturation and protease digestion

| Trypsin solution | 0.1% (w/v) Trypsin, 0.1% (w/v) Calcium chloride, 20mM Tris-HCl solution (pH 7.8). |

Slides were pre-soaked in 500mLs of citrate buffer pH 6.0 for 5mins in a plastic beaker as described above. The slides were microwaved (in a 800W microwave) at high power for 5mins, then at mid-high power for 5mins. After which they were immediately washed in pre-warmed (37°C) deionised water for 5mins. The sections were digested with pre-warmed (37°C) trypsin solution for approximately 1min in a humid box, then washed 2x5mins in deionised water on a shaker.

2.14.3 Signal amplification systems

A) ABC signal amplification system

| 10% NGS/PBS | 10% (v/v) Normal goat serum (NGS) in PBS. |
| Wash buffer | 0.05% (v/v) TWEEN20™ in PBS. |
| DAPI | Working concentration 100µg/mL (w/v), use 1/200 (Sigma-Aldrich Company Ltd, UK). |
| PI | Working solution 100µg/mL (w/v), use 1/200 (Sigma-Aldrich Company Ltd, UK). |
| SG | Working concentration 25µM, use 1/100 (Molecular Probes Europe BV, Netherlands). |

The StreptABComplex/HRP (horse radish peroxidase) or StreptABComplex/AP (alkaline phosphatase) signal amplification kits (DAKO Ltd, UK) were used. The ABCComplex mixtures were made up according to the manufacturer’s protocol. If HRP-labelled antibodies were used, the tissue sections were first treated with 3% (v/v) Hydrogen Peroxide (Sigma-Aldrich Company Ltd, UK) at room temperature for 10mins in a humid box, before it was incubated with 10% NGS/PBS. Biotin-labelled species specific secondary antibodies (Sigma-Aldrich Company Ltd, UK) were diluted to 1:300
in PBS and used on the sections at room temperature for 30mins. The tissue sections were washed 3x5mins in wash buffer, then incubated with the mixture containing the ABCComplex and a nucleic acid dye (DAPI, PI or SG). The appropriate enzyme substrates (described below) were added after the tissue sections were washed 3x5mins in the wash buffer.

**B) FAST RED - substrate for alkaline phosphatase activity**

| Tris-buffered saline buffer (TBS) | 2.42g Trizma base, 8g NaCl, 3.8mLs of 1M HCl. Adjust pH to 7.6. Make up to 1L solution. |

The Fast Red TR/Napthol AS-MX substrate (Sigma-Aldrich Company Ltd, UK) was made up according to the manufacturer’s protocol. Tissue sections were washed in TBS for 5mins before the substrate was added. The reaction was stopped once substrate deposition was observed on the tissue section under a light microscope by rinsing in wash buffer.

**C) Tyramide Signal Amplification (TSA)-Direct - substrates for HRP activity**

| Wash buffer | 0.05% (v/v) TWEEN20™ in PBS. |

Coumarin tyramide (BlueFISH TSA-Direct kit; NEN Life Science Products Incorporated, USA), tetramethylrhodamine tyramide (RedFISH TSA-Direct kit, NEN Life Science Products Incorporated, USA) or fluorescein tyramide (GreenFISH TSA-Direct kit, NEN Life Science Products Incorporated, USA) substrates were used. The substrate solution was made up according to the manufacturer’s protocol. Tissue sections were incubated with the substrate solution for 8mins in a dark humid box, then washed 3x5mins in wash buffer.
2.15 DNA fluorescent *in situ* hybridisation (FISH) on tissue sections

2.15.1 Generation of Digoxigenin (DIG)-labelled DNA probes

DIG-labelled DNA probes used in FISH were synthesised using the random priming DIG-labelling kit (Roche Diagnostics Ltd, UK) according to the manufacturer's protocol. Restriction enzyme linearised PV genome DNA was used as template. 3μg of template was used in each labelling reaction.

2.15.2 Pre-hybridisation protocol

**Proteinase K solution** - 20mg/mL (w/v) stock solution. Dilute to 50μg/mL (w/v) in PBS.

Formalin-fixed paraffin-embedded tissues were dewaxed and hydrated as described in Section 2.14.1. Tissue sections were treated with 3% (v/v) Hydrogen Peroxide at room temperature for 10mins in a humid box, and washed in PBS for 2x2mins. This was followed by a 15min incubation with the proteinase K solution at 37°C in a humid box, and 2x2mins PBS washes.

2.15.3 DNA-DNA hybridisation

**Dextran sulphate** - 50% (w/v) in deionised water. Sigma-Aldrich Company Ltd, UK.

**Low molecular weight salmon sperm DNA** - dissolve at 10mg/mL (w/v) in deionised water at room temperature, then denatured by boiling for 10mins. Sigma-Aldrich Company Ltd, UK.

**20xSSC** - 88.23g Tri-sodium citrate, 175.32g NaCl per litre. Adjust pH to 7.8 with Sodium hydroxide.

**Hybridisation buffer** - 50% (v/v) Deionised formamide (Sigma-Aldrich Company Ltd, UK), 1x Denhardt’s (Sigma-Aldrich Company Ltd, UK), 5% (w/v) Dextran sulphate, 200μg/mL (w/v) Salmon sperm DNA and 4x SSC made up with deionised water. Store at -20°C.

DIG-labelled probes were mixed in hybridisation buffer at a ratio of 1:25. 3-10μL of the probe mixture was added onto each tissue section. The section was covered with a glass
coverslip, which was sealed around the edges with cow gum. The slides were heated at 93°C on a heating block for 5 mins, then quenched in a bucket of ice. Hybridisation was carried out at 42°C overnight in a humid box.

2.15.4 Post-hybridisation washes

Formamide wash buffer - 50% (v/v) Formamide, 2xSSC, 0.05% (v/v) TWEEN20™.
20xSSC - 88.23g Tri-sodium citrate, 175.32g NaCl per litre. Adjust pH to 7.8 with NaOH.

The cow gum was removed and the coverslip loosened in PBS with shaking. The tissue section was washed with pre-warmed (45°C) wash buffers on a shaker as follows: 2x5 mins formamide wash buffer, 2x5 mins 2xSSC solution.

2.15.5 Detection of DIG-labelled probes

TNB Blocking buffer - 5% (w/v) Blocking reagent (TSA-Direct Kit, NEN Life Science Products Incorporated, USA), 0.1M Tris-HCl solution (pH 7.5), 0.15M NaCl.
Wash buffer - 0.05% (v/v) TWEEN20™ in PBS.
DAPI - Working concentration 100μg/mL (w/v), use 1/200 (Sigma-Aldrich Company Ltd, UK).
PI - Working solution 100μg/mL (w/v), use 1/200 (Sigma-Aldrich Company Ltd, UK).
SG - Working concentration 25μM, use 1/100 (Molecular Probes Europe BV, Netherlands).

Sections were blocked with TNB Blocking buffer at room temperature for 1 hr in a humid box. A mixture containing 1:500 dilution of anti-DIG-POD (HRP) antibody (Roche Diagnostics Ltd, UK) and a nucleic acid dye (DAPI, PI or SG) was prepared in TNB Blocking buffer, and added onto the tissue section for 1 hr at room temperature, in a humid box. The section was washed 3x5 mins in wash buffer at room temperature on a shaker. HRP activity was detected using the TSA Direct Kits (NEN Life Science Products Incorporated, USA) as described in section 2.14.3.C, and the section was mounted with one drop of citifluo (Agar Scientific, UK) and a glass coverslip, then viewed under a fluorescent microscope.
CHAPTER 3 - E4 PROTEIN SEQUENCE ALIGNMENTS

3.1 Introduction ............................................................................................................ 93

3.2 The E1^E4 transcripts and proteins ....................................................................... 96
   3.2.1 The E4 ORFs and gene products ................................................................. 96
   3.2.2 E4 proteins ................................................................................................. 97

3.3 E4 alignments between viruses from the same phylogenetic Supergroup/Group ................................................................. 103
   3.3.1 Supergroup C (BPV-1 and BPV-2) ............................................................... 103
   3.3.2 Supergroup E ............................................................................................ 106
   3.3.3 Group A10 ................................................................................................ 112

3.4 Other conserved features between the E4 proteins ............................................. 114
   3.4.1 The proline-rich region in the rabbit and the Group A10 PVs ...................... 114
   3.4.2 Nuclear and nucleolar localisation signals ................................................ 116
   3.4.3 Nuclear export signal (NES) ...................................................................... 118

3.5 Discussion .............................................................................................................. 121
3.1 Introduction

Cladistic phylogenetic classification is a way in which the evolutionary history of organisms can be reflected through the analysis of their sequence structures (Wiley, 1978 & 1981). The first phylogenetic tree, consisting of nine PV types, was constructed in 1986, when an alignment of the conserved domains of the viral E1 gene was performed (Giri & Danos, 1986). As more PV types were included in the alignment analysis, the PV phylogenetic trees evolved into a consistent binary structure where mucosal and cutaneous tropic PVs formed the two main branches of PVs respectively (Chan et al., 1992; Chan et al., 1995). This enabled the viruses to be classified into supergroups and groups (Chan et al., 1995; Myers, 1997). New PV types were identified based on percentage of sequence dissimilarity of specific conserved regions within the virus genome, and placed in the appropriate groups in the phylogenetic trees. It is important that only regions exhibiting conserved amino acid sequences between different PV types, such as specific regions in the E6, E7 and L1 genes, are used for sequence alignment studies. For this reason, the highly variable E4 gene is considered unsuitable for such studies. So far, alignment analyses carried out by independent groups based on conserved regions of the E6 or L1 genes (Chan et al., 1995; Myers, 1997), have generated phylogenetic trees with similar topologies. Only minor dissimilarities with virus placements were found between them.

Most PVs can be categorised into five distinct Supergroups, namely Supergroups A, B, C, D and E. Supergroup A contains all the genital and mucosal tropic HPV types. The majority of the viruses classified in Supergroup B are the EV-associated HPVs. Supergroups C and D consist of ungulate PVs, which are unique in both their viral genome organisation and pathology. PVs in Supergroup C can infect fibroblasts and keratinocytes and cause fibropapillomas, while PVs in Supergroup D infect only keratinocytes, and cause true papillomas. Supergroup E is the least defined group in the PV phylogeny, and encompasses cutaneous
PVs (HPV-1, HPV-41, HPV-63, and CRPV), as well as mucosal animal PVs (COPV, and ROPV) (Chan et al., 1995).

The E4 ORF is found in the genome of all known PVs. The preservation of this viral ORF suggests an important role for the E4 gene product in the virus life cycle. Although the function of the E4 protein is elusive, it is assumed that E4 may play a common role in the life cycle of different PV types. Conserved motifs and domains within the E4 protein of some HPV types have been identified (Doorbar & Myers, 1996). These include a leucine cluster, a proline-rich region, a hydrophilic charged domain and a C-terminal motif (Doorbar et al., 1989; Doorbar & Myers, 1996). In addition, hyper-phosphorylation, multimerisation, and proteolysis of the E4 gene products have been described in some HPV infections (reviewed in Doorbar & Myers, 1996). So far, the identification and analysis of these conserved features were done only on a selected number of HPV E4 proteins.

This chapter attempts to examine the characteristics of the E4 proteins of BPV-1, COPV, CRPV, ROPV and HPV-11. This range of PV types was chosen based on their relevance to the aims of this thesis, previously described in chapter 1. E4 gene sequences were obtained from the PV sequence website (http://hpv_web.lanl.gov/), and the E4 protein sequences were translated from the respective E1^E4 spliced mRNA sequence using the EditSeq program (DNASTAR Incorporated, USA). E4 protein alignments were systematically carried out among viruses of the same phylogenetic supergroups using the MegAlign (DNASTAR Incorporated, USA) and Lalign (ISREC bioinformatics group) sequence alignment programmes. The first allows multiple E4 sequences to be simultaneous aligned, and the latter finds the best local sequence homologies between two entry sequences. Putative phosphorylation sites, as well as conserved motifs and domains in the E4 proteins were also examined. Last but not least, E4 protein sequences were also analysed using the blastp
(protein sequence database) program of BLAST (version 2.0; BLAST2.0), where other proteins with sequence homologies to E4 were listed.
3.2 The E1^E4 transcripts and proteins

3.2.1 The E1^E4 ORFs and gene products

The E4 protein is expressed from a spliced E1^E4 mRNA message following the activation of the viral differentiation dependent promoter. Table 3.1 describes the nucleotide positions of the E4 ORF and E1^E4 message, the size of the E1^E4 gene product, and the PV classification.

<table>
<thead>
<tr>
<th>Virus</th>
<th>E4 ORF</th>
<th>E1^E4 mRNA</th>
<th>E4 protein</th>
<th>Tissue tropism</th>
<th>Phylogenetic Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPV-1</td>
<td>3173-3529</td>
<td>(849-864)(^{\wedge}) (3225-3529)</td>
<td>107</td>
<td>C</td>
<td>C1</td>
</tr>
<tr>
<td>CRPV</td>
<td>3377-4015</td>
<td>(1362-1371)(^{\wedge}) (3714-4015)</td>
<td>103</td>
<td>C</td>
<td>E</td>
</tr>
<tr>
<td>ROPV</td>
<td>3299-3935</td>
<td>(1146-1155)(^{\wedge}) (3543-3935)</td>
<td>134</td>
<td>M</td>
<td>E</td>
</tr>
<tr>
<td>COPV</td>
<td>3116-3463</td>
<td>(816-831)(^{\wedge}) (3144-3463)</td>
<td>111</td>
<td>M</td>
<td>E</td>
</tr>
<tr>
<td>HPV-11</td>
<td>3255-3581</td>
<td>(832-847)(^{\wedge}) (3325-3581)</td>
<td>90</td>
<td>M</td>
<td>A10</td>
</tr>
</tbody>
</table>

Table 3.1. Description of the E4 ORF (nucleotide positions), E1^E4 message (nucleotide positions) and protein (number of amino acids), and the virus classification. ^ - spliced message; C - Cutaneous; M - Mucosal.

Two points could be made from the data shown in Table 3.1. Firstly, the sizes of the E4 ORFs vary between the viruses. When compared to another member of the Supergroup E PVs, the CRPV E4 ORF (639 bp) is more than twice as large as the HPV-41 E4 ORF (306 bp). However, this did not affect the expression of an E4 protein of about one hundred amino acids (comparable to other E4 proteins), since the ^E4 splice acceptor consensus is located near the middle of the CRPV E4 ORF. Secondly, the size of the ROPV E4 proteins is almost 30% larger than the CRPV E4 protein.
3.2.2 E4 proteins

The E4 protein sequences of the selected range of PVs were examined for the presence of conserved motifs and domains, such as the leucine cluster (LLXLL), the proline-rich region, a highly charged region, and the C-terminal consensus sequence for cutaneous PVs (DLXDYW) (reviewed in Doorbar and Myers, 1996). Putative phosphorylation sites of protein kinase C (PKC), mitogen-activated protein kinase (MAPK), casein kinase II (CKII), and protein kinase A (PKA) activities were identified. The consensus sequences used for the identification of these phosphorylation sites are - protein kinase C (PKC) [S/ TXR/K, K/RXXS/T, K/RXS/T] (Russo et al., 1992), mitogen-activated protein kinase (MAPK) [PXXS/TP, PXS/TP] (Davis, 1993), casein kinase II (CKII) [S/TXXD/E] (Russo et al., 1992) and protein kinase A (PKA) [RXS/T, RR/KXS/T] (Pearson & Kemp, 1991). These kinase consensus sequences were previously used for the analysis of HPV-11 E4 phosphorylation by Bryan et al. (2000b). Lastly, E4 protein sequences were analysed using the basic local alignment sequence software (BLAST2.0). Sequences were entered into the BLAST database (non-redundant) and proteins sharing sequence homology with E4 were listed (data in Appendix 1/1).

**BPV-1**

BPV-1 E4 contains five E1 and one hundred and two E4 amino acids (Figure 3.2a), has a calculated molecular weight of 11.8 kDa, and a charge of +11.185 at neutral pH. It contains a leucine cluster (SLSLL; positions 31-35) that resembles the (LLXLL) motif near the N-terminus of some HPV E4. A short proline-rich region is present after the leucine cluster. Two charged domains are found, one close to the N-terminus and the other the C-terminus of the protein. BPV-1 E4 has nine putative phosphorylation sites. Altogether, there are six PKC, no MAPK, two CKII, and one PKA recognition sites. Unlike the other E4 proteins analysed here, BPV-1 E4 does not have any hydrophobic residues at the C-terminus.
From the BLAST database search entry, the closest relatives to BPV-1 E4 were the E4 proteins of other members of the PV Supergroup C classification, such as the European elk PV (EPV) and deer PV (DPV) (Appendix 1/1). BPV-2 E4 protein sequence is almost identical to BPV-1 E4, with 89% sequence identity between the proteins (Figure 3.3).

**CRPV (Washington b strain)**

CRPV E4 contains three E1 and one hundred E4 amino acid residues (Figure 3.2b), has a calculated molecular weight of approximately 12 kDa, and a charge of +6.363 at pH 7.0. There are nine putative phosphorylation sites in the protein, seven sites by PKC, one site by MAPK, one site by CKII, and two sites by PKA. CRPV E4 does not have a N-terminal leucine cluster. Instead, a long leucine-rich domain, stretching from residue 68 to the end of the protein, is present. In this domain, leucine (L) or isoleucine (I) residues constitute 30% of the protein sequence. A proline-rich region is present near the N-terminus. Although charged residues are found all along the length of the protein, CRPV E4 does not contain a charged domain following the proline-rich region. However, CRPV E4 has a highly charged domain near the C-terminus of the protein (positions 82-91). The C-terminal consensus sequence present in most cutaneous HPVs is not found here.

When queried in the BLAST 2.0 protein database, CRPVb and ROPV E4 proteins showed the highest degree of sequence identity to CRPV E4 (Appendix 1/1). Other interesting matches included HPV-6c and HPV-11 E4.

**ROPV**

ROPV E4 contains four E1 and one hundred and thirty E4 amino acid residues (Figure 3.2c), has a calculated molecular weight of about 15 kDa, and a net charge of -4.895 at neutral pH. There are eleven putative phosphorylation sites along the protein, and four PKC, five MAPK, two CKII, and one PKA consensus motifs. Similar to CRPV E4, ROPV E4 does not have
Chapter 3 – E4 Protein Sequence Alignments

a N-terminal leucine cluster, but a leucine-rich domain (30% leucine/isoleucine content) at the C-terminus, from residue 96 to the end of the protein. A N-terminal proline-rich stretch is also present. ROPV E4 has two charged domains, the first stretching from residues 33-71, and the second from residues 97-121. The first charged domain can be sub-divided into an arginine (R) and lysine (K) rich basic cluster, followed by an acidic stretch rich in aspartic acid (D) and glutamic acid (E) residues. The second charged zone consists predominantly of acidic residues, aspartic acid (D) and glutamic acid (E). A region made up entirely of polar (serine (S) and threonine (T)) and proline residue repeats is present between the two charged domains. A C-terminal domain with similar biochemical properties to the cutaneous HPVs C-terminus consensus sequence is present in ROPV E4. The ROPV E4 C-terminus domain (DLEEFG; positions 118-24), has an additional glutamic acid (E) residue in the middle of the consensus, when compared to the (DLXDYW) consensus sequence of the cutaneous HPV E4 proteins.

Blastp analysis of the ROPV E4 protein sequence listed a variety of non-related membrane associated proteins and glycoproteins (Appendix 1/1). Among these include (M7) membrane virion glycoprotein 150 (murine herpesvirus 72), proline-rich cell wall protein (Gossypium hirsutum) and an assortment of proline-rich proteins. CRPVb E4 was also found to be remotely related to ROPV E4.

**COPV**

COPV E4 contains five E1 and one hundred and six E4 amino acid residues, and has a calculated molecular weight of approximately 12.6 kDa (Figure 3.2d). At neutral pH, it has a charge of -2.76. Five putative phosphorylation sites are identified along the protein based on consensus motifs for PKC (three sites), MAPK (one site), CKII (one site), and PKA (two sites). Two postulated PKC and PKA phosphorylation consensus motifs are also found. Unlike the other E4 proteins, COPV E4 does not have a leucine cluster. A proline-rich domain is present at the N-terminus of the protein. Interestingly, COPV E4 is a highly charged protein, with charged
residues (histidine (H), lysine (K), arginine (R), aspartic acid (D), and glutamic acid (E)) constituting 40% of its protein composition. Although these charged residues are distributed along the entire length of the protein, two regions in particular, contain a series of mono-charged residues. The first region begins at position 50 and contains a bipartite string of ten acidic residues (D, E), and the second region starts seven amino acids downstream, and consists of a string of basic residues from positions 67-79. Another charged domain, made up of both basic and acidic residues, is also be found near the C-terminus of the protein. COPV E4 has a C-terminal motif (DWEDFC) similar to the HPV-1 C-terminal domain sequence (DLDDFC) (Doorbar & Myers, 1996). COPV and ROPV E4 are the first animal PV proteins shown to share a C-terminal domain sequence similar to that found in HPV-1 E4.

COPV E4 was found to share sequence identity with the E4 proteins of HPV-5, HPV-21, and HPV-36 from the blastp query. These HPVs are cutaneous tropic PVs and belong to the main groups of viruses in Supergroup B1. Other proteins listed from the BLAST database search included Pseudorabies ORF1, ORF2 and ORF3 gene sequence product, and nucleolin (protein C23) [Gallus gallus] (Appendix 1/1).

**HPV-11**

HPV-11 E4 contains five E1 and eighty-five E4 amino acid residues (Figure 3.2e), has a charge of +3.32 at neutral pH, and shown to be expressed as an 10/11kDa protein in infected tissues (Brown et al., 1988 & 1991). HPV-11 E4 has a N-terminal leucine cluster (LLNLL), followed by a proline-rich domain, and charged region (Doorbar and Myers, 1996). It is shown to be phosphorylated in vitro by PKA (threonine and serine at positions 36 and 44, respectively) and MAPK (threonine at position 53) (Bryan et al., 2000b). No C-terminal consensus is present in this protein.
From the blast query list, the E4 proteins of the Group A10 PVs showed the most sequence identity with HPV-11 E4. In decreasing order of sequence identity, the list include HPV-6, HPV-13, HPV-55, HPV-44, common chimpanzee PV type (ChPV), and pigmy chimpanzee PV type 1 (PCPV1) (Appendix 1/I).
Chapter 3 - E4 Protein Sequence Alignments

**Figure 3.2.** E1^E4 protein sequences of BPV-1, CRPV, ROPV, COPV and HPV-11.

Conserved features were identified in the E4 proteins expressed by the different PVs, such as the proline-rich regions, charged domain(s) and multiple putative phosphorylation sites. Sequences translated from the E1^ORE are in italics. **Bold** - leucine cluster sequences. Grey highlight - proline-rich regions. Underlined domains - charged-residue rich domains. **Colour highlights** - putative phosphorylation sites. Bryan *et al.* (2000) showed that the HPV-11 E1^E4 protein can be phosphorylated *in vitro* by PKA (positions 36 and 44) and by MAPK (position 53).
3.3 E4 protein alignments between viruses from the same phylogenetic Supergroup/Group

E4 protein sequence alignments were performed using Lalign and MegAlign (DNASTAR Incorporated, USA). Lalign compares sequences of protein segments between two target proteins, whereas MegAlign analyses the alignment of complete amino acid sequences of multiple target proteins. Percentage identity scores between two entry proteins could also be obtained in both analysis programmes. The alignment profiles of the E4 proteins of Supergroup E PVs were analysed more extensive in this section as this supergroup of PVs consists of an interesting combination of animal and human, as well as mucosal and cutaneous PV types.

3.3.1 Supergroup C (BPV-1 and BPV-2)

Supergroup C consists of the ungulate fibropapillomaviruses. BPV-1 and BPV-2 are further grouped together as Group C1, whereas DPV, EPV, and ovine (sheep) PV types 1 and 2 (OvPV) form Group C2 in this PV supergroup.

Alignment of BPV-1, BPV-2, DPV and EPV E4 revealed the presence of a number of conserved motifs, particularly in the central region of the proteins (positions 30-90) (Figure 3.3a). The E4 proteins of these fibropapillomaviruses, except DPV, share a similar leucine cluster (defined as LXLL) as the mucosal HPV (Doorbar & Myers, 1996), near the N-terminus of E4. A proline-rich region is found in BPV-1 and BPV-2 E4, but this region is not as prominent in the DPV or EPV proteins. The E4 C-terminus contains a domain that is rich in charged residues. A (RXTXQ) motif is found at the C-terminus of BPV-1, BPV-2, and EPV, but not DPV E4.

The amino acid sequence of BPV-1 and BPV-2 E4 are highly homologous, with an identity score of approximate 90% (Figure 3.3b). On the other hand, DPV and EPV E4 are more
closely related to each other than the Group C1 PVs, with a amino acid similarity percentage of 64.7% (data not shown).
Figure 3.3. Alignment analysis of the E1\(^{\wedge}E4\) protein sequences of the supergroup C PVs. The E4 proteins of BPV-1, BPV-2, DPV and EPV were aligned using the MegAlign sequence analysis program. Conserved amino acid residues, are represented by the consensus ruler. The strengths of amino acid conservation are represented by the height of the coloured bars (in decreasing order: red, green, light blue, dark blue). This report shows the greatest amount sequence homology in the central region of the E4 proteins. Alignment of the BPV-1 and BPV-2 E4 also showed that the primary protein sequence of the proteins are almost identical.
3.3.2 Supergroup E

Supergroup E consists of a mixture of different species and tissue type specific PVs. There are three animal (COPV, CRPV, and ROPV) and three human (HPV-41, HPV-1 and HPV-63) types PVs in this supergroup. There is one sub-group (Group E1) of PVs consisting of HPV-1 and HPV-63, based on previous alignment studies of the E6 and L1 gene sequences (Chan et al., 1995). Interestingly, sequence analysis of the COPV E6 and L1 genes has found that this mucosal PV is a relative of the cutaneous HPV-1 and HPV-63 (Bernard, 1997). Since CRPVb have been reported to show a high degree of homology in its genome organisation and gene sequences with CRPV (Salmon et al., 1997 & 2000), CRPVb has been included in this study as a supergroup E PV.

The alignment of E4 also showed a similar relationship as described by previous researchers. A phylogenetic tree was generated using the MegAlign program, based on alignments of E4 protein sequences of all seven PVs in this supergroup (Figure 3.4a). From the data, this group of PVs were separated into two branches, with HPV-1, HPV-41, HPV-63 and COPV in one branch, and the rabbit PVs in the other. As expected, HPV-1 and HPV-63 shared a high degree of sequence homology in their E4 sequences. Based on the E4 alignment phylogenetic tree, COPV is evolutionary closer to HPV-41 than HPV-1 and HPV-63. Among the rabbit PVs, the cutaneous tropic CRPV and CRPVb seemed to share a more recent ancestor than with ROPV (mucosal tropic PV).

The E4 alignment report generated from MegAlign revealed a common C-terminal consensus domain in HPV-1, HPV-63, HPV-43 and COPV (Figure 3.4b). This C-terminal consensus can be represented in single letter amino acid translation as D-hpb-E-D-F-X-R-K-L-G-I, where hpb represents a hydrophobic residue. This domain shares some characteristics with the C-terminus motif (DLXDYW) found in most cutaneous HPVs in Supergroup B (reviewed in Doorbar & Myers, 1996). Although the D-hpb-E-D-F-X-R-K-L-G-I consensus is most defined in
the HPV and COPV E4 proteins in this supergroup, a domain with similar chemical properties, i.e. charge and solubility, is also found in ROPV (Figure 3.4c), but not in the CRPV or CRPVb E4 proteins (data not shown). Apart from this shared C-terminus feature, the E4 proteins of the two mucosal animal PVs in this group (COPV and ROPV) did not show high levels of sequence homology with each other in terms of their protein sequence (Figure 3.4c). Alignment of the cutaneous PVs, CRPV, HPV-1 and HPV-41, E4 proteins also showed very little sequence homology between them (data not shown).
### Figure 3.4. Alignment analysis of the E1^E4 protein sequences of the supergroup E PVs.

(a) Shows the phylogenetic tree of the Supergroup E viruses generated from the multiple alignment of the E1^E4 protein sequences by MegAlign. PVs in this supergroup can be divided into two subgroups based on their E1^E4 protein sequences, where the rabbit PVs are grouped together, and the HPVs and COPV are found in the other subgroup. (b) MegAlign analysis of the E1^E4 protein sequences of COPV and the HPVs in this supergroup identified conserved regions of the C-terminus of the proteins (highlighted in bold). (c) Alignment of the E1^E4 proteins of the two mucosal animal PVs (COPV and ROPV) in this supergroup found conserved amino acid residues in both proteins. Based on the amino acid composition, both viral proteins share similar chemical properties at their C-terminus. These residues are highlighted in grey.
E4 alignment of CRPV, CRPVb and ROPV showed that a number of common features were shared between the proteins. All three E4 proteins have a common N-terminal protein sequence, M-A-E-A, translated from their respective E1^E4 transcripts. Another conserved region is present near the C-terminus of the proteins. This region is a nineteen amino acid stretch, made up of a mixture of charged and hydrophobic residues. This domain can be defined as D-X-L-L-Q-R-X-L-X-E-hpb-R-X-L-X-E-X-L. Interestingly, the C-terminus of CRPVb E4 contains sequence elements that align with both CRPV and ROPV E4. Besides sharing an identical C-terminal protein sequence with CRPV E4, CRPVb E4 has an additional eight amino acids at its C-terminus end (Figures 3.5). This eight amino acid stretch is very similar to the last eight amino acids of the ROPV E4 protein. The significance of this in terms of the evolution of these viruses is discussed later. CRPV and CRPVb showed a 95% sequence identity in the first hundred and three amino acids of their E4 proteins (data not shown). When compared to ROPV E4, both MegAlign and Lalign results showed that ROPV E4 shared more amino acid homology with CRPVb than CRPV (data not shown).
Figure 3.5. Alignment analysis of the E1^E4 protein sequences of the rabbit PVs. MegAlign alignment report showing conserved domains and amino acid sequences in the E4 proteins of CRPV, CRPVb and ROPV. All three E4 proteins share the same N-terminus amino acid sequences derived from the respective viral E1 ORF, as well as a charged region near the C-terminus. The C-terminus of CRPVb E4 contains common sequences found in the E4 proteins of CRPV and ROPV.
3.3.3 Group A10

Supergroup A makes up one of the main branches of the PV phylogeny and contains mainly the genital/mucosal tropic PVs. Group A10 consists of HPV types 6, 11, 13, 44, 55 and 74, as well as the chimpanzee PVs (PCPV, and ChPV) (Chan et al., 1995; Myers, 1997). This group features a number of low-risk genital HPVs which also infects oral and laryngeal, as well as penile epithelia.

E4 alignments showed that the N-terminal leucine cluster (LLXLL), proline-rich motif, core charged region and C-terminal consensus motif were highly conserved among this group of PVs (Figure 3.6a). Percentage similarity scores of >90% were obtained for the comparison of E4 sequence between HPV-11 and HPV-6c, and HPV-44 and HPV-55 (data not shown). These data suggest that these PVs probably shared a common ancestor. A phylogenetic representation of the evolution distance between the viruses is shown in figure 3.6b.
Figure 3.6. Alignment analysis of the E1^E4 protein sequences of the Group A10 PVs. (a) Multiple alignment of the E1^E4 protein sequences using the sequence analysis program, MegAlign. All proteins contain a leucine cluster, proline-rich region, and C-terminal domain consensus (shown in bold). (b) Shows the phylogenetic representation of the MegAlign alignment report in (a). A high level of sequence homology was found between the E4 proteins of HPV-11 and HPV-6, as well as HPV-44 and HPV-55.
3.4 Other conserved features between the E4 proteins

3.4.1 The proline-rich region in the rabbit and the Group A10 PVs

The proline-rich region is a conserved feature of the PV E4 protein. Alignment of CRPV, CRPVb and ROPV E4 showed a consensus domain in the proline-rich region of the proteins, which can be described as L-($\pm$)-T-P-P-(+)-R-P-P-X-hpb-Q-C-P-P-hpb, where (+) represents a basic residue (K, H, or R) (Figure 3.7). Interestingly, this consensus domain was also found to be conserved in HPV-6 and HPV-11 E4 proteins. Derivatives of this consensus sequence were present in HPV-13, HPV-44, HPV-55 and PCPV1 E4 proteins. This may suggest a common ancestral link between the rabbit PVs and the Group A10 mucosal PVs.
### Consensus

<table>
<thead>
<tr>
<th>Protein</th>
<th>Consensus Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRPV</td>
<td>L- (+) - T- P- (+) - R- P- P- X - hpbl - Q- C- P- P- hpbl</td>
</tr>
<tr>
<td>CRPVb</td>
<td>L K T P S P K R P P T V Q C P P L&quot;</td>
</tr>
<tr>
<td>ROPV</td>
<td>L T T P R R P P - L Q Y P Q A&quot;</td>
</tr>
<tr>
<td>HPV-6c</td>
<td>L H T F P H R P P P L Q C P P A&quot;</td>
</tr>
<tr>
<td>HPV-11</td>
<td>L H T F P H R P P P L Q C P P A&quot;</td>
</tr>
<tr>
<td>HPV-13</td>
<td>L H T F P P P P H R P P P Q C P A A&quot;</td>
</tr>
<tr>
<td>HPV-44</td>
<td>L H T F P P P P P L H C P L A&quot;</td>
</tr>
<tr>
<td>HPV-55</td>
<td>L H T F P P P P P L H C P P A&quot;</td>
</tr>
<tr>
<td>PCPV-1</td>
<td>L H T P Q P P L H R P P A Q C H P S&quot;</td>
</tr>
</tbody>
</table>

*Figure 3.7. Proline-rich region consensus in the E4 proteins of rabbit and Group A10 PVs. Interestingly, a high level of sequence homology was found in the proline-rich region of the CRPV, CRPVb, ROPV, HPV-6c and HPV-11 E4 proteins. The consensus sequence is more variable in the HPV-13, HPV-44, HPV-55 and PCPV-1 E4 proteins. **Bold** - conserved amino acid residues. The number at the end of each line denotes the amino acid position in the respective E1^E4 protein sequences. (+) - basic residue. hpbl - hydrophobic residue.*
3.4.2 Nuclear and nucleolar localisation signals

Nuclear localisation signals (NLSs) are the short peptide sequences that are required for the translocation of proteins from the cytoplasm into the nucleus compartment. NLS acts as an entry rather than a retention signal, and usually remain intact during transport. Although NLSs cannot be described by any tight consensus sequence, the two best characterised classes of NLSs consist of a string of four to seven basic amino acids (also known as the classical NLS), or a longer bipartite sequence composed of two stretches of basic amino acids separated by a ten-amino acid spacer (Wen et al., 1995). NLSs are very hydrophilic, usually preceded by a β-turn or random coil region, and highly antigenic surface structures (Jans & Hubner, 1996; Kaffman & O’Shea, 1999). A class of NLS motif that contained a helix-breaking N-terminal amino acid (P or G), followed by at least three basic amino acids in a stretch of seven to nine residues has also been reported (Dang & Lee, 1989). Putative NLSs were found in the BPV-1, COPV, and CRPV E4 sequences (Figure 3.8a).

Unlike the short NLS motifs, a longer stretch of basic residues might be required for protein targeting to the nucleolus (Dang & Lee, 1989). This trend was also found by Liu et al., (1999) when they compared the nucleolar localization sequences (NoLSs) of known nucleolar proteins. Since COPV E4 has been previously shown to be localised in the nucleolus (Nicholls et al., 2001; Peh et al., in press), a stretch of basic residues that might serve as a putative NoLS was found in the protein (Figure 3.8b).
Chapter 3 - E4 Protein Sequence Alignments

(a) 

**Putative classic NLS**

BPV-1  \[\text{G K K K S S R}^\text{14}\]

ROPV  \[\text{P R T I K P R S S R Y R G}^\text{14}\]

COPV  \[\text{P R R G R R R L}^\text{19}\]

**Putative bipartite NLS**

Consensus \[\text{K/R K/R 10-12 amino acid spacer K/R K/R K/R}\]

CRPV  \[\text{K R P P T V Q C P P L K R K}^\text{24}\]

(b) 

**Putative NoLS**

COPV  \[\text{P S R S R P R R G R R}^\text{35}\]

---

Figure 3.8. Putative NLSs and NoLSs in E4.

(a) Shows sequences in BPV-1, ROPV, COPV and CRPV E4 proteins that resemble the classical (a stretch of basic amino acid) or bipartite NLS consensus sequence. (b) Shows a putative NoLS motif in the COPV E4 protein sequence made up a long stretch of basic amino acid residues. **Bold** - basic residues. The number at the end of each line denotes the amino acid position in the various E1^E4 protein sequences.
3.4.3 Nuclear export signal (NES)

From the study of the NES domain of the HIV-1 Rev protein, a new class of NES, which consists of a stretch of hydrophobic residues (especially leucines), was described (Fischer et al., 1995). The leucine-rich NES was able to mediate rapid and active transport of proteins into the nucleus, and did not require complex tertiary structures for its function (Gerace, 1995; Wen et al., 1995). However, these leucine-rich NESs had been suggested to form an α-helical folding based on the neighbouring residue composites (Fukuda et al., 1996). Comparative studies of some NESs by the helical wheel representation showed that a general feature of these motifs is the formation of a hydrophobic surface of 3 leucine residues, plus an additional hydrophobic residue on the other side (Fukuda et al., 1996).

Using the protein analysis program (Protean, DNASTAR Incorporated, USA), formation of α-helical and β-sheet structures can be predicted based the amino acid sequence of a protein. By studying the protein sequence of BPV-1, COPV, CRPV, CRPVB, ROPV and HPV-11 E4, putative NESs were found in each protein. These regions were also predicted to form α-helixes by the Protean analysis program. When represented as an α-helix wheel structure, these putative NES regions formed a leucine-rich hydrophobic face, with an additional hydrophobic residue on the other side (Figure 3.9). This configuration is a common feature in other leucine-rich NESs, and be recognised by a common NES receptor (Fukuda et al., 1996).
Figure 3.9. Alpha-helical wheel representation of the putative leucine-rich NES motifs in the E1^E4 proteins of CRPV (a; positions 77-89), CRPVb (b & c; positions 77-89 & 103-109), ROPV (d; positions 99-119), COPV (e; positions 103-112), HPV-11 (f; positions 7-17) and HPV-1a (g; positions 89-96).

These helical wheel representations illustrate the formation of a hydrophobic surface, and a hydrophobic residue on the other side of the alpha-helical folding.
Chapter 3 - E4 Protein Sequence Alignments

(d) ROPV

(e) COPV

(f) HPV-11

(g) HPV-1a
3.5 Discussion

**E4 proteins between different PV types**

E4 is considered one of the most variable PV protein in terms of its genetic and protein sequence. Therefore, not many alignment studies had been done on this viral protein. The analysis of the E4 protein sequences of BPV-1, COPV, CRPV, and ROPV have allowed the identification of conserved and unique features in these proteins, which had previously only been examined in selected groups of HPVs. The information provided here could also be useful in future studies that examine the function(s) of this list of animal PV E4 proteins.

At neutral pH, the different E4 proteins differed greatly in their overall protein charge. This reflected the great diversity in the primary sequence of the E4 proteins expressed by different PVs. However, it has been shown that *in vivo* phosphorylation of E4 proteins can affect the overall charge of the protein (Grand *et al.*, 1989). Phosphorylation is a major post-translation modification mechanism which can regulate protein function(s), translocation, and protein-protein associations, and other intracellular processes. Phosphorylated forms of E4 were first identified from HPV-1 infected wart extracts (Breitburd *et al.*, 1987), and more recently in HPV-11 infected genital lesions (Bryan *et al.*, 2000b). *In vivo* and *in vitro* phospho-labelling studies have shown that both HPV-1 and HPV-11 E4 are phosphorylated by more than one kinase at the serine (S) and threonine (T), but not tyrosine (Y) residues. HPV-1 E4 was shown to be phosphorylated at more than four different sites *in vitro* by a preparation of bovine cardiac cAMP-dependent protein kinase, which included PKA, and a still unknown kinase (Grand *et al.*, 1989). Similarly, papilloma extracted HPV-11 E4 was found to be phosphorylated at four different sites along the protein, although only three phospho-residues at positions 36, 44 and 53, were identified in *in vitro* assays, when MBP-HPV11E1^E4 fusion proteins were phosphorylated by PKA and MAPK, but not CKII and PKC (Bryan *et al.*, 2000b).
Multiple putative phosphorylation sites (by PKA, PKC, CKII and MAPK) were also identified in the E4 protein sequences analysed here. The number of phosphorylation sites as well as the frequency of phosphorylation by each kinase differed between the proteins (Figure 3.2). This suggests that E4 may be phosphorylated to different extents, depending on the E4 protein and its role in its host cell environment. All the E4 proteins contained one or more putative phosphorylation site(s) near the N-terminus, except HPV-11 E4, where the first putative phosphorylation site was found just before the central charged domain. Despite the close evolutionary relationship between CRPV and ROPV, the phosphorylation patterns, in terms of the frequency and position of the putative phosphorylation sites, of the viral E4 proteins did not resemble each other. Only two kinase sites were found to be common between the two rabbit PVs. The first is a threonine (T) residue (MAPK site) located in front of the proline-rich region, and the second is a CKII site near the C-terminus of CRPV (phosphothreonine) and ROPV (phosphoserine) E4. In both rabbit PV E4 proteins, the C-terminal CKII sites were located just before a stretch of charged residues. Even though multiple phosphorylation is thought to be a conserved feature of PV E4 proteins, its significance for the role of the protein during the virus life cycle remains to be evaluated.

From the analysis of E4 protein sequences, the proline-rich region is the only conserved domain found in all E4 proteins. A leucine cluster was not found in COPV, CRPV or ROPV E4, although a leucine-rich motif was identified near the central region of the fibropapillomavirus E4 proteins. The C-terminal domain previously described as a conserved feature in cutaneous HPV E4 proteins was found in COPV E4. The sequence homology between COPV and HPV-1 E4 proteins supports previous reports of an evolutionary link between COPV and HPV-1 (Delius et al., 1994). E4 alignments of COPV and the supergroup E HPVs showed that these viruses shared a conserved C-terminal domain in their E4 proteins which can be described as (D-hpb-E-D-F-X-R-K-L-G-l). Interestingly, the C-terminal domain of ROPV E4 was comparable to COPV E4 in terms of its amino acid properties. In earlier studies,
the C-terminal region of HPV-1 E4 has been shown to be involved in the multimerisation of E4 proteins (Doorbar et al., 1988 & 1996). Mutant studies of the C-terminus of HPV-1 E4 identified two hydrophobic residues in this region that were required for E4 multimerisation (Ashmole et al., 1998). When compared with the C-terminal sequences of COPV and ROPV E4, hydrophobic residues were also found in both positions in ROPV E4, but only in the first position in COPV E4. Although multimeric forms of COPV and ROPV E4 have not been identified, immuno-detection of these proteins in infected tissue sections showed the presence of E4 inclusions, which could result from the formation of E4 multimers in vivo (Nicholls et al., 2001; Peh et al., in press; Chapter 5).

**Sequence homologies between E4 and other proteins**

Using the BLAST2.0 protein database and analysis program (blastp), the PV E4 protein appears to be unique. The most common hits following an E4 protein query in the blastp program were proteins with a proline-rich region. These hits were usually dismissed as insignificant since proline-rich regions are widely found in many different proteins with biologically distinct functions. Despite the diversity in the E4 protein sequence, short segments of protein sequence identity were found between the E4 proteins of PVs from the same phylogenetic groups with the exception of COPV E4. When queried, COPV E4 was found to share sequence homology with the E4 proteins of three EV-associated HPVs, HPV-21, HPV-5 and HPV-36. A 20-30% sequence identity score was obtained from the alignment of COPV and the EV-associated HPV E4 proteins using MegAlign and Lalign (data not shown). Since the length of the EV-associated HPV E4 proteins are almost twice as long as COPV E4, it is assumed that the sequence homology may be due to the conserved motifs of the E4 proteins, and the viruses are unlikely to be close evolutionary relatives.
However, other interesting proteins were also listed from the COPV E4 query. They included the ORF2 gene product of Pseudorabies virus (PRV) (GenBank accession number M57505) (Cheung, 1991), and chick nucleolin (protein C23) (GenBank accession number P15771). These proteins share a common feature which consists of a stretch of at least ten acidic amino acid residues (D and E). The organisation of this charged region was also similar between COPV E4 and the PRV ORF2 gene product. PRV (or Suid Herpesvirus type 1) belongs to the alpha subfamily of the Herpesviridae family (Roizman, 1996). Like COPV, this virus can infect the mucosal epithelium and canine hosts (http://www.aphis.usda.gov/nc/vs/prv.html). Unfortunately, the function of the PRV ORF2 gene product is still unknown. The other protein showing homology, chick nucleolin, was thought to be interesting because COPV E4 has been detected in the nucleoli of COPV infected mucosal canine lesions (Nicholls et al., 2001; Peh et al., in press; Chapter 5). Analysis of the protein sequences using Protean predicted an alpha-helical folding structure in this charged region in COPV E4, PRV ORF2 and chick nucleolin. When represented as alpha-helical wheels, these regions formed negatively charged loops in all three proteins (data not shown). Thus, it may be interesting to investigate the role of this charged region in the protein function.

**Identification of new consensus motifs and other features in E4**

Alignment of the rabbit PVs (CRPV, CRPVb and ROPV) revealed a conserved domain (D-X-L-L-Q-R-X-L-X-E-E-hpb-R-X-L-X-E-X-L) near the C-terminus of E4. This region is made up of a mixture of hydrophobic and negatively charged amino acids and was found to exhibit a low surface probability and antigenic index (data not shown). The conservation of this nineteen amino acid region in all three proteins may suggest a role of this domain in the function of the proteins, possibly in the formation of a protein binding pocket.

A consensus motif was found in the proline-rich region of CRPV, CRPVb, ROPV, HPV-6c and HPV-11 (Figure 3.7). Epitope mapping of high affinity monoclonal antibodies has shown
that the proline-rich region of HPV-1 and HPV-16 E4 were highly antigenic (Doorbar et al., 1988 & 1997). On the other hand, functional studies involving HPV-1 and HPV-16 E4 deletion mutants showed that the proline-rich region is not important in the cellular distribution and cytokeratin association properties in SV40 recombinant virus infected cells (Roberts et al., 1994). Perhaps the proline-rich region of HPV-1 and HPV-16 E4 may have a structural instead of functional role. Thus, the identification of a consensus motif in the proline-rich region in the rabbit PVs, HPV-6c and HPV-11, may suggest a similar folding organisation in this region of the proteins.

Putative nuclear targeting sequences had been identified in BPV-1, CRPV, COPV and ROPV E4. Putative NESs were also found in CRPV, ROPV, COPV, HPV-11 and HPV-1 E4. Thus the cellular localisation of E4 may be governed by the translocation of the protein through the nuclear pore via its NLS or NoLS, and NES. The translocation of proteins via the NLS or NES is an active and regulated process, mediated by nuclear transport receptors (Kaffman and O'Shea, 1999). Therefore, the cellular localisation of E4 may be determined by the rate of nuclear import and export at the time when an equilibrium between the two processes is reached. In order to be predominantly cytoplasmic, the NLS in E4 may be weaker than the NES, so that at equilibrium, the rate of protein nuclear export is faster than import. The translocation of E4 into the nucleus may also be inhibited by the inactivation of the NLS by masking, phosphorylation, or cleavage (Jans and Hubner, 1996; Gerace, 1995; Wen et al., 1995). In addition, the association of E4 with the cytokeratin network (Doorbar et al., 1991) may also serve as an anchor by which the movement of E4 into the nucleus is prohibited. However, as a small protein, E4 may diffuse passively through the nuclear pore, especially since the protein is shown to be expressed in high levels in cells (Doorbar et al., 1997). In the case of COPV and CRPV E4, these proteins can be easily detected at high levels in the nucleus and cytoplasm of cells using specific E4 antibodies (Chapter 5). Since passive diffusion of proteins across the nuclear pore occurs at a very slow rate compared to the active transport of proteins via the NES
Chapter 3 – E4 Protein Sequence Alignments

and NLS, the accumulation of COPV and CRPV E4 in the nucleus may be directed by the activities of their nuclear import and export signals. Although the role of E4 is elusive, its cytoplasmic localisation is thought to be important for its function, since most E4 proteins are predominantly detected in the cytoplasm of infected cells. It remains to be seen if the nuclear import of E4 is important to the function of E4 during the virus life cycle.

Virus-host co-evolution

The isolation and characterisation of a new CRPV subtype (CRPVb) provided an experimental model to show that intratypic variation of the CRPV genome could affect disease outcome, and extent of viral replication and capsid production of CRPV in NZW rabbits (Salmon et al., 1997 & 2000). Sequencing of the two different CRPV subtypes showed that sequence variations between CRPV and CRPVb were not uniform in all the viral ORFs (Salmon et al., 1997). The greatest degree of significant sequence divergence was seen in the E6 ORF, which showed a 13.4% and 16.2% amino acid variation in the LE6 and SE6 protein sequences respectively. On the other hand, only 0.8% of the L1 protein sequence varied from the prototype. Using CRPV and CRPVb chimeric viral genomes, Salmon et al. (2000) showed that the LRR/E6/E7 region, together with host genetic constitution, play a major role in the outcome of disease. They have also showed that CRPVb infections in NZW rabbits were able to initiate high levels of viral replication and capsid protein production, in contrast to the non-productive infections in NZW rabbits following CRPV infections.

E4 protein alignments of CRPV, CRPVb, and ROPV E4 proteins revealed an interesting evolution link between these three rabbit PVs. CRPV and CRPVb were both recovered from cutaneous papillomas found on wild cottontail rabbits (Shope, 1933; Salmon et al., 1997). ROPV was isolated from oral papillomas in NZW rabbits, and is specifically mucotropic and antigenically unrelated to CRPV (Parson & Kidd, 1942; DiGiacomo & Mare, 1994; Harvey et al., 1998). From the E4 protein alignment analyses, it is clear that CRPVb E4 shares characteristics
of both the CRPV and ROPV E4 proteins. Sequence homology levels were found to be approximately 40% between ROPV and the CRPV subtypes, despite the size difference of about thirty amino acids between them. The most interesting evolutionary feature between the rabbit PVs is found at the C-terminus of E4, where an additional eight amino acid segment can be found that is common to CRPVb and ROPV E4, but not in CRPV E4. This protein segment may have been generated in CRPVb by point mutations (at positions 4015 and 4038) in the E4 ORF, which may have been from a common ancestor.

From the evolution point of view, PVs are genetically stable entities, which have evolved through the selection of host specificity during mild and productive infections as well as during disease persistence. In addition, the genetic stability of PVs may also be attributed to the dependence on the host error-correcting machinery for genome replication and maintenance of episomal DNA, making the evolution of the viruses intricately tied to that of the cellular replication processes (Shadan & Villarreal, 1993). Since the expression of E4 coincides closely with viral DNA amplification (Doorbar et al., 1997; Peh et al., in press), the role of E4 may be related to amplification of the virus genome. Therefore, based on the observations of the non-productive CRPV infections, and the productive CRPVb and ROPV infections in NZW rabbits, it may be speculated that the sequence similarities between CRPVb and ROPV E4 may influence the efficiency of virus production in the NZW rabbit hosts. In this case, the additional protein segment in the C-terminus of E4 may be a key factor which facilitated the activation of CRPVb late events during the infection of the NZW rabbit. It will be interesting to examine the relevance of the C-terminal tail of CRPVb and ROPV E4, in the productive infections of NZW rabbits.
CHAPTER 4 – E4 SPECIFIC ANTIBODIES

4.1 Introduction ................................................................. 129

4.2 Preparation and purification of the E4 immunogens ................. 130
   4.2.1 Preparation of the GST- and MBP-E1^E4 immunogens ................. 130
      A) E1^E4 fragments ......................................................... 130
      B) Cloning of the E1^E4 fragments into pGex4T-1 or p-Malc vectors ...... 132
      C) Verification of the positive clones ...................................... 132
      D) Fusion protein expression ................................................ 132
      E) GST-E4 or MBP-E4 immunogens ....................................... 133
   4.2.2 Preparation of the KLH (and BSA) conjugated peptide immunogens .... 135
      A) E1^E4 and E4 peptides .................................................. 135
      B) Peptide conjugation ....................................................... 135
      C) Preparation of the KLH-peptide immunogens ....................... 136

4.3 Immunisation of animals ................................................. 136

4.4 E4 specific antisera ....................................................... 138
   4.4.1 E4 specificity on western blots ....................................... 138
   4.4.2 E4 specificity and titre with ELISA ................................... 140

4.5 Discussion ................................................................. 144
4.1 Introduction

The generation and use of specific antibodies has become an integral part of molecular biology. Specific antibodies are widely used to detect and visualise proteins in cells and on western blots and in ELISAs. Previously, specific antibodies against BPV-1, HPV-1, HPV-6, HPV-11, HPV-16 and HPV-31b E4 proteins were used to detect E4 expression in infected cells and tissue sections (Doorbar et al., 1988 & 1997; Brown et al., 1991; Jareborg & Burnett, 1991; Rogel-Gaillard et al., 1993; Pray & Laimins, 1995). In order to characterise the life cycle of BPV-1, COPV, CRPV, ROPV and HPV-11 in infected tissue samples, specific antibodies against the viral E4 proteins were generated.

E4 polyclonal antibodies were generated in rats and rabbits following a standard immunisation protocol (see Chapter 2) with E4 containing fusion proteins or conjugated E1^E4 peptide preparations. The animals were inoculated subcutaneously with the immunogens, and boost injections were given periodically until satisfactory levels of antibody titres were attained. Antisera from the animals were tested for specific reactivity against the respective E4 immunogens on western blots and ELISAs. These E4 antisera were subsequently used to detect the E4 gene product in infected tissue sections.
4.2 Preparation and purification of the E4 immunogens

Two different types of E4 immunogens, namely E4 fusion proteins and conjugated E4 peptides, were used in the immunisation of rats and rabbits. To generate E4 fusion proteins, the complete length of the spliced E1^E4 ORF was cloned in-frame into two bacterial expression vectors, pGex4T-1 (Amersham Pharmacia Biotech, UK; Appendix 2) and pMAL-c2 (New England BioLabs Ltd, UK; Appendix 2). The pGex4T-1 (or pMAL-c2 E4) constructs were amplified in DH5α E. Coli cells, where the expression of glutathione sepharose-transferase (GST) (or maltose binding protein, MBP) E4 fusion proteins were induced by the addition of IPTG to the bacterial culture (see Chapter 2). The E4 peptide immunogens were made by conjugating short E4 peptides to the carrier protein keyhole limpet hemocyanin (KLH). E4 peptides were coupled to melaimide-activated KLH molecules via a cysteine amide group, which was synthetically added to the C-terminus of the E4 peptide sequences. All immunogen preparations were dialysed in PBS before they were used to inoculate the animals.

4.2.1 Preparation of the GST- and MBP-E1^E4 immunogens

A) E1^E4 fragments

The E1^E4 ORF of BPV-1, COPV, CRPV, ROPV and HPV-11 were amplified from the respective virus genome by PCR. The PCR primers used in these reactions were specifically designed to facilitate the cloning of the amplified E1^E4 fragment into the pGex and pMAL vectors. An additional thirteen nucleotides were added to the 5' end of the E1^E4 sequence of each forward primer, creating BamHI and EcoRI restriction enzyme sites in the 5' end of the PCR amplified E1^E4 fragment (Figure 4.1). An additional nucleotide segment was also added to the reverse primers, so that a SalI and SmaI cloning site would be created in the 3' end of the PCR product (Figure 4.1). The E1^E4 spliced genes were amplified using Taq polymerase (see Chapter 2).
Chapter 4 - E4 Specific Antibodies

**PCR amplified E1^E4 fragment**

BamH I/EcoR I  
5'  3'  
Sal I/Sma I  
E1^E4 sequence

**Forward primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>BamH I</th>
<th>EcoR I</th>
<th>5' - 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPV-1</td>
<td>GCA AAC GAT ARA GAG ATC GCC CAG ACG GA</td>
<td>(42)</td>
<td></td>
</tr>
<tr>
<td>CRPV</td>
<td>GCT GAA GCT CCC CCC AGC CCG TGC TCA</td>
<td>(40)</td>
<td></td>
</tr>
<tr>
<td>COPV</td>
<td>GCG GCT AGA ARA GTC CCG CCG GAA CCT CCG</td>
<td>(43)</td>
<td></td>
</tr>
<tr>
<td>ROPV</td>
<td>GCT GAA GCT CAA CCC CCC TAC GCC</td>
<td>(37)</td>
<td></td>
</tr>
<tr>
<td>HPV-11</td>
<td>GCG GAC GAT TCA GCA C TG TAC GAG AAG TAT CC</td>
<td>(45)</td>
<td></td>
</tr>
</tbody>
</table>

**Reverse primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sma I</th>
<th>Sal I</th>
<th>5' - 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPV-1</td>
<td>TCA CTG GTT CTT CCT CNG TG</td>
<td>(38)</td>
<td></td>
</tr>
<tr>
<td>CRPV</td>
<td>TTA TAA GCT CGC GAA GCC GTC TAT</td>
<td>(39)</td>
<td></td>
</tr>
<tr>
<td>COPV</td>
<td>CTA AAA CAA CAA CTG CCG GAT CGT</td>
<td>(42)</td>
<td></td>
</tr>
<tr>
<td>ROPV</td>
<td>TCA CTG CGC GAT CCC GAG</td>
<td>(36)</td>
<td></td>
</tr>
<tr>
<td>HPV-11</td>
<td>CTA TAG CGG TAG CTG CAC TGT GAC</td>
<td>(42)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.1. Specific PCR primers for the synthesis of E1^E4 sequence fragments.  
These specially designed primers contained specific restriction enzyme recognition sequences (BamH I, EcoR I, Sma I and Sal I) that would enable the cloning of the fragments into the prokaryotic fusion protein expression vectors pGex 4T-1 and pMalc-2. An additional stop codon (TAG) was also introduced at the end of the E4 ORF. The number at the end of each sequence represents the length of each primer. Bold - annealing bases. (*) - stop codon.
B) Cloning of the E1^E4 fragments into pGex4T-1 or pMalc vectors

The E1^E4 PCR products were purified prior to digestion with Sal I and EcoR I (refer to Chapter 2). The digestion reactions were run on a 2% agarose gel (Figure 4.2a), and the DNA fragments were extracted from the gel. pGex4T-1 and pMAL-c2 bacterial expression vectors were digested with Sal I and EcoR I, then 5' dephosphorylated by incubation with alkaline phosphatase at 37°C for 1hr. Linearised plasmid DNA was gel purified. The digested vector and E1^E4 fragment were used in a ligation reaction overnight at 14°C with T4 ligase. The DNA from the ligation mix was ethanol precipitated and transformed into competent E. coli cells. Ampicillin resistant transformants were selected for on LB agar plates containing 50µg/mL (w/v) ampicillin. Five bacterial colonies from each transformation were grown in 5mL LB (+50µg/mL (w/v) ampicillin) overnight, from which plasmid DNA was purified using the Wizard Plus Miniprep DNA Purification System (Promega Ltd, UK)

C) Verification of the positive clones

The purified plasmid DNA was examined by Sal I/EcoR I digestion, to confirm the presence of an approximately 300bp E1^E4 insert (Figure 4.2b). Plasmid clones containing the E1^E4 insert were sequenced, and data were analysed using the ABI PRISM DNA Sequencing Analysis Software (PE Applied Biosystems, Perkin Elmer, USA). Sequences were read in both directions across the 5' EcoR I insertion site, as well as the entire E1^E4 ORF, to ensure that the cloning was in-frame and no base changes had been introduced (Figure 4.2b).

D) Fusion protein expression

Once the in-frame insertions were confirmed, GST-E4 and MBP-E4 fusion proteins were expressed in liquid culture, and purified as described in Chapter 2. The E4 fusion proteins were purified (see Chapter 2 for the purification protocols for GST and MBP fusion proteins), and run on mini SDS-PAGE gels alongside a high molecular weight range (14.3kDa - 220kDa) rainbow coloured protein ladder (Amersham Pharmacia Biotech, UK). Besides the full length
protein, other cleaved forms of the E4 fusion proteins were also present in the preparations (Figure 4.2c).

Protein expression was optimised for both the GST and MBP fusion proteins in the following ways. Bacterial protein contaminants were reduced by using a shorter incubation time (25mins) for protein-resin binding at 4°C, followed by repeated washing with almost forty times the resin volume. Problems with proteolysis were minimised by reducing the samples handling times, as well as ensuring that all the apparatus, reagents and protein solutions were kept cooled.

**E) GST-E4 or MBP-E4 immunogens**

Bacterial expressed GST fusion proteins of BPV-1, CRPV, COPV, ROPV and HPV-11 E4 proteins were prepared for the inoculation of rabbits and rats. Once purified, the protein samples were dialysed in PBS and concentrated (refer to Chapter 2). The final concentration of each immunogen preparation was at least 0.4mg/mL (w/v) for rabbit immunisation, and 0.5mg/mL (w/v) for rat immunisation. The samples were divided into 0.5mL (rabbit immunisation) or 0.1mL (rat immunisation) aliquots, and stored at -20°C until ready to be used.

To improve the binding affinity of the GST-CRPVE4 antisera, an immunogen containing both native and heat denatured forms of MBP-CRPVE4 was prepared (see Chapter 2). This was inoculated into rats using a slightly modified immunisation protocol as described in Chapter 2.
PCR amplified E1^E4 fragments were checked on a 2% agarose gel (+EtBr) against a 1kb DNA ladder (M). B, CR, CO, RO, H-11 represents the five different PV types described in this chapter. Their amplified fragments were approximately 348, 339, 363, 432 and 300 base pairs respectively.

The E1^E4 PCR product and vector were digested with EcoR I and Sal I. The linearised DNA are showed on lanes 3 and 4. Lane 2 shows a typical ladder of ligated products following an overnight ligation with T4 ligase.

DNA from selected bacterial clones were purified and sequenced. Sequences were checked in both directions to ensure that the E1^E4 insert was cloned into the right ORF and that no other mutations were present. Sequence data on the left shows an intact EcoR I site and part of the E1^E4 sequence.

Bacterial expressed GST-E1^E4 proteins were checked on SDS-PAGE mini protein gels. Full length, cleaved and multimerised products were usually present in each protein preparation as shown by the numerous bands. The approximate sizes of the GST-E1^E4 proteins are 37.8kDa (BPV-1), 37.6kDa (CRPV), 38.6kDa (COPV), 41kDa (ROPV) and 35.8kDa (HPV-11). These protein preparations were dialysed and used for the immunisation of the rabbits or rats. Arrows - full-length GST-E1^E4 proteins.

Figure 4.2. Preparation of the E1^E4 fusion proteins used for the generation of E1^E4 specific antibodies. (a), (b) and (c) shows the main steps leading up to the production of the fusion protein immunogens, namely the synthesis of the E1^E4 DNA fragments, the validation of E1^E4-vector plasmid clones, and the expression of the E1^E4 fusion proteins.
4.2.2 Preparation of the KLH (and BSA) conjugated peptide immunogens

A) E1^E4 and E4 Peptides

Short E1^E4 or E4 peptides with an additional C-terminal cysteine amide group were synthesised for BPV-1 and CRPV. Peptide 1211 ((ac)-Ala-Asn-Asp-Lys-Ile-Ala-Gln-Thr-Glu-Ser-Gly-Cys*amide) contains the first 12 amino acid residues of the BPV-1 E1^E4 protein after the N-terminal methionine residue, and was used to generate antibodies to the N-terminus of the protein. This peptide was synthesised in two forms with or without N-terminal acetylation (ac). Peptide 1212 (Arg-Ser-Ser-Arg-Pro-Thr-Pro-Gln-Arg-Lys-Asn-Gln-Cys*amide) contains the last 12 amino acids of the BPV-1 E4 protein, and was used to generate antibodies to the C-terminus of the protein. Peptide CRE4PT2 (ac-Ala-Glu-Ala-Pro-Pro-Ser-Arg-Trp-Ser-Val-Pro-Leu-Cys*amide) contains the first 12 amino acid residues of the CRPV E1^E4 protein after the N-terminal methionine residue, and was used to generate antibodies to the N-terminus of the protein.

B) Peptide conjugation

The E4 peptides were conjugated to KLH and BSA carrier proteins. Activated KLH carrier solution was prepared fresh before each conjugation reaction by incubating reconstituted KLH with freshly prepared sulfo-MBS (m-Maleimidedobenzoyl-N-hydroxysuccinimide ester) solution at room temperature. The activated KLH carrier molecules were collected in fractions, pooled and used immediately for peptide conjugation (refer to Chapter 2 for protocol). Maleimide activated bovine serum albumin (BSA) carriers (Sigma-Aldrich Company Ltd, UK) and used according to the manufacturer's protocol. Due to the high hydrophobic content (62%) of the CRPT2 peptide, this peptide was dissolved in a small volume of dimethylformamide (DMF) before it was made up to 8mg/mL (w/v) with conjugation buffer. In contrast, peptides 1211 and 1212 were readily soluble in conjugation buffer.
The degree of peptide conjugation in each reaction was determined by the concentration of free cysteine molecules remaining after conjugation, calculated using the cysteine assay method (described in Chapter 2; Table 4.3)

<table>
<thead>
<tr>
<th>Peptides</th>
<th>Percentage KLH-peptide conjugation</th>
<th>Percentage BSA-peptide conjugation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1211 (+ ac)</td>
<td>~ 40%</td>
<td>~ 69%*</td>
</tr>
<tr>
<td>1211 (- ac)</td>
<td>~ 21%</td>
<td></td>
</tr>
<tr>
<td>1212</td>
<td>~ 50%</td>
<td>~ 66%</td>
</tr>
<tr>
<td>CRPT2 (+ac)</td>
<td>~ 90%</td>
<td>~ 60%</td>
</tr>
</tbody>
</table>

Table 4.3. Percentage of peptide conjugation as determined by the cysteine assay. Values are brought to the nearest whole number. (+/- ac) - acetylated or non-acetylated peptide. (*) - with equal amounts of acetylated and non-acetylated peptides.

C) Preparation of KLH-peptide immunogens

Only KLH-conjugated E4 peptides were used as immunogens for the generation of E4 specific antibodies. Each KLH-peptide immunogen was dialysed and concentrated before aliquoting. KLH-1211 and KLH-1212 preparations were used to inoculate rabbits, while KLH-CRPT2 was used to inoculate rats. The KLH-1211 immunogen was prepared with two parts of the KLH-1211 (+ac) and one part the KLH-1211 (-ac) conjugation reaction.

4.3 Immunisation of animals

Antisera generation was performed by Murex Biotech under specific project licences for rabbit and rat immunisation. More than one animal was used for with each immunogen to ensure that a sufficient amount of antisera was collected, and to compensate for variations in the level of immune response of different animals. Immunogens containing CRPV and ROPV E4 protein sequences were inoculated into rats so that the use of anti-rabbit secondary antibodies on infected rabbit tissues may be avoided. This will prevent any complication that may arise from
same species antibody cross-reactivity during the immunodetection experiments on the tissue sections. BPV-1, COPV and HPV-11 E4 immunogens were injected into rabbits.

All protein preparations, except MBP-CRE4, were inoculated according to a standard immunisation protocol as described in Chapter 2. Rats immunised with MBP-CRE4 were given a 21-day, instead of a 14-day, interval period between each booster injection. Pre-immune sera samples were collected before the start of the immunisation course to be used as negative controls. Up to three test bleed samples were requested for each protein approximately two months after the first injection, before the animals were sacrificed. For easy reference, each E4 antisera was re-named with an abbreviation as listed in Table 4.4.

<table>
<thead>
<tr>
<th>Immunogen</th>
<th>Animals</th>
<th>Final Bleed Antisera</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BPV-1</strong></td>
<td>GST-BPV-1E4</td>
<td>Rabbits F/G</td>
</tr>
<tr>
<td></td>
<td>KLH-1211</td>
<td>Rabbits H/J</td>
</tr>
<tr>
<td></td>
<td>KLH-1212</td>
<td>Rabbits K/L</td>
</tr>
<tr>
<td><strong>CRPV</strong></td>
<td>KLH-CRPT2</td>
<td>Rats 1-6</td>
</tr>
<tr>
<td></td>
<td>GST-CRPVE4</td>
<td>Rats 13-18</td>
</tr>
<tr>
<td></td>
<td>MBP-CRPVE4</td>
<td>Rats 25-30</td>
</tr>
<tr>
<td><strong>COPV</strong></td>
<td>GST-COPVE4</td>
<td>Rabbits A/B</td>
</tr>
<tr>
<td><strong>ROPV</strong></td>
<td>GST-ROPVE4</td>
<td>Rats 19-24</td>
</tr>
<tr>
<td><strong>HPV-11</strong></td>
<td>GST-HPV-11E4</td>
<td>Rabbits C/D</td>
</tr>
</tbody>
</table>

Table 4.4. Abbreviations used to denote the immunogens, animals and antisera during the generation of the E4 specific antibodies. Rabbits or rats were referred as (Rb) or (r) respectively. Each rabbit was given a representative alphabet (A or B), and each rat a representative number (1-30). Pre-immune sera were labelled as PIm, test bleed samples as TBn (e.g. TB2), where n is the order of test bleed collected, and the final bleed antisera as FB.
4.4 E4 specific antisera

Antisera specificity for the respective E4 immunogen was determined by western blotting, followed by ELISA. For each animal, only the Plm sera, last TB sample and FB were tested. MBP-E4 fusion proteins and BSA-peptide conjugates were used as positive controls to test specific E4 binding affinity of the GST-E4 and KLH-peptide induced antisera. GST-CRPVE4 and BSA-CRPT2 were used as positive controls for the MBP-CRPVE4 generated antisera. MBP, GST and BSA protein solutions were prepared as negative controls.

4.4.1 E4 specificity on western blots

Specific E4 binding was found in the antisera collected from all the immunised animals. Positive E4 reactivity was visualised by the deposition of the insoluble Fast-DAB horse-radish peroxidase (HRP) substrate (Sigma-Aldrich Company Ltd, UK) on western blots with specific E4 antigen tracks. A dilution of 1:250 of the Plm, TB, and FB antisera from each animal was used to probe the E4 antigens on the nitrocellulose membranes. HRP conjugated anti-rabbit (Amersham Pharmacia Biotech, UK) or anti-rat (Amersham Pharmacia Biotech, UK) monoclonal antibodies, were used to detect antibody binding to the E4 fusion proteins (Figure 4.5). No binding activity was detected in the protein lanes containing the GST or MBP proteins.
Figure 4.5. E4 specific binding on western blots.
(a) Shows a three western blot set-up typically used to test E4 specific antibodies of the rabbit antiserum. The set of results show the PIm (pre-immune) sera, and TB (test-bleed) 2 and 3 of rabbit A, which was inoculated with GST-COPVE1^E4 (G-COE1^E4). The PIm sera did not show immuno-reactivity to all the antigens. E1^E4 specific antibodies were shown by substrate deposition bands (dark grey) on the MBP-COE1^E4 (M-COE1^E4), but not the MBP protein track. M - prestained protein ladder. (b) Shows a similar western blot set-up for testing E1^E4 reactive rat antiserum. Due to the large number of rat samples collected, only two protein samples were used for each blot. In this case, antibodies against E1^E4 was detected from the test bleeds taken from rat 27, following the immunisation of MBP-CRPVE1^E4. The testbleeds (TB) but not pre-immune (PIm) sera, showed immuno-reactivity to GST-CRE1^E4 (G-CRE1^E4) but not GST alone. RM - prestained rainbow protein marker.
4.4.2 E4 specificity and titre with ELISA

E4 binding specificity and affinity of the collected antisera (except anti-GST-COPVE4 sera) were measured using ELISAs. Serial dilutions (1:20 - 1:10240) of the Plm sera and FB antisera were tested against specific E4 containing antigens, or negative controls. The colourmetric development of the HRP substrate, tetramethylbenzidine dihydrochloride (TMB) (Sigma-Aldrich Company Ltd, UK), was used as an indicator of the level of specific antibody binding. Typical sigmoidal antibody binding curves were plotted from the spectrophotometer values at 450nm against the dilution factor of the antisera (Figures 4.6 & 4.7). No two antisera samples produced the same antibody binding curve, even when animals were immunised with the same immunogen. Antisera showing strong E4 specific reactivity on western blots were not always reflected in the ELISAs, for instance in the case of GST-CRPVE1^E4, GST-HPV-11E1^E4 antisera samples (data not shown). Fortunately, strong and specific antibody activities were shown from the ELISA data for the antisera samples immunised with GST-BPV1E1^E4, GST-ROPVE1^E4, MBP-CRPVE1^E4, KLH-CRPVPT2 and KLH-1212 (Figures 4.6a, e & d, and Figures 4.7c & b).
Chapter 4 - E4 Specific Antibodies

Figure 4.6. ELISA binding curves of antisera against GST or MBP fusion E1\^E4 proteins. Graphs were plotted with the absorbance values at OD450 versus the antisera dilution factor. (a) Anti-GST-BPV1\^E4 antisera (FB from rabbits F & G) against MBP-BPV1\^E4 & MBP (RbF\_M-B1E4 & RbG\_M-B1E4). Good E4 specific immunity were elicited in both rabbits. (b) Anti-GST-HPV11\^E4 antisera (FB from rabbits C & D) against MBP-HPV11\^E4 & MBP (RbC\_M-H11E4 & RbD\_M-H11E4). The curves indicated poor E4 binding affinities from both rabbits. (c) Anti-GST-CRPVE1\^E4 antisera (FB from rats 13-18) against MBP-CRPVE1\^E4. All six antisera showed weak E4 binding affinity and specificity. (d) Anti-MBP-CRPVE1\^E4 antisera (FB from rats 25-30) against GST-CRPVE1\^E4. In comparison, these antisera showed much stronger and specific E4 binding affinities than the earlier immunisation with GST-CRPVE1\^E4. (e) Anti-GST-ROPVE1\^E4 antisera (FB from rats 19-24) against MBP-ROPVE1\^E4. Specific ROPV E4 binding activities were found in all six antisera.
Chapter 4 - E4 Specific Antibodies

c) r13-18 FB and r14 & r15 Pl against MBP-CRPVE1^E4

![Graph](image)

d) r25-30 FB and r25 & r27 Pl against GST-CRPVE1^E4

![Graph](image)

e) r19-24 FB and Pooled Pl against MBP-ROPVE1^E4

![Graph](image)
Figure 4.7. ELISA binding curves of antisera against KLH-peptide immunogens. Graphs were plotted with absorbance readings at OD_{450} versus the antisera dilution factor. (a) Anti-KLH-1211 antisera (FB from rabbits H and J) against MBP-BPVE1^E4 (RbH_M-B1E4 & RbJ_M-B1E4) and MBP (RbH_MBP & RbJ_MBP). The curves indicate poor E4 binding specificity in both antisera. (b) Anti-KLH-1212 antisera (FB from rabbits K and M) against MBP-BPVE1^E4 (RbK_M-B1E4 & RbM_M-B1E4) and MBP (RbK_MBP & RbM_MBP). Specific E4 antibodies were generated in both rabbits. The curves show that RbKFB contains E4 antibodies binding to the E4 antigen at a higher affinity than RbMFB. (c) Anti-KLH-CRPT2 antisera (FB from rats 1-6) against MBP-CRPVE1^E4 (r1-M-CRE4) and MBP (r1_MBP). All the rat antisera showed highly specific E4 binding activities, but with variable affinities.
4.5 Discussion

Specific polyclonal antibodies were generated for BPV-1, CRPV, COPV, ROPV and HPV-11 E4 proteins following the immunisation of rabbits or rats. Specific E4 reactivity was shown against the E4 part of the immunogen on western blots and in ELISAs.

The five most commonly used animals for the production of antisera are rabbits, mice, rats, hamsters, and guinea pigs, of which rabbits are considered the most favourable choice. This is because rabbits have been shown to generate good immune responses against immunogens when they administered with adjuvants. In addition, rabbits are readily available and economical to purchase, and are able to produce a large amount of antisera at the end of the immunisation course.

Bacterially expressed, non-denatured GST and MBP E4 fusion proteins were used as immunogens for antibody production as they have previously been used successfully in our laboratory to generate specific antibodies against HPV-1, HPV-2, HPV-16 and HPV-65 E4 proteins. Heat denatured MBP-CRPV E4 proteins were used in a latter immunisation protocol when the anti-GST-CRE4 antibodies showed poor binding affinity on infected rabbit tissues. In some cases, protein denaturation has been showed to increase the immunogenicity of protein antigens by exposing new epitopes to the host immune system (Harlow & Lane, 1988). In addition, protein aggregates produced as a result of heating have been shown to be more immunogenic than soluble antigens (Harlow & Lane, 1988). An extended immunisation protocol was used for the inoculation of MBP-CRE4 as this was thought to maximise the effective dose of the booster injections, which might be reduced by the presence of moderate levels of circulating antibodies 2 weeks post-injection (Harlow & Lane, 1988). Indeed, E4 antibodies obtained from the immunisation of MBP-CRE4 showed a higher binding affinity to the antigen in infected tissue sections than the GST-CRE4 antisera.
Chapter 4 – E4 Specific Antibodies

The KLH conjugated E4 peptides also proved to be effective in generating specific E4 antibodies. The main advantage of using short peptide sequences as immunogens was the ability to obtain antibodies which recognised epitopes within a defined region of the E4 protein. The CRPVPT2 antibodies have proved to be particularly useful for the detection of the CRPV E4 in infected tissue sections (Chapter 5). In addition, since the E1^E4 protein has been shown in some lesions to be progressively cleaved at the N-terminus as the infected cells move towards the surface of the lesion (Doorbar et al., 1988), peptide antibodies against the C-terminal of the BPV-1E4 protein may be useful for detecting all possible forms of the E4 protein in lesions.

Valuable advice was given by Dr Tony McGee (formerly from the Division of Membrane Biology, NIMR) and members of his laboratory, regarding the post-translational modification of proteins in vivo. Since some proteins are known to undergo N-terminal modifications in cells, it was useful to know if the N-terminus of E4 proteins would be subjected to specific chemical modification in vivo. If so, appropriate chemical modifications could be made to the E1^E4 peptides immediately following synthesis, so that the E1^E4 peptide might closely mimic the in vivo E4 gene product. The type of post-translational chemical modification can be predicted by the N-terminal primary amino acid sequence of the protein (Utsumi et al., 2001). Based on the N-terminal amino acid sequence of CRPV and BPV-1 E4 proteins, the predicted post-translational modification involved the replacement of the first methionine residue with an acetyl (ac) group at the N-terminus of the protein. Although it was not certain if these E4 proteins were in fact modified in such a way in vivo, antisera generated against the acetylated CRPT2 and 1211 peptides showed positive binding activities to the E4 protein in infected tissue sections, on western blots and in ELISAs.
CHAPTER 5 - CHARACTERISATION OF THE EVENTS IN THE LATE STAGE OF THE VIRAL LIFE CYCLE: E4 AND L1 PROTEINS EXPRESSION AND VIRAL DNA AMPLIFICATION PATTERNS IN FORMALIN-FIXED, PARAFFIN-EMBEDDED PAPILLOMAVIRUS INFECTED TISSUES

5.1 Introduction ............................................................................................................... 147
5.2 Infected tissue specimens ...................................................................................... 149
5.3 Staining protocols and specifics ............................................................................. 150
5.4 E4 antibody staining and the general distribution of the E4 protein in infected tissue sections ....................................................................................................... 150
5.5 Viral DNA amplification and E4 expression correlate exactly in infected tissues 155
5.6 E4 expression proceeds L1 in all human and animal PV infections and model systems ................................................................. 160
5.7 E4 shows different intracellular expression patterns in lesions caused by different PVs ............................................................................................................. 162
5.8 The viral life cycle in mouse xenograft systems .................................................... 167
5.9 Discussion ................................................................................................................ 170
5.1 Introduction

The late stage of the PV life cycle can be characterised by three major events, namely viral DNA amplification, expression of the viral late proteins (E4, L1 and L2), and virus particle assembly and release. Since the life cycle of the virus is dependent on differentiation of the epithelium, the occurrence of these events are intricately linked to the transition of the host cells. Previously, it has been shown that the onset of viral DNA amplification and E4 protein expression correlated exactly in HPV-1 and HPV-16 infected tissues (Doorbar et al., 1997). Production of L1 capsid proteins was triggered after viral genome amplification, but only in E4 positive cells. Although the timing of the late events in the cutaneous HPV-1 papillomas and mucosal HPV-16 lesions were vastly different, the general pattern of viral DNA amplification and E4 correlation, and delayed L1 expression, was conserved. Thus, the late stage of the virus life cycle can be divided into two phases, namely DNA amplification and virus particle formation. In order to study the viral late events, specific labelled-DNA probes and E4 antibodies were made for HPV-1 and HPV-16, and commercial L1 antibodies were used for the detection of the viral major capsid protein. In this chapter, I have expanded on this report and analysed the late events in other PV types, mainly animal PVs. PV type specific labelled DNA probes and E4 antibodies (Chapter 4) were generated for BPV-1, COPV, CRPV, ROPV and HPV-11, and L1 antibodies were obtained from several external sources. Viral DNA amplification was detected by DNA-DNA in situ hybridisation (ISH), and viral protein expression by fluorescent immuno-histochemistry. By examining a wider range of viruses, a more comprehensive assessment can be made between conserved events in the late stage of the viral life cycle in animal and human PV infected tissues.

The use of animal PVs in many aspects of PV research has partially compensated for the problems involved in working with human viruses. Animal model systems has provided numerous advantages for the study of the virus infection in the natural host, particularly in immunological studies during papilloma development, regression, and persistence as well as
progression to malignancy. Infections in the animals can be constantly monitored and can be used to study the effects of molecularly modified DNA or proteins. For these reasons, animals are heavily utilised for vaccination and pharmacological studies and are essential during the early stages of vaccine and medicine development. Data collected from studies in these animal models are very often applied to HPV infections. However, no actual study has been conducted to compare the life cycle events between animal and human PVs. The most widespread animal PVs used in laboratories today are BPV-1, COPV, CRPV and ROPV. Tissue samples infected with these viruses were collected here and used in comparative studies of late events in the virus life cycle. These included naturally or experimentally infected tissue samples, which were either virus or DNA-induced. In addition, xenograft systems which have been shown to support the life cycle of various PV types had also been included, such as the kidney capsule implants in NUDE mice, and the skin implants in SCID mice.

Problems with inconsistency and inconvenience have been a difficulty in the study of mucosal HPVs in laboratories. The life cycle of these viruses can now be reconstructed in vitro in organotypic cultures, as well as in murine implants. However, these systems are not only laborious to perform, but also difficult to reliably reproduce. In this chapter, I will also compare the relevance and suitability of using animal systems as a model for the study of mucosal type HPVs. For this reason, HPV-11 infected materials were also included as part of this comparison study.
### 5.2 Infected tissue specimens (Table 5.1)

<table>
<thead>
<tr>
<th>Virus</th>
<th>Host</th>
<th>Tissue type</th>
<th>Fixation</th>
<th>Origin</th>
<th>No. of lesions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPV-1</td>
<td>Bovine</td>
<td>Penile fibrilloma, cutaneous.</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>N.I.</td>
<td>4</td>
<td>U.S.A.</td>
</tr>
<tr>
<td></td>
<td>Bovine</td>
<td>Penile fibrilloma, cutaneous.</td>
<td>Frozen, then formalin-fixed, paraffin-embedded.</td>
<td>N.I.</td>
<td>4</td>
<td>Wen-Jun, Australia.</td>
</tr>
<tr>
<td>CRPV</td>
<td>NZW rabbit</td>
<td>Cutaneous</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>&gt;15</td>
<td>Brandsma, U.S.A.</td>
</tr>
<tr>
<td></td>
<td>CT rabbit</td>
<td>Cutaneous</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>2</td>
<td>Christensen, U.S.A.</td>
</tr>
<tr>
<td></td>
<td>Mouse</td>
<td>CRPV infected kidney capsule xenograft</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>1</td>
<td>Christensen, U.S.A.</td>
</tr>
<tr>
<td></td>
<td>Mouse</td>
<td>CRPV infected skin xenograft</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>2</td>
<td>Brietburd, France.</td>
</tr>
<tr>
<td></td>
<td>Beagle</td>
<td>Gingival, mucosal.</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>1</td>
<td>Nicholls, U.K.</td>
</tr>
<tr>
<td></td>
<td>Beagle</td>
<td>Oesophagus, mucosal.</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>3</td>
<td>Nicholls, U.K.</td>
</tr>
<tr>
<td></td>
<td>Beagle</td>
<td>Oral lesions, weekly time course</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>10</td>
<td>Nicholls, U.K.</td>
</tr>
<tr>
<td></td>
<td>Mouse</td>
<td>Penile, skin xenograft ROPV infected kidney capsule xenograft</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>1</td>
<td>Christensen, U.S.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E.I.</td>
<td>2</td>
<td>Christensen, U.S.A.</td>
</tr>
<tr>
<td>SCID</td>
<td>Mouse</td>
<td>HPV-11 infected skin xenograft</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>3</td>
<td>Roche, U.K.</td>
</tr>
<tr>
<td></td>
<td>Mouse</td>
<td>HPV-11 infected kidney capsule xenograft</td>
<td>Formalin-fixed, paraffin-embedded.</td>
<td>E.I.</td>
<td>3</td>
<td>Roche, U.K.</td>
</tr>
</tbody>
</table>

Table 5.1. Types and sources of infected tissue samples used for immuno-staining. All tissues were prepared as formalin-fixed and paraffin embedded samples. E.I. = experimentally infected, and N.I. = naturally infected.
5.3 Staining protocols and specifics

Tissue sections were double stained for E4 and L1 or E4 and viral DNA, by combinations of protocols listed in Chapter 2, although positive staining for both antigens was not possible in all cases. In most cases, specific antigen exposure steps were required for the detection of E4 in the formalin-fixed paraffin-embedded tissue sections. Various different methods and reagents were used here, including microwaving, pressure-cooking and protease digestion, as well as different heat treatment buffers. In some occasions, combinations of antigen exposure methods were also explored in order to improve staining results.

Most antibody and ISH detections were visualised using immuno-fluorescence. To further increase the sensitivity of the protocols, tyramide fluorophores (HRP-substrates; NEN Life Science Products Incorporated, USA) were used to detect ISH signals, and occasionally for immuno-histochemistry. Alternatively, increased sensitivity was also achieved by using Fast-Red (Sigma-Aldrich Company Ltd, UK), a substrate for alkaline phosphatase (AP) activity.

5.4 E4 antibody staining and the general distribution of the E4 protein in infected tissue sections

E4 has been shown to be expressed in abundance and in high levels in HPV-1 papillomas (Doorbar et al., 1988). In this study, specific E4 antibodies were generated against five other PV types (Chapter 4) and used to detect the viral proteins in formalin-fixed paraffin-embedded tissue sections. Besides HPV-11 antibodies which have been shown to cross-react with HPV-6 E4 proteins on ELISA, all other E4 antibodies did not show cross-reactivity within our range of test proteins. To ensure that the positive E4 stains were specific and not due to the
The E4 staining varied in intensity and distribution in BPV-1, COPV, CRPV, ROPV and HPV-11 infected tissue samples. E4 expression could be triggered in the basal layer, as seen in the case of COPV, or much higher in the differentiated epidermis, for instance in CRPV and HPV-11 infections. Staining patterns also varied from sporadic staining throughout the lesion (BPV-1 and COPV) to expression in almost every cell in specific regions of the tissues (CRPV, ROPV and HPV-11).

COPV and ROPV E4 could be easily detected in the tissue sections without any antigen exposure treatment. RbA anti-COPV E4 antibodies were used at 1:600 at room temperature for approximately one hour on the sections. Rat18 antisera (1:400) was used to detect ROPV E4 at room temperature for approximately an hour. However, in some cases, an antigen exposure step prior to immuno-detection was required for the detection of E4 in the formalin-fixed tissues. Various antigen exposure methods were used. For HPV-11 a two-step treatment was required. Formalin-fixed tissues were treated in the microwave in a citrate buffer (pH6.0) followed by a brief trypsin enzyme digestion step before staining could be obtained. According to the ELISA data (Chapter 4), the binding affinity of HPV-11 E4 antisera to the antigen was low and 1:80 (RbC antisera) had to be used for at least an hour at room temperature to achieve satisfactory signals. For BPV and CRPV, E4 staining on the cutaneous warts was difficult to achieve and reproduce. BPV E4 was rarely detected in BPV-1/BPV-2 bovine warts even after various combinations of antigen exposure treatments, including heat and enzyme treatments or combinations of both, and the use of antisera of different specificity and from different animals. In the case of the CRPV infected cottontail rabbit tissues, E4 was detected only after sections were heat-treated by either microwave or pressure cooking in a modified citrate buffer (pH8.0). It was also found that background staining could be reduced when
immuno-detection was carried out at 37°C instead of room temperature. Antibodies generated from three different preparations of CRPV E4 were found to give the same E4 staining patterns, but with variable positive staining and background levels. Best staining results were obtained with rat antisera generated against MBP-CRPVE1^E4 (rat30) and KLH-PT2 (rat2), used at 1:100 dilution.

Positive E4 staining on the tissues showed a heterogeneous E4 distribution pattern in the different animal and human tissues. These are described and shown in Figure 5.2.
Chapter 5 – Characterisation of the Events in the Late Stage of the Viral Life Cycle

(a)

<table>
<thead>
<tr>
<th>PV type</th>
<th>E4 staining distribution patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mucosal type virus</strong></td>
<td></td>
</tr>
<tr>
<td>COPV</td>
<td>Abundant E4 was detected in mature COPV infected oral lesions in 50-60% of cells. E4 expressing cells usually appeared rounded, and were distributed throughout the lesions in sporadic clusters. E4 expression in the basal layer is a unique feature of COPV infections. This staining pattern is typical in all 21 lesions examined.</td>
</tr>
<tr>
<td>ROPV</td>
<td>High levels of E4 were expressed in almost every cell in the ROPV infected areas. E4 was first detected in the parabasal layers, where E4 staining extended to the top of the lesion. This staining pattern is typical in all 16 lesions examined.</td>
</tr>
<tr>
<td>HPV-11</td>
<td>The extent of E4 staining was variable between lesions. Typically, E4 would be detected only in specific areas of the HPV-11 lesions. E4 staining was first seen weakly in sporadic cells in the intermediate layers of the epithelium and increased in intensity and distribution in the layers above the region. This staining pattern was seen in all 8 lesions examined, but was best shown in the cervical lesions and infected SCID mouse xenografts.</td>
</tr>
<tr>
<td><strong>Cutaneous type virus</strong></td>
<td></td>
</tr>
<tr>
<td>BPV-1</td>
<td>E4 was detected in sporadic cell clusters throughout the epidermal layer of the BPV-1/BPV-2 infected lesions. It was thought that E4 staining started as low as the immediate parabasal layer and extended to the edge of the papilloma. This pattern was seen 2 out of 4 formalin-fixed paraffin-embedded lesions examined.</td>
</tr>
<tr>
<td>CRPV</td>
<td>Like HPV-11, CRPV E4 was detected in specific regions in the cutaneous rabbit papilloma tissues. E4 was detected from the spinous layer and extended to granular layer. The abrupt loss of E4 staining in the uppermost layers of the papilloma was thought to be attributed by the loss of antigenic epitopes in the E4 protein so that our CRPV E4 antibodies could no longer detect the protein. This may be due to protein modifications, cleavage or degradation in terminally differentiated layers. This staining pattern was seen in 3 out of 5 CT rabbit lesions examined.</td>
</tr>
</tbody>
</table>

Figure 5.2. General distribution patterns of the viral E4 protein in infected tissue sections as detected by specific E4 antibodies. (a) Describes the E4 distribution patterns in tissues infected by COPV, ROPV, HPV-11, BPV-1/BPV-2 and CRPV. (b) Shows the immuno-staining patterns of E4 in tissues sections infected by (I) mucosal type viruses, and (II) cutaneous type viruses. Sections were immunostained with specific E4 antibodies for each PV type. E4 proteins were detected throughout the lesions in sporadic clusters in COPV and BPV-1 infected lesions, whereas ROPV, HPV-11 and CRPV E4 expression was confined to specific regions. The timing of E4 expression differed between the different PV types. Basal layer expression of E4 was unique to COPV infections. ROPV and BPV-1/BPV-2 E4 were detected from the parabasal layers, while HPV-11 and CRPV E4 were first detected in the intermediate layers. The position of the basal layer is marked with a dotted line. ( ) – represents the microscope objective magnification used to take the picture.
(b)

I - Mucosal Type Viruses

II - Cutaneous Type

BPV (10x)  CRPV (10x)
5.5 Viral DNA amplification and E4 expression correlate exactly in infected tissues

PVs are highly tissue and host specific and each virus type is assumed to have adapted specifically to its particular environment. However, the viral life cycle is thought to be conserved even between different virus types. For instance, the viral late events were shown to occur in a similar pattern in HPV-1 and HPV-16 infected tissue sections (as described earlier; Doorbar et al., 1997).

In double stained tissue sections, a correlation between viral DNA amplification and E4 expression was observed. This correlation pattern is now considered a general property of the late stage of the PV life cycle. A close correlation was also seen between the levels of E4 and viral DNA in these cells.

**Mucosal type viruses**

In all the COPV lesions, amplified viral DNA and E4 protein were detected from the basal layer to the top of the lesion throughout the papillomas (also described in Nicholls et al., 2001). High levels of E4 expression and viral DNA amplification were easily correlated in double stained oral, gingival and oesophageal lesions (Figure 5.3a). However, exceptions where this correlation did not occur were also found in both naturally and experimentally induced COPV lesions. Figure 5.4a shows cells which contained high levels of viral DNA but no E4, and vice versa. The occurrence of these cells were rare and experimental variability could not be ruled out as a contributing factor. In addition, mis-timed or ‘slipped’ activities within the host cells might be the other causative factors that led to unregulated activation of the individual viral events.

In rabbit tongue and penile tissue sections, high levels of viral DNA and E4 were detected in almost every cell nuclei in the ROPV infected areas. The distribution of the two
events coincided exactly. Onset of these late events began in the lower intermediate or mid-spinous layers, and extended to the surface of the epidermis. Levels of viral DNA and E4 increased simultaneously in these areas as the cells continued to move upwards. This indicated the activation of viral late promoter activity as the cells migrated through the differentiating epidermis, leading to the accumulation of viral late products in the higher layers. In one exception, low levels of viral DNA was detected in the basal and parabasal layers in a small isolated area within the penile lesion. E4 and stronger viral DNA signals was detected only one to two layers above this weak replication zone (Figure 5.4b). One possible explanation for this is the detection of increased viral DNA replication in the lower layers during the early stage of papilloma formation. Viral DNA may have increased several folds to levels that are detectable by FISH.

Although only a limited number of HPV-11 samples were obtained for this study, the onset of the viral late events appeared to be variable in the different tissue samples. From the staining, viral DNA and E4 could be detected from as low as the intermediate layers, to a few layers from the top of the epidermis. Like ROPV, the distribution patterns and levels of the immuno-stains correlated exactly in these lesions (Figure 5.3c).

**Cutaneous type viruses**

Unfortunately, a double stain of E4 and viral DNA could not be done on the BPV-1 and -2 infected tissues. In this case, individually stained sections were compared instead. Both the viral DNA and E4 staining pattern appeared scattered and sporadic in the bovine warts (Figure 5.3d). Although an exact alignment of the adjacent tissue stained with viral DNA and E4 were not possible, similarities in their general distribution patterns suggested that both events were likely to correlate in these tissues.
Chapter 5 – Characterisation of the Events in the Late Stage of the Viral Life Cycle

In CRPV infected cottontail rabbit warts, viral DNA and E4 were first detected in the spinous layer of the epidermis. The late stage of the viral life cycle appeared to be supported in specific regions of the entire papilloma tissue. As described previously, CRPV E4 could not be detected in the upper layers of the epidermis using two different antibodies, probably due to the loss of the specific antigenic epitope in E4. However, viral DNA amplification and E4 expression still showed an exact correlation pattern in the lower layers (Figure 5.3e).
Chapter 5 - Characterisation of the Events in the Late Stage of the Viral Life Cycle

(a) Mucosal Type Viruses

Viral DNA amplification (red) was detected by in situ hybridisation (ISH) using DIG-labelled DNA probes, and E4 expression (green) by specific antibodies by immunohistochemistry. Exact correlation of the two events were observed in double immuno-stained COPV, ROPV and CRPV tissue sections. Adjacent HPV-11 tissue sections were used to correlate the distribution patterns of viral DNA amplification and E4. Both events were found to be located in the same regions of a genital lesion biopsy. Due to the poor and inconsistent BPV-1 E4 immuno-stains, the sporadic distribution pattern of viral DNA and E4 (see Figure 5.2) suggested that the two events could be correlated in the infected tissues. Basal layers are represented by dotted lines, m - represents the merged image of DNA, E4 and DAPI triple stains, except in the case of the BPV-1/m image which shows a DNA and DAPI double stain. Images were taken using a 10x magnification objective, except for ROPV which was taken using a 20x magnification objective.

Figure 5.3. Viral DNA amplification and E4 expression correlate exactly in infected tissues. Viral DNA amplification (red) was detected by in situ hybridisation (ISH) using DIG-labelled DNA probes, and E4 expression (green) by specific antibodies by immunohistochemistry. Exact correlation of the two events were observed in double immuno-stained COPV, ROPV and CRPV tissue sections. Adjacent HPV-11 tissue sections were used to correlate the distribution patterns of viral DNA amplification and E4. Both events were found to be located in the same regions of a genital lesion biopsy. Due to the poor and inconsistent BPV-1 E4 immuno-stains, the sporadic distribution pattern of viral DNA and E4 (see Figure 5.2) suggested that the two events could be correlated in the infected tissues. Basal layers are represented by dotted lines, m - represents the merged image of DNA, E4 and DAPI triple stains, except in the case of the BPV-1/m image which shows a DNA and DAPI double stain. Images were taken using a 10x magnification objective, except for ROPV which was taken using a 20x magnification objective.
Figure 5.4. Rare events during the late stage of the viral life cycle.
(a) Cells with high levels of viral DNA by in situ hybridisation (ISH) but no detectable E4 proteins were rarely detected in COPV infected oral lesions (arrows). Others with E4 expression but no viral DNA amplification were also found (arrowheads). The levels of viral DNA and E4 did not always correlate in some areas of the lesion. (b) A region in the ROPV infected penile lesion showed moderate levels of viral DNA but no E4 expression in the basal and parabasal layers of the epithelial (arrow). This region could represent increased levels of virus replication, which preceded the activation of the late events (viral DNA amplification and E4 expression). All images were taken using a 40x magnification objective.
5.6 E4 expression precedes L1 in all human and animal PV infections and model systems

Recently, L1 expression was showed to be triggered after the expression of E4 in HPV-1 and HPV-16 infected tissue sections (Doorbar et al., 1997). L1 was detected only in the nuclei of differentiated cells that were also positive for E4, although not all E4 positive cells would express L1 in the terminally differentiated layers.

This pattern of sequential viral late protein expression was observed in all the animal and human tissues examined in this study. A spatial gap was always present between the first expression of E4 and the first appearance of L1, even though the size of this gap was variable in different PV infected tissues samples.

In BPV, COPV, and ROPV lesions, where the late events were supported in the lower layers of the epidermis, large gaps of up to four to six cell layers separated the first appearance of E4 and that of L1. In comparison, smaller gaps of two to four layers were found between the expressions of E4 and L1 in CRPV and HPV-11 induced lesions (Figure 5.5). Although both E4 and L1 have functional roles in the late stage of the viral life cycle and can be translated from the same bi-cistronic mRNA message E1^E4^L1 (Nasseri et al., 1987; Palermos-Dilts et al., 1990; Hummel et al., 1992), the expression of the proteins appears to be regulated differently. This observation is in contrast to an earlier report by Brown et al. (1992), which described a correlation in the expression pattern of E4 and L1 in HPV-11 infected condylomas.
Figure 5.5. E4 expression precedes the expression of L1 in human and animal PV infections. Specific E4 and L1 antibodies were used to detect the antigens on the same tissue section. L1 is usually detected in the terminally differentiated layers of the epithelium after E4 expression. A spatial gap of variable depth is usually present between the start of E4 and L1 expression. E4 - green; L1 - red; Nuclei - DAPI blue. Basal layers are represented by dotted lines.
5.7 E4 shows different intracellular expression patterns in lesions caused by the different PVs

So far, several different E4 cellular localisation patterns have been reported in HPV infected warts. Most have reported a predominant cytoplasmic localisation while others detected E4 in the nucleus and cytoplasm, and some showed E4 to be membrane associated (Doorbar et al., 1988 & 1997; Rogel-Gaillard et al., 1993; Bryan & Brown, 2000). The heterogeneity of the E4 expression patterns is interesting since the functional role of the protein is still largely unknown, although it is assumed that it serves a common function even in the different PV infections.

The intracellular distribution patterns of E4 were vastly different in the COPV, CRPV, and ROPV infected tissues. Each PV E4 protein showed a characteristic intracellular expression pattern and differed in their ability to associate to host membranes or structures, form intracellular structures, and translocate within the host cells.

**Mucosal Type Viruses**

COPV E4 showed the most heterogenous intracellular distribution. E4 could be detected in the cytoplasm as well as the nuclei, although the level of protein was usually much higher in the cytoplasm. Cytoplasmic E4 appeared diffuse, but was also able to form cytoplasmic structures that resemble inclusion granules in some cells. Occasionally, cells with increased levels of E4 at the cell and nuclear membranes were found. It was not certain if the proteins were directly associated with the membranes or had accumulated at the membrane periphery. When treated with protease, COPV E4 was detected in the nucleoli. Attempts to do a nucleolin-E4 double stain were not successful as the available anti-human nucleolin antibody (CN Biosciences Ltd, UK) did not cross-react with the canine protein. E4 nucleolar staining was confirmed by phase contrast confocal microscopy. A surprising finding in COPV infected lesions
was the presence of keratohyalin granules in the terminally differentiated epithelium, as normal oral mucosal epithelial did not contain these structures. However, it was found that the keratohyalin granules were not present in the E4 expressing cells (Figure 5.6a).

ROPV E4 was predominantly cytoplasmic and able to form cytoplasmic inclusion structures that were similar to those seen in the canine oral tissues and in cutaneous HPV-1 infected tissues. A gradual increase in the size of the inclusion granules was observed as cells moved towards the top of the lesions (Figure 5.6b). This characteristic was previously described for the HPV-1 E4 inclusions in infected warts, but was not observed in COPV E4 inclusions (Rogel-Gaillard et al., 1993; Egawa et al., 1993; Egawa, 1994).

HPV-11 E4 was found to be predominantly cytoplasmic and do associate with and accumulate at the cell membrane in the higher layers of the mucosal epithelial (as described by Bryan and Brown, 2000). Further analysis of the intracellular distribution of HPV-11 E4 was not investigated here.

Cutaneous Type Viruses

Detailed studies of the BPV-1 E4 in cells could not be performed as the quality of the immuno-staining could not be optimised. However, from the staining that was carried out, it was thought that BPV-1 E4 was predominantly cytoplasmic.

In the cottontail rabbit papillomas, CRPV E4 seemed to be expressed at high levels in both the cytoplasm as well as nuclei. Although E4 did not appear to associate with any cellular structures, E4 positive cells were spotted with numerous dark cytoplasmic structures. Hematoxylin and eosin stained sections revealed an extensive presence of densely stained, rounded granules, variable in size and number in the differentiated layers. These were thought to be a mixture of melanin and keratohyalin granules, and the cause of the numerous non-
staining 'holes' seen in E4 positive cells. Therefore, unlike the E4 proteins of COPV and HPV-1, CRPV E4 and profilaggrin were thought to be expressed in the same cells (Figure 5.6c).
Chapter 5 - Characterisation of the Events in the Late Stage of the Viral Life Cycle

Figure 5.6. The different intracellular localisation patterns of E4 in (a) COPV, (b) ROPV and (c) CRPV infected lesions.

(a) COPV E4 (green) was detected in the cytoplasm and nucleus, and also showed cellular and nuclear membrane association in some cells (arrow in top panel [COPV/E4]). In the immunofluorescent images, the cell nuclei were counterstained with either propidium iodide (red, top panel) or DAPI (blue, middle and bottom panels). Cytoplasmic E4 sometimes gave a granular appearance in some lesions (arrows in middle panel [COPV/E4]). Canine lesions stained with haematoxylin and eosin (COPV/H+E in the top and middle panels) showed presence of keratohyalin granules (KH in the top panel) in the infected lesion which are not found in normal canine oral epithelium (inset in top panel). Under higher magnification (COPV/H+E in the middle panel), permissive cells (P) that are enlarged and rounded in the infected lesion, are packed with inclusion granules which may have attributed to the granular staining pattern of COPV E4. NP represents the non-permissive cells. COPV E4 was also seen to be localised at the nucleoli (bottom panel in COPV/m). This unique E4 pattern was confirmed using phase contrast confocal microscopy (COPV/m and COPV/phase in the bottom panel where the NUC represents the nucleoli) as well as in E4 immuno-peroxidase stained (brown) sections (COPV/E4(DAB)). (b) ROPV E4 (red) is predominantly cytoplasmic. Nuclei were counterstained with DAPI (blue). Inclusion granules (g) similar to those seen in cutaneous lesions caused by HPV-1 are also found in the infected rabbit tongue lesions. Inclusion granules in the haematoxylin and eosin stained (ROPV/H+E) section correlates the formation of the E4 granules in the differentiated layers of the epithelium. (c) CRPV E4 (green) showed an equal distribution pattern in the cytoplasm and nucleus. Nuclei were counterstained with DAPI (blue). Dark granules (arrows in CRPV/E4) of variable sizes and shapes are scattered in the cytoplasm. These are thought to be keratohyalin granules (KH), seen in the haematoxylin and eosin stained sections (CRPV/H+E). All images were taken using a 40x objective, unless otherwise stated.
Chapter 5 - Characterisation of the Events in the Late Stage of the Viral Life Cycle

(a) COPV E4

(b) ROPV E4

(c) CRPV E4
5.8 The viral life cycle in mouse xenograft systems

_In vitro_ cell culture systems and transplantation techniques have been employed mainly for the study of human PV infections. In these model systems, some viruses have been shown to be able to complete their productive life cycle. Examples include the propagation of HPV-16, HPV-18, and HPV-31 in organotypic cultures (Meyers _et al._, 1992 & 1997; Flores _et al._, 1999), and HPV-1, HPV-11, HPV-16 in athymic mice kidney capsule and SCID mice skin xenografts (Kreider _et al._, 1987 & 1990; Brandsma _et al._, 1995; Bonnez _et al._, 1998, reviewed in Bonnez, 1998). Although the viral life cycle could be mimicked in these model systems, very few had compared the viral events in these model systems with natural infections.

NUDE mice kidney capsule and SCID mice skin human skin implants have been stained for occurrence of viral late events. The timing of the late events in these model systems were similar to those seen in the real infections. In terms of the reconstruction of the viral life cycle, the SCID mouse skin xenograft model was the superior system between the two.

**NUDE Mice Kidney Capsule Implants**

CRPV, ROPV and HPV-11 infected mouse kidney capsule xenografts were obtained and stained with the respective antibodies. The distribution of the viral DNA and protein in these implant materials resembled that found in natural infections (Figure 5.7). The onset of viral DNA amplification and E4 expression correlated exactly, and L1 capsid proteins were detected in higher layers only in E4 positive cells. However, the typical histological characteristics of the natural infections, such as the thickening of the differentiated epithelial layers and papillae formation, could not develop fully and were not found in these implants.
**SCID Mice Skin Xenografts**

Similarly, HPV-11 skin xenografts showed the distribution of the viral late events as found in naturally-occurring lesions (Figure 5.7). Expression of E4 and viral DNA coincided in the same cells, and L1 expression was triggered afterwards in E4 positive cells. The transplanted human materials were able to support papilloma formation and provided a better morphological model for the virus infection.
Chapter 5 - Characterisation of the Events in the Late Stage of the Viral Life Cycle

Figure 5.7. The PV life cycle in murine xenograft model systems. The late stage of the viral life cycle were mimicked in murine xenograft implants either in the kidney capsule of NUDE mice or skin of SCID mice. The late events of ROPV, CRPV and HPV-11 infections in these model systems were similar to those of the natural infection. Viral DNA amplification and E4 expression coincided, and L1 was expressed after E4. Since the skin xenograft enables the outward growth of the papilloma instead of inward differentiation seen in the kidney capsule model, the SCID presents a better model system for the study of the papillomavirus life cycle. Basal layers are represented by dotted lines. Images were taken using a 10x magnification objective.
5.9 Discussion

Onset of the viral late events

The reliance on animal PV models and in vitro systems for the study of the PV life cycle has made these models important research tools. In order to be useful for the study of human infections however, it is essential to access how useful these experimental systems resemble the natural HPV infections. By examining infected tissue samples using immuno-staining methods, events during an infection can be determined both spatially and sequentially. Other approaches such as DNA and protein sequence analyses or the investigation of viral promoter activity would not have been able to provide information on the organisation of the virus life cycle described here. By mapping viral late events, each PV type was shown to exhibit a distinct pattern of life cycle organisation in its host tissue. No two PV types studied here exhibited the same viral DNA, E4 and L1 staining patterns. Generally, the onset of viral late events followed two different patterns. In COPV and BPV-1 infections, cells supporting the late stage of the viral life cycle were found in sporadic clusters throughout the lesions. In lesions caused by ROPV, HPV-11 and CRPV late events were supported only in specific areas of the papilloma. In some tissues, the late stages of the virus life cycle were initiated in the basal and parabasal layers (COPV, ROPV and BPV-1) while others showed initiation only in the more differentiated layers (CRPV and HPV-11). As a unique feature of its life cycle, COPV was shown to be able to initiate viral DNA amplification and E4 expression in the basal layer of the oral mucosa. Although the basal layer may include some post-mitotic differentiating cells (Li et al., 1998), it remains to be established how the COPV late events can be supported in such cells. It is possible that the lack of an E5 open reading frame in the COPV genome, and the additional non-coding region found between the E2 and L2 open reading frames could contribute to this unique pattern of events (Nicholls et al., 2001). By comparison, the COPV life cycle seems to share some similarities with cutaneous HPV types, such as HPV-1, HPV-2 and HPV-63, based on the pattern of late events in infected tissue sections. All of these virus types are able to initiate viral genome amplification.
and high levels of E4 expression in the lower layers of the epidermis and to support efficient L1 synthesis during their productive life cycles (Peh et al., in press). Interestingly, COPV was found to have a close evolutionary relationship with HPV-1, HPV-41 and HPV-63, based on protein alignments of the E4, E6 and L1 sequences (Chan et al., 1995; Chapter 3). The results of this chapter suggest that the initiation of the late stage of the virus life cycle may be differently regulated at the transcriptional and translational levels in infections caused by different PV types. Infections caused by CRPV, ROPV and HPV-11 showed similar patterns of late event activation.

Previous studies examined the late stage of the virus life cycle in HPV-1 and HPV-16 infected tissues have suggested that the virus life cycle is tightly regulated (Doorbar et al., 1997). The results of this chapter support this observation and have involved the analysis of five other PV types. In an established productive infection, viral DNA amplification and E4 expression are linked events. Since the E1, E2 and E1^E4 proteins are thought to be expressed from a single transcriptional unit from the viral differentiation dependent promoter, it is perhaps not surprisingly that E1^E4 expression and genome amplification coincide.

Expression of L1 was only detected in cells which already contain high levels of viral DNA and E4. In Chapter 8 of this thesis, I will describe the effects observed in the late stage of the viral life cycle when E4 mutant viral DNA was introduced into domestic and cottontail rabbits. Preliminary results indicate that E4 is required for the initiation of viral DNA amplification and L1 expression. When taken together, these results suggest that L1 expression is also intricately linked to the completion of earlier late events. The delayed expression of L1 results in a spatial gap between the onset of E4 expressions and that of the virus coat proteins. This gap was found to be consistent in infections caused by the same virus type, but differed in lesions caused by different PV types. Up to now, the exact mechanisms that control the regulation of L1 expression has not been solved. Besides promoter activity, the timely expression of L1 is likely
Chapter 5 - Characterisation of the Events in the Late Stage of the Viral Life Cycle

to involve other factors. These are likely to involve both viral and host cell proteins and regulatory pathways. Among the possible factors that have been proposed are changes in codon usage during host cell differentiation processes (Zhou et al., 1999), a threshold level for viral genome copy number in cells that are not supporting genome amplification (Grassman et al., 1996), changes in mRNA processing patterns and stability (Furth et al., 1994; Sokolowski et al., 1998) and perhaps the presence of the E4 protein. We do not know if the same regulatory factors apply in all PV infections, but this may well be the case, as the general L1 pattern is consistent in different virus infected tissues. The host and tissue type specificity of each PV, as well as the different mechanisms of virus infection, suggest that a multitude of different viral and host cell elements must be co-ordinated in order to create a suitable environment in which the productive life cycle of the virus can be supported.

Abortive and non-productive infections

From an evolutionary perspective, production of infectious virions is important for the success of a particular virus type. Viruses which are incapable of completing their productive life cycle efficiently are expected to be evolutionarily short-lived. Non-productive infections can however occur during carcinogenesis, and in 'secondary' hosts, where host cell conditions are altered or unfavourable for the virus. In this chapter, I have described two examples where late events do not occur in the expected manner (i.e. in some COPV and CRPV infected lesions). In these instances, the lesion contained cells that are supporting abortive or non-productive infections. Cells with high levels of viral DNA but no or very little detectable E4, or vice versa, were rare but could be found in COPV infected lesions. For this observation to be significant, we assume that the levels of staining are directly related to the levels of DNA or protein in the cells. Thus, cells with strong ISH signals are assumed to contain high levels of viral DNA. Cells with no E4 staining suggested the absence of protein expression. On the other hand, high levels of E4 did not necessarily ensure viral DNA amplification. This suggests that additional viral and/or host cell factors may be necessary. In the tissue sections, the vast majority of cells showed that the
onset and levels of viral DNA amplification and E4 expression correlated very closely with each other. In a small number of cells, it appears that E4 expression is not required for viral copy number to increase. Without appropriately high levels of viral DNA and E4 in the cells, the virus is assumed to be unable to trigger the synthesis of its capsid proteins.

In the case of CRPV infected NZW rabbit papillomas, the absent or small amount of viral DNA amplification detected was assumed to be due to the fact that CRPV has not evolved to complete its life cycle in this 'secondary' host. Although the presence of specific host inhibitory factors has not been fully investigated, differences in host genetics and the virus' inability to activate transcription from the viral late promoter are likely to be contributing factors (Shope, 1933; Zeltner et al., 1994; Han et al., 1992, Salmon et al., 2000). NZW rabbits are permissive for papilloma development, but do not support the full late stage of the viral life cycle following experimental inoculation with plasmid CRPV DNA and virus suspensions (Noya & Mellors, 1957). The frequency of wart regression and progression to carcinoma has been reported to be significantly different between cottontail and NZW rabbits infections (Brietburd et al., 1997, Rous & Beard, 1935). The rare detection of low level of any viral late events in NZW rabbit papillomas may be linked to the high rate of malignant transformation observed in these rabbit hosts (approximately 75%), and to the inability of the virus to activate its differentiation-dependent promoter (Phelps et al., 1985). The occasional presence of low levels of viral DNA amplification in some NZW rabbit papillomas suggests leaky transcription from the differentiation-dependent virus promoter, which would lead to translation of the E1 and E2 proteins. We assumed that viral DNA cannot be efficiently replicated in these cells, when compared to those in which genome amplification occurs in cottontail rabbit papillomas.

It is possible that the inability of CRPV to complete its life cycle in domestic rabbits might be related to the use of experimental methods in the induction of the virus infection. Most of the animal tissues where abnormal life events were observed were collected from experimental
Chapter 5 – Characterisation of the Events in the Late Stage of the Viral Life Cycle

infections. Differences in cellular responses, particularly the immune response, to natural or experimental virus infections might contribute to the differences in viral life cycle progression. The different methods used for producing experimental lesions include particle-mediated (gene gun) DNA inoculation, epidermal puncture and wounding followed by inoculation using virus or DNA preparations. These different approaches of DNA and virus delivery are employed in various animal systems for the study of vaccine therapy, infection and viral protein expression. However, it has been found that wound healing after particle inoculation could affect the outcome of the experiments directly or indirectly. The activation of specific immune responses was one of the many suggestions made to explain this result (Meuli et al., 2000). Abnormal wound healing may influence the ability of the virus to produce an abortive rather than a productive infection.

**Intracellular E4 expression patterns in tissue sections**

The intracellular E4 expression patterns exhibited by the three PV types (COPV, CRPV and ROPV) were found to be highly heterogeneous between the different viruses. Generally, E4 is detected predominantly in the cytoplasm. When expressed in cells in vitro, the E4 proteins of HPV-1, HPV-16 and HPV-31 have previously been shown to exhibit keratin network association (Doorbar et al., 1992; Rogel-Gaillard et al., 1992; Pray & Laimins, 1995). In addition, cytopathic effects such as the formation of inclusion granules, disruption of the cytokeratin network and vacuolisation could variously be found in HPV-1, HPV-16, HPV-31, and BPV-1 E4 expressing cells (Doorbar et al., 1991; Jareborg & Burnett, 1991; Rogel-Gaillard et al., 1993; Pray & Laimins, 1995). Unlike the HPV-1 inclusion granules which were also detected in infected tissue sections (Rogel-Gaillard et al., 1993), the association of E4 with cytotkeratins, either in filamentous networks or collapsed granules, had not been shown in immuno-stained tissue sections. Failure to detect the precise pattern of protein expression and distribution, particularly the cytokeratin network, is not uncommon in immunohistology. Tissue fixation, processing, embedding and sectioning can affect the immuno-detectability of proteins by
antigen masking (Fox et al., 1985; Mason & O'Leary, 1991). For E4 detection in the formalin-fixed paraffin-embedded tissues examined here, antigen exposure steps, such as heat denaturation, protease digestion, and the use of tyramide signal amplification were sometimes necessary in order to effectively detect and visualise the protein (Shi et al., 1991; Wu et al., 1992; Cattoretti et al., 1993; Zaidi et al., 2000). Heat denaturation was used to reveal the antigenic epitopes for BPV-1, and CRPV, and protease digestion for the nucleolar localisation of COPV E4. A combination of both heat and protease action was required for the detection of HPV-11 E4. The inclusion granules were most clearly shown in ROPV infected tissue sections with the use of tyramide substrates.

The significance of the different intracellular distribution patterns of E4 in tissues infected by different PV types is not entirely understood. Sequence alignments of the E4 proteins (Chapter 3) shows the presence of conserved motifs between some PVs. Despite the sequence similarities at the protein level, the E4 proteins of CRPV and ROPV did not share much in common in terms their cellular localisation. CRPV E4 was localised in almost equal levels in both the cytoplasm and nucleus, whereas ROPV E4 was a predominantly cytoplasmic protein with the ability to form cytoplasmic inclusion granules. Inclusion granule formation is a common feature of the E4 proteins of cutaneous type HPVs in infected tissues and in vitro. The ability to multimerise to form protein aggregates and inclusion granules was found to be mediated by the C-terminal domain of E4 (Ashmole et al., 1998). Similar C-terminal sequences are apparent in the two mucosal type viruses COPV and ROPV (Chapter 3), and both were seen to be able to form inclusions in infected tissues. It is not clear why some viral E4 proteins should form inclusion granules while others do not. This E4 cytopathic effect does not appear to be an exclusion characteristic of the cutaneous viruses however.

The presence of the COPV E4 in the nucleus and more specifically in the nucleolus of infected cells might suggest a role for the protein in nuclear-related processes, such as
transcription, replication and some cell cycle events. Since E4 proteins contain a putative NLS and NES motifs (Rogel-Gaillard et al., 1993; Roberts et al., 1994), the final distribution of the protein in cells may depend on the rate of protein import and export from the nucleus at equilibrium (Kaffman & O'Shea, 1999). This could explain the differences in nuclear and cytoplasmic levels of E4 in COPV and ROPV infected lesions. However, the presence of a putative nucleolar localisation signal in COPV E4 (Chapter 3) and the detection of COPV E4 in the nucleoli in tissue sections may suggest a more defined role for the protein. E4 is essential for viral genome amplification in the CRPV life cycle (Chapter 8). Perhaps the role of E4 involves its ability to translocate within the cellular compartments. A similar role was also described for the Rous Sarcoma src gene product. When expressed in cells, this viral protein exhibited selective translocation between the nucleus, nucleolus and cytoplasm, depending on the differentiation state of the cells, and accumulation of the protein in the cytoplasm coincided with a G2 cell cycle arrest (David-Pfeuty & Nouvian-Dooghe, 1995).

The best model system for the study of HPV infections

HPV infections can cause a wide range of diseases from benign skin tumours, to malignant carcinomas such as invasive cervical cancer. The clinical significance of mucosal type HPVs, particularly the high-risk type viruses, is well established. HPV associated cervical cancer is the second most common female cancer world-wide. Although research on the molecular biology of HPV infections had been carried out for the last few decades, it is not until recently that an in vitro cell culture system has been developed that allows the complete life cycle of some HPV types to be supported in the laboratory (Meyers et al., 1992 & 1994; Frattini et al., 1997). The re-creation of the HPV life cycle in differentiating organotypic cultures following DNA transfection has provided opportunities for the molecular manipulation of the viral genomes, and the study of the HPV life cycle events in vitro. Despite the ability to complete its life cycle in differentiating cultures, several aspects of virus infection cannot be examined in these in vitro systems. These include host immunity, the events that take place during papilloma formation,
and disease progression (malignancy, persistence and latency). Moreover, these cell culture techniques are not only laborious to perform and can be difficult to reproduce. Fortunately, other approaches have also been used to study the HPV life cycle. These include the use of animal PV infections and the use of murine transplantation models (Bonnez, 1998).

The most commonly used animal model systems in PV research are BPV-1 and CRPV. Housing and handling considerations have resulted in the CRPV-rabbit model becoming the most popular model for the study of PV infections in vivo. As a model system for the study of mucosal HPV infections however, CRPV does not appear at first sight to be a suitable choice. Most importantly CRPV infects cutaneous epidermis and causes productive infections in wild cottontail rabbits, but does not complete its life cycle in domestic rabbits (Shope, 1933; Kidd & Parsons, 1935). From the analysis of late events during infection by animal PVs, ROPV appeared to be the best candidate as an animal model for human mucosal PV infections. Not only can ROPV infect the oral cavity, it was also shown to be able to infect genital epithelial (Harvey et al., 1998). Moreover, the initiation and distribution of ROPV late events was similar to the pattern of events seen in HPV-11 infections. Both virus types initiated their late events in confined areas in the intermediate layers of the mucosal epithelial. Other advantages offered by the ROPV-rabbit model system include the readily available natural host (domestic rabbits), the availability of established protocols for their housing and handling, and the ability to study infection in hosts with a specific immuno-background.

Based on the pattern of viral late events, COPV did not qualify as a good model for HPV infection, even though this model has been the choice for immunological and vaccine studies (Stanley et al., 2001). The sporadic pattern of cells supporting late events within the papilloma lesion differed significantly from that seen in HPV-11 infections. Furthermore, the difficulties and expense involved in the housing of dogs for experimentation makes the use of this animal model undesirable. Last but not least, the unique large non-coding region within the COPV
genome, as well as the initiation of the late viral events in the basal layer are sufficient to set this canine virus in a separate category within the PV family (Delius et al., 1994; Nicholls et al., 2001).

Murine transplantation models are regarded as valuable alternatives to the use of CRPV and COPV, and have been used for a variety of immunological and vaccine studies. In PV research, mouse kidney capsule implants and skin grafts are successfully used to propagate viruses of various HPV types. Although these systems are widely used, very few studies had been done to compare the pattern of virus life cycle events in these murine xenografts with that of natural HPV infections. By studying the pattern of late events in HPV-11 infected foreskin implants propagated under the renal capsules (NUDE mice) or on the epithelial surface of SCID mice, the latter was undoubtedly a better model for mimicking the natural HPV-11 infections in the transplant systems. Immuno-detection of viral late markers in these tissues showed that the HPV-11 life cycle events were recreated in almost the same pattern in skin xenografts and in infected cervical lesions. Moreover, previous studies which looked at the tissue histology of the internal and surface xenografts of CRPV infected tissues also found that internal implantations resulted in abnormal papilloma tissue growth (Rous & Beard, 1960). The proliferating layer of the epithelial formed the outermost circular boundary, where dividing cells were pushed inwards towards a centre of thick cornified cells. On the other hand, papilloma growth from the skin grafts extended outwards, and was similar to that seen in natural virus-induced papillomas (Rous & Beard, 1960). Although both systems have been shown to support the full HPV life cycle, the distortion of the papilloma morphology reduced the usefulness of the kidney capsule xenograft as a model to study virus life cycle events. Skin grafting not only enabled the proper growth of the papilloma at the right tissue site, but would also provide a good system for topical therapeutic applications and studies of HPV infection.
# Chapter 6

CHAPTER 6 – EXPRESSION OF CELL CYCLE MARKERS IN INFECTED TISSUES IN RELATIONS TO THE ONSET OF THE VIRAL LATE EVENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Introduction</td>
<td>180</td>
</tr>
<tr>
<td>6.2 Staining protocol and antibodies</td>
<td>183</td>
</tr>
<tr>
<td>6.3 Expression of PCNA and the onset of the viral late events</td>
<td>183</td>
</tr>
<tr>
<td>6.4 Expression of G1/S transition cyclins and the onset of the viral late events</td>
<td></td>
</tr>
<tr>
<td>6.4.1 Cyclin E and the viral late events</td>
<td>189</td>
</tr>
<tr>
<td>6.4.2 Cyclin A and the viral late events</td>
<td>190</td>
</tr>
<tr>
<td>6.5 Expression of cyclin B (G2/M transition) and onset of the viral late events</td>
<td>192</td>
</tr>
<tr>
<td>6.6 Discussion</td>
<td>196</td>
</tr>
</tbody>
</table>
Chapter 6 - Expression of Cell Cycle Markers in Relations to the Viral Late Events

6.1 Introduction

The PV life cycle is intricately linked to epithelial cell differentiation. From the infection of the basal layer, followed by the initiation of early events in the lower layers, to the onset of late events in the differentiated cells. The complex network of regulatory mechanisms which coordinate the viral life cycle and epithelial differentiation is still not completely understood. Studies on the HPV early proteins E6, E7, E1 and E2 in vivo and in vitro model systems have shown that these viral proteins can associate with many host cell cycle regulators (Thomas et al., 1999; Zwerschke & Jansen-Durr, 2000). These viral-host interactions result in the interruption of normal epithelial differentiation and cell cycle regulation (Southern & Herrington, 1998 & 2000; Southern et al., 2000).

Shortly after virus infection, the expression of the early proteins E6 and E7, as well as very low levels of E1 and E2, are required for cell transformation and viral DNA episome replication respectively (Chow & Broker, 1994; Munger et al., 1989). Unfortunately, the direct analysis of these proteins in infected tissue sections had been impeded by the difficulties in the immuno-detection of the viral early proteins using specific antibodies. With the current knowledge of some of the viral-host protein interactions, an indirect method for detecting the presence of the early viral events in cells is to analyse the altered expression patterns of specific host cell surrogate markers. These included the proliferative markers PCNA (proliferating cell nuclear antigen) and ki67, and cyclins D, E, A and B1. The expression patterns of these cell-cycle regulators had been well-documented for HPV infections by low- and high-risk viruses, that cause low- and high-grade cervical lesions (Southern & Herrington, 1998; Southern et al, 2000). To determine the relationship between the early and late events of the viral life cycle in animal and human PV infected tissues, the expression patterns of surrogate markers for the viral E7 protein were examined, together with E4 expression or viral DNA amplification.
The expression of the viral transformation proteins, E6 and E7 disrupts the host cell cycle process through various different mechanisms and molecular pathways. The main associate of the E7 oncoprotein is the retinoblastoma (pRB) protein, which plays multiple roles in regulating cell cycle progression. E7 binds hypophosphorylated pRB and leads to the release of the host cell transcription factor, E2F. E2F in turn activates the transcription and expression of several S-phase genes, including PCNA, and cyclins E and A (Chellappan et al., 1992; Harbour & Dean, 2000). Up-regulation of PCNA (an auxiliary protein for polymerase delta), cyclin E (a G1-S-phase transition regulator) and cyclin A (a S-phase entry and progression cyclin), indicates a direct role of E7 in activating the replication machinery in the infected host cells. In addition to positively promoting cell proliferation, E7 was also shown to be able to disrupt cell cycle checkpoints at the G1/S transition phase by binding to the cyclin-dependent kinase inhibitors (CDKIs) p21 and p27, and abolishing their inhibitory activities in vitro (Funk et al., 1997; Jones et al., 1997; Zerfass et al., 1996). The role of E7 is crucial for the establishment of a suitable S-phase cellular environment that will allow viral DNA replication to occur in post-mitotic, differentiated keratinocytes.

Working together with E7 to induce cell cycle arrest is E6, another viral oncoprotein, which has been shown to be able to bind the tumour suppressor p53 and cause the rapid degradation of the protein complex (Werness et al., 1990; Hubbert et al., 1992). Loss of p53 function as a result of E6 binding, not only abrogates the G1/S, but also the G2/M transition checkpoint, which leads to the accumulation of tetrasomy and aneuploidy cells. This could be achieved by the failure to activate p21 and loss of the mitotic spindle checkpoint (Thompson et al., 1997). Recently, it has been shown that the expression of cyclin B, a G2/M transition regulator, was up-regulated in high-grade squamous intraepithelial lesions (SIL) (Southern et al., 2000). The expression of E4 in cell cultures and S. pombe yeast cells has also been shown to increase cyclin B accumulation in the cytoplasm and caused cell cycle arrest at the G2 phase (Davy, 2002).
Immuno-detection of the host cell cycle proteins together with viral late event markers could be useful for the analysis of viral life cycle events in different PV infections and tissue samples. So far, only HPV infected biopsies have been extensively studied using this approach. Similar observations regarding the altered expression patterns of cyclins were obtained here following the study of animal and human PV infections. Among these included elevated levels of PCNA, cyclins A and E in differentiated cells, and increased cyclin B expression in the infected epithelium. PCNA, and cyclins A and E, were used as surrogate markers for the early stage of the viral life cycle. The expression patterns of the early markers and the onset of the viral late events were collectively used to describe a general pattern in the viral life cycle as it occurred during host infections.
6.2 Staining protocol and antibodies

Formalin-fixed paraffin-embedded tissue sections were microwave-treated in citrate buffer as an epitope exposure method for the immuno-detection of the cell cycle protein (see Chapter 2). Coupled with this step was the use of the tyramide signal amplification (TSA) system (NEN Life Science Products Incorporated, USA) to increase the sensitivity of the immuno-detection protocol. For more fastidious antigens, such as the HPV-11 E4 proteins, double immuno-stains were achieved with a combination of antigen exposure steps as described in Chapter 5. Unfortunately, not all commercially available cell cycle protein antibodies were pre-tested for, or showed cross-reactivity with some of the animal species analysed here. Table 6.1 shows the summary of the antibodies used and the cross-reactivity in the different animal tissues.

<table>
<thead>
<tr>
<th>Antibodies</th>
<th>Clone, source and immunogen</th>
<th>Antigen recognition in PV infected animal tissues sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCNA</td>
<td>PC10 mouse monoclonal; Neomarkers; protein A-rat PCNA fusion protein</td>
<td>Bovine, canine, rabbit and human</td>
</tr>
<tr>
<td>Cyclin A</td>
<td>mouse monoclonal; Novocastra; human cyclin A</td>
<td>Human</td>
</tr>
<tr>
<td>Cyclin B</td>
<td>clone V152 mouse monoclonal; NeoMarkers; hamster cyclin B1</td>
<td>Bovine, canine, rabbit and human</td>
</tr>
<tr>
<td>Cyclin E</td>
<td>mouse monoclonal; Novocastra; bacterial expressed fusion full-length human cyclin E</td>
<td>Canine, rabbit and human</td>
</tr>
</tbody>
</table>

Table 6.1. Cell cycle protein antibodies and their species cross-reactivity on formalin-fixed paraffin-embedded tissue sections.

6.3 Expression of PCNA and the onset of the viral late events

PCNA was first identified as an auxiliary protein for polymerase delta in 1984 (Mathews et al., 1984), and has since been shown to be essential for cellular and viral DNA replication. PCNA is a highly conserved acidic nuclear protein in mammalian cells, and has an approximate
molecular size of 36kDa (Almendral et al., 1987). PCNA homologues have also be identified in plants and viruses (Suzuka et al., 1989; O’Reilly et al., 1989). Using the same antibody (PC10), PCNA was detected in all the PV infected tissues analysed here.

In general, PCNA was detected in the nuclei of cells from the basal to the differentiated layers in PV infected tissues. In normal epithelium, PCNA is expressed only in the basal and parabasal layers. Staining patterns showed that the number of cells expressing PCNA was increased by several fold in all papilloma tissues. In dual stained sections, an overlap region could be clearly seen where cells were positive for both PCNA and the viral late stage markers, E4 expression or viral DNA amplification. There was no exception to this observation in all the tissue types tested.

Mucosal type viruses

In COPV infected oral lesions, PCNA was detected in the nuclei of cells from the basal layer to the top of the lesions and persisted until just a few layers from the epithelial surface (top panel in Figure 6.2a). Almost every cell in the basal and immediate parabasal layers expressed high levels of PCNA. As these cells moved into the intermediate layers, PCNA staining became sporadic, and in the higher layers, protein levels gradually decreased. In the lower region of the epithelium, all the E4 positive cells were also PCNA positive, but not all PCNA positive cells were found to express E4. These double positive cells persisted until the intermediate layers, while only E4 positive, PCNA negative cells were detected in the higher region. Rarely, PCNA was present in differentiated keratinocytes without E4 expression. In these cells, PCNA may be incorporated into a quaternary protein complex, that is stable and inactive (Jian et al., 1999).

In the mucosal epithelium of the normal rabbit lower tongue, PCNA was detected mainly in the basal and parabasal layers (middle panel in Figure 6.2a). In ROPV infected papillomas, up-regulation of PCNA extended to the upper layers, before it became undetectable in the
uppermost layers of the mucosal epithelium. In PCNA/E4 double immuno-stained sections, a large overlap region containing double positive cells marked the cells that are actively supporting viral DNA amplification, since it was previously showed that the E4 expression pattern coincides exactly with the detection of viral DNA amplification (Chapter 5). A small number of PCNA-positive, E4-negative post-mitotic differentiated cells were also found in these lesions.

In HPV-11 infected cervical lesions, almost all cells from the basal to highly differentiated layers near the surface of the epithelium displayed increased PCNA levels. As previously reported, immuno-detection of the proliferation marker ki67 did not extend into the superficial cells in low-grade SIL (Southern & Herrington, 1998). However, the up-regulation of PCNA did not seem to be uniform throughout the lesion as discerned by the level of fluorescence (bottom panel in Figure 6.2a). PCNA/E4 double immuno-staining showed that the double positive overlap region was limited to the upper layers of the lesion, as the onset of the viral late events was usually triggered only in the upper intermediate layers of the epithelium.

*Cutaneous type viruses*

Despite the poor immuno-staining signals in the BPV-1 and BPV-2 infected tissue sections, PCNA appeared to be up-regulated throughout the lesions, as found in other PV infected tissues, and the presence of double positive cells could be just about distinguished.

In CRPV infected cottontail rabbit tissue sections, PCNA could be detected from the basal to the superficial layers of the epidermis (lower panel in Figure 6.2b). Interestingly, in all the papillomas analysed, the extent of PCNA staining in the differentiated layers was highly variable, even within the same warty tissue. Expression of PCNA could reach the uppermost layers of the epidermis, or become undetectable in the upper intermediate layers. Double immuno-stained sections showed a relatively large overlap region where cells containing PCNA, and which were also supporting viral DNA amplification were found. In some sections, PCNA appeared to have
both a nuclear and a cytoplasmic intracellular distribution. The cytoplasmic PCNA staining pattern was thought to be caused by poor tissue fixation in aqueous formaldehyde, as previous reports have noticed marked alterations in cellular membrane permeability following tissue-fixation in this fixative. Alternatively, tissue damage might have resulted from the harsh procedures carried out before and during wax embedding (Fox et al., 1985).
Chapter 6 - Expression of Cell Cycle markers in Relation to the Viral late Events

Figure 6.2. A population of PCNA positive cells in the differentiated layers of the epithelium coincides with the onset of viral late events.
(a) Mucosal type virus infections. (b) Cutaneous type virus infections. The level and extent of PCNA up-regulation varied in the different tissues. With the detection of E4 expression or viral DNA amplification as an indication of the onset of the viral late events, the double immuno-stained sections showed an overlap region where PCNA and viral late events were detected simultaneously in the same cells. The overlap regions were variable in thickness between different virus types but consistent for infections caused by the same virus type at the same site. The double positive population is thought to represent replication competent cells where viral DNA amplification is occurring. The ROPV, CRPV and BPV-2 images were taken with a 10x magnification objective, and the COPV and HPV-11 images with a 20x objective. Nuclei were counterstained with either DAPI (blue) or sytox green (green) in the CRPV lesion. m, (r), (g) and (b) represent merged colours, red, green and blue respectively. Double positive PCNA/E4 or PCNA/DNA cells in the overlap region are indicated with bold arrowheads, while thin arrows show PCNA-negative cells that have exited the replication competent region. The basal layers are represented by dotted lines.
Chapter 6 - Expression of Cell Cycle Markers in Relations to the Viral late Events

(a) Mucosal type viruses

ROPV/m (10x)

ROPV/PCNA(r)+E4(g)

ROPV/PCNA

COPV/m (20x)

COPV/PCNA(r)+E4(g)

COPV/PCNA

HPV-11/m (20x)

HPV-11/PCNA(r)+E4(g)

HPV-11/PCNA

(b) Cutaneous type viruses

BPV-2/m (10x)

BPV-2/PCNA(r)+E4(g)

BPV-2/PCNA

CRPV/m (10x)

CRPV/PCNA(r)+DNA(b)

CRPV/PCNA
6.4 Expression of G1/S transition cyclins and onset of the viral late events

E7-induced E2F transcription factor release and activation leads to the expression of E2F regulated genes. Among the effector gene products are the G1/S transition regulators, cyclins E and A.

6.4.1 Cyclin E and the viral late events

Cyclin E is not usually detectable in tissue sections because of its low level expression and rapid protein degradation. It has been proposed, that in the presence of E7, stabilised forms of cyclin E can be detected in cells (Jian et al., 1999). Cyclin E was detected in the parabasal and differentiated layers of PV infected tissues. Double staining of cyclin E and E4 or viral DNA showed that the presence of the stabilised cyclin E complex and onset of the viral late events were mutually exclusive in infected tissues. However, double positive cells were rarely detected, and these were thought to represent replication competent cycling cells that are supporting the late stage of the viral life cycle.

Mucosal type viruses

Up-regulated levels of cyclin E were detected in ROPV and HPV-11 infected sections. In both rabbit and human tissues, cyclin E was present only in a small number of differentiated cells from the parabasal to higher layers of the mucosal epithelia. In HPV-11 infected cervical lesions, the sporadic distribution of cyclin E was clearly in a different subset of cells from the E4 positive population, in accordance with the hypothesis presented by Jian et al. (1999) (upper panel in Figure 6.3). However, this mutually exclusive pattern was not as apparent in ROPV infected tongue lesions. As shown in the middle panel of Figure 6.3, the difficulties in discerning cell-cell boundaries in the rabbit tongue lesions, prevented proper comparison of the two fluorescence signals. Numerous cyclin E and E4 double positive cells were apparent in these lesions.
**Cutaneous type viruses**

In CRPV infected rabbit lesions, abundant cyclin E staining was seen from the parabasal to the intermediate layers of the epidermis. Interestingly, the stable cyclin E complex appeared to be degraded just before the onset of viral DNA amplification, as cyclin E was undetectable in the upper layers of the epidermis (lower panel in Figure 6.3). The difference between the cyclin E and viral late event staining patterns may be a unique characteristic of CRPV infections. It may also suggest that a different sequence of cellular events can occur in these rabbit lesions following viral-host protein interactions that leads to the rapid degradation of the cyclin E complex prior to the activation of the viral late events activation.

Cyclin E was not detected in the COPV and BPV-1 infected lesions (data not shown). Two possible explanations were given for this, namely the non-species cross-reactivity of the cyclin E antibody, or the absence of the suggested cyclin E protein complex.

**6.4.2 Cyclin A and the viral late events**

As an S-phase marker in normal epithelium, cyclin A would be predicted to be found only in proliferating basal and parabasal cells, but not in differentiated cells. The presence of cyclin A in the differentiated layers reflects the up-regulation of the protein as a result of E7 expression in the cell. In HPV-11 infected cervical lesions, cyclin A was detected in sporadic cells from the basal layer to differentiated layers near the top of the lesion. Due to the poor quality of staining carried out with this antibody, a correlation of cyclin A with the viral markers was not done. Unfortunately, the pattern of cyclin A expression could not be analysed in the animal tissues examined here. This was presumed to a lack of cross-species reactivity of the cyclin A antibody used.
Figure 6.3. Detection of cyclin E expression did not generally correlate with the onset of the viral late events, as represented by E4 expression or viral DNA amplification. Cyclin E was detected from the intermediate layers of the epithelium in the HPV-11 infected tissue, as previously observed. (Jian et al, 1999) Whereas, cyclin E was detected in the parabasal and intermediate layers in ROPV and CRPV infected tissues. Detectable cyclin E was mutually exclusive from the cells which were supporting the viral late events in HPV-11 and CRPV infections, although this was not always the case in the ROPV lesions. Some cyclin E and E4 double positive cells were found in the ROPV infected tongue lesions. Interestingly, the cyclin E protein complexes appeared to be degraded just before the onset of DNA amplification in the CRPV infected tissues. This was in contrast to ROPV and HPV-11 infections where cyclin E complexes were present in the same cell layers where the viral late events were triggered. A 20x magnification objective was used to take all the images in this figure. Nuclei were counterstained with DAPI (blue) or sytox green (green) in the CRPV lesion. m represents merged colours. Bold arrowheads indicate cyclin E positive, viral late events negative cells, and thin arrows indicate cyclin E and E4 positive cells. The basal layers are represented by dotted lines. All images were taken using a 20x magnification objective.
6.5 Expression of cyclin B (G2/M transition) and onset of the viral late events

Cyclin B is crucial for the G2-M phase transition and can be detected at high levels in the cytoplasm of cells that are arrested in the G2-phase. In a differentiating epithelium, cyclin B could be detected in the cytoplasm of sporadic cells in the basal and differentiating intermediate layers. In PV infected tissues, cyclin B expression was up-regulated. Double immuno-stained COPV and ROPV infected lesions showed exact co-localisation of cyclin B and E4. In CRPV papillomas, cyclin B1 was detected in the same regions as E4 expressing cells in adjacent sections. Low staining qualities in HPV-11, BPV-1 and BPV-2 infected tissues did not allow the comparison of cyclin B1 and E4 to be convincingly determined.

Mucosal type viruses

Cyclin B1 and E4 expressions correlated closely in COPV and ROPV infected oral and tongue lesions (first and second panels respectively on figure 6.4). Both proteins showed an exact immuno-staining pattern, in which the E4 positive cells had high levels of cyclin B1, but not vice versa. Cells with cyclin B1 but no E4 expression were thought to represent actively cycling cells that had not yet triggered the late stage of the viral life cycle. Since rat anti-ROPV E4 antisera and mouse anti-cyclin B1 antibodies were used for the E4/cyclin B1 double staining on the rabbit tongue lesions, concerns were raised regarding possible cross-reactivity of the anti-rat and anti-mouse secondary antibodies which could result in false positive staining patterns. These concerns were abolished as adjacent tissues tested with either cyclin B1 or E4 antibodies produced the same staining patterns in the different sections.

Unfortunately, the high background of the cyclin B1 immuno-stain on the HPV-11 infected tissue sections did not allow the correlation of the cyclin and E4 expression patterns to be established in these lesions (sixth panel on Figure 6.4). Despite the low levels of positive immuno-staining, it was however obvious that the cyclin B1 and E4 immuno-staining patterns did
not resemble that seen in the COPV and ROPV infected lesions. Double positive cells were rarely observed, and most cyclin B1 positive cells did not have high levels of detectable E4 (and vice versa).

**Cutaneous type viruses**

Cyclin B1 immuno-stains were difficult to obtain in BPV-1 and BPV-2 infected tissues. When fluorescence signals were visible, they were weak and were accompanied by a high general background level (third panel of Figure 6.4). Careful digital enhancement was required to bring out the positive fluorescence signals. In some cases, digital enhancement of double cyclin B1 and E4 immuno-stains showed a correlated staining pattern. However, no definite conclusion was drawn from these results.

In CRPV infected cottontail rabbit papillomas, cyclin B1 could be detected in the cytoplasm of sporadic basal and differentiating cells. Cyclin B1 and E4 double immuno-stains were difficult to perform due to the fastidious protocol required for the detection of CRPV E4 in the tissue sections. Correlation of cyclin B1 and E4 expression was shown on neighbouring tissue sections from the same rabbit tissue sample (fourth and fifth panels in Figure 6.4). In most cases, cyclin B1 was detected in the cytoplasm in a diffuse manner (CRPV/cycB_mag; fourth panel in Figure 6.4). Interestingly, a different cyclin B1 staining pattern was observed in one of the CT rabbit papilloma samples. Cyclin B1 appeared to be diffusely cytoplasmic in the basal and parabasal layers, but exhibited a granular staining pattern in the upper layers of the epidermis (CRPV/cycBgran_mag; fifth panel in Figure 6.4). The cyclin B1 granules differed in sizes and shapes in the cells, and were found in an area previously shown to supporting high levels of E4 expression. The different staining patterns were speculated to be an effect of different formalin-fixation conditions. The significance of these cytoplasmic structures needs further investigation.
Chapter 6 - Expression of Cell Cycle Markers in Relations to the Viral late Events

Figure 6.4. Cyclin B expression patterns showed variable degrees of correlation with the onset of the viral late events (E4 expression or viral DNA amplification). The best correlation was found in COPV and ROPV infected oral lesions, where cytoplasmic cyclin B was detected in almost every E4 positive cell. BPV-2 cyclin B and E4 correlation was difficult to discern from the high immuno-signal background levels, although a close correlation pattern eventually emerged when the individual immuno-stain images were digitally enhanced. A link between the expression patterns of CRPV E4 and cyclin B were shown in the same region of neighbouring sections of the same rabbit papilloma tissue. (CRPV/cycB+DAPI and CRPV/E4+DAPI) A 4x magnified image showed that cyclin B is predominantly cytoplasmic in the rabbit tissues. (CRPV/cycB_mag) In a separate experiment, similar cyclin B and E4 localisation patterns were again detected in neighbouring sections from a different CT rabbit papilloma tissue (figures CRPV/cycBgra+DAPI and CRPV/E4gra+DAPI). In this cyclin B immuno-stained rabbit tissue section, the intracellular localisation pattern of cyclin B exhibited an interesting cytoplasmic granular appearance, which resembled the inclusion granules commonly found in these cutaneous rabbit papillomas. (CRPV/cycBgra_mag) Poor cyclin B and E4 correlation patterns were found in the HPV-11 infected cervical lesion. Only a few rare cells were double positive for cytoplasmic cyclin B and E4. Bold arrowheads indicate cells with cyclin B and late viral events correlation. Thin arrows indicate cells that are only cyclin B positive. The basal layers are represented by dotted lines. Nuclei were counterstained with either DAPI (blue) or propidium iodide (red). m and (g) represent merged colours and green respectively. The COPV, ROPV, BPV-2 and HPV-11 images were taken using a 20x magnification objective. The low-magnification CRPV images were taken using a 10x objective, and the magnified images (mag) using a 40x objective.
6.6 Discussion

Expression of the viral early proteins E6, E7, E1 and E2 shortly after virus infection leads to transformation of the infected epithelial cells, and the creation of a pro-replication environment that allows maintenance of the viral genome copy number at a low level. In the previous chapter, the late stages of the viral life cycle were found to follow a common pattern. In order to link these events to the early stages of the viral life cycle, the expression patterns of surrogate markers of E7, such as PCNA and cyclins E and A, were used to show that a region of overlap exists between the early and late stages of the viral life cycle, and this is preserved in all the PV infections studied.

Altered expression patterns of cyclin have previously been reported in HPV infected lesions (Southern & Herrington, 1998 & 2000; Southern et al., 2000). It has also been shown that the expression of the viral E7 protein in differentiating cell culture systems could induce altered expression patterns of cyclin E, p21 and p27, and the prolongation of the replicative state in differentiated cells (Jian et al., 1999; Noya et al., 2001). Similar observations were made in the animal PV infected tissue sections analysed here. Proteins such as PCNA, and cyclins E and A, which are markers for the G1/S phases of the cell cycle, were found to be up-regulated in all the PV infected tissues examined. The induction of these proliferation markers in infected cells was the downstream effect of the expression of the viral E6 and E7 proteins. Increased PCNA expression was detected in all the infected tissues, and extended from the basal layer into the differentiated epithelial cell layers. While different virus types varied in their ability to up-regulate PCNA expression throughout the papilloma tissue, a region where the expression of PCNA coincided with the onset of the viral late events was found to be conserved in all tissues. This overlap region is assumed to represent a belt of replication competent cells that are supporting the vegetative replication of the episomal viral DNA genome. According to this hypothesis, cells which have migrated out of this replication competent zone will no longer be
able to support viral DNA amplification. Such non-replicative cells will have high levels of viral DNA but no PCNA. This appears to be the case in the tissues examined here. From the double stains of the viral late markers and PCNA, different virus types showed different degrees of overlap between the two stages of the viral life cycle. From our results, ROPV showed the most extensive overlap, while HPV-11 the least. This may reflect the ability of the different viruses to support a productive life cycle in their host environment.

During the normal cell cycle, the up-regulation of cyclin E is generally thought to be undetectable by immuno-detection. However, it has been shown that high levels of cyclin E can be detected in the differentiated layers of PV infected tissues, or E7 expressing raft cultures. The population of cells expressing detectable levels of cyclin E is thought to contain the protein as part of an inactive complex that is resistant to proteasome degradation in cells that are expressing E7 (Noya et al., 2001). Therefore, instead of an indicator of S-phase entry, cells with detectable cyclin E are thought to represent a population of post-mitotic cells that are incapable of DNA replication. Interestingly, the cyclin E staining results shown here support this hypothesis. In general, cells with detectable cyclin E belonged to the population that were not supporting the late stages of the viral life cycle. In other words, the detection of cyclin E may mark the lack of viral late activity, since the onset of viral amplification and cellular proliferation are generally exclusive from this cell population. However, exceptions were found where high levels of cyclin E were detected in some ROPV infected lesions, in cells which also expressed E4 (thin arrows; Figure 6.3). Unfortunately, the formation of the cyclin E complexes was not further investigated in the animal tissues, and the double immuno-stains of cyclin E and p21 or p27 were not done.

The suggestion that a stable quaternary complex exists that is resistant to protease or proteosome degradation does not seem to explain the results obtained from the analysis of HPV infected low-grade SILs, where cyclin E was only detected in the intermediate layers. In
these tissue sections, the cyclin E stable complex appears to be degraded in the upper layers, as shown by the lack of staining in the superficial layers (Southern and Herrington, 1998). This was in contrast to the observations made in HPV-11 E7 expressing rafts, which showed the presence of cyclin E stable complexes throughout the differentiated epidermis. Cyclin E was detected in the intermediate layers of HPV-11 and ROPV infected lesions, and in the parabasal layers of CRPV induced papillomas, but appeared to be degraded in the higher layers. From the staining patterns of cyclin E in the animal tissue samples, it was evident that the cyclin E distribution differed in the timing of its initial appearance as well as in the timing of its degradation. CRPV infected papillomas presented the most unique cyclin E staining pattern as the cyclin appeared to be rapidly degraded prior to the onset of viral DNA amplification. The significance of the cyclin E staining patterns could not be addressed in this thesis.

Recently, work in our lab had led to the observation that the expression of E4 in monolayer tissue culture cells resulted in cell cycle arrest in the G2 phase (Davy, 2002). Cyclin B is the main G2/M phase transition regulator, and can be detected in the cell cytoplasm during the approach towards mitosis. This prompted us to investigate a possible role for E4 in causing a G2 cell cycle arrest in cells which are amplifying the viral genome. From the PV infected tissues, increased expression of cyclin B was detected by immuno-staining. Interestingly, cyclin B was found to co-localise exactly with E4 in COPV and ROPV infected lesions, supporting the suggested role for E4 in inducing G2 arrest in these cells. Co-expression of the two proteins was also observed in CRPV and HPV-11 infected tissues, although the extent of the colocalisation was far lower than that seen in COPV and ROPV. In HPV-11 infected cervical lesions, only a few cells were found to be double positive for cyclin B and E4, and since double immuno-staining was not possible in CRPV infected papillomas, similar cyclin B and E4 staining patterns were shown in adjacent sections. E4 is thought to be associated with cyclin B, either directly or indirectly, and this association may be important during the onset of viral genome amplification. In fact, it has been shown than most E4 protein sequences contain a cyclin binding motif (RXL),
which may act as a potential site for direct cyclin B association (Davy, 2002). E4 may prevent cells from entering mitosis by sequestration of cyclin B in the cytoplasm. From the staining patterns of cell cycle markers and markers of viral late events, the PV life cycle appears to involve the up-regulation of S phase proteins, including PCNA and cyclin A, by the expression of the early viral proteins. This provides a replication competent environment in which the host replication machinery is available for viral DNA replication. Virus infected cells may thus be maintained in a pseudo-S phase state when the viral late events are activated and the amplification of viral DNA takes place in the differentiated cells. At the same time, the expression of E4 in these cells results in a G2 cell cycle state arrest. This replication competent zone is restricted to several cell layers until the S phase cyclins are no longer expressed and the virus continues with the late stage of its life cycle.
CHAPTER 7 - CONSTRUCTION OF HPV-16 AND CRPV E4 C-TERMINAL TRUNCATION PROTEIN MUTANTS

7.1 Introduction ............................................................................................................. 201

7.2 Design of the HPV-16 and CRPV E4 mutants ................................................... 203
  7.2.1 HPV-16 E4 mutants ...................................................................................... 203
  7.2.2 CRPV E4 mutants ...................................................................................... 204

7.3 Construction of recombinant plasmids containing the E4 ORF ....................... 206
  7.3.1 Vector and virus genome plasmids .............................................................. 206
  7.3.2 Cloning of the E4 mutagenesis template into pET-23b(+) ......................... 206

7.4 Kunkel site-directed mutagenesis ........................................................................ 207
  7.4.1 Verification of the E4 mutant -pET clones .................................................. 207

7.5 QuikChange site-directed mutagenesis ............................................................... 208
  7.5.1 Verification of the E4 mutant -pET clones .................................................. 208

7.6 Reconstruction of the HPV-16 E4 mutant viral genome ..................................... 209
  7.6.1 Construction of pTZW12 - a HPV-16 parental plasmid ................................ 210
  7.6.2 Cloning of mutant E4 from W12st-pET plasmids into pTZhW12 ............... 210
  7.6.3 Verification of the pTZhW12st clones ....................................................... 211
  7.6.4 Sequencing pTZhW12st E4 mutants ............................................................ 211
  7.6.5 Re-circularisation of the wt and E4 mutant viral genome ......................... 212

7.7 Reconstruction of the CRPV E4 mutant viral genome ....................................... 215
  7.7.1 Cloning of mutant E4 fragments from CrsE4-pET plasmids into pLACRPV ................................................................. 215
  7.7.2 Verification of pLACRst E4 mutants by restriction enzyme digestion ........ 215
  7.7.3 Sequencing pLACRst E4 mutants .............................................................. 216
  7.7.4 Final plasmid verification and re-circularisation of CRPV viral genome .... 217

7.8 Discussion .............................................................................................................. 220
7.1 Introduction

A straightforward way to investigate protein function(s) is to analyse the effect(s) of mutants. In PV research, mutated forms of several viral proteins, such as E4, E5, E6, E7 and L2, have been studied. Analysis of these mutants in in vitro assays as well as in vivo systems, has elucidated the role of each viral protein and has provided information as to the molecular mechanisms involved in PV pathogenesis.

Mutant forms of the BPV-1, HPV-1 and HPV-16 E4 proteins have previously been investigated in undifferentiated cell culture systems (Neary et al., 1987; Roberts et al., 1994 & 1999). However, the role of E4 in differentiated epithelia where the protein is normally expressed in abundance has not been studied. In order to address the function of E4 during natural virus infection, a series of E4 mutants for HPV-16 and CRPV were constructed. This chapter describes the mutagenesis process and generation of the E4 mutant DNA, which was used to examine the virus life cycle following experimental inoculation into primary human keratinocytes (for HPV-16) and rabbit hosts (for CRPV; described in Chapter 8).

There are problems in the construction of E4 mutants in the context of the whole viral genome. Firstly, the E4 gene is located within the E2 gene in the viral genome which makes genetic manipulation of the E4 ORF very restricted. Secondly, the E4 protein is translated from a spliced E1^E4 viral late transcript in differentiated cells (described previously in Chapter 1). Activation of the viral differentiation-dependent promoter also directs the expression of E1, E2, L1, and L2 transcripts during the late stage of the virus life cycle (Klumpp & Laimins, 1999; Stubenrauch & Laimins, 1999). Thus, severe genetic disruptions of the E4 ORF may be detrimental to viral transcriptional activity during the late stage of the viral life cycle. To overcome these potential problems, point mutations were made within the E4 gene, so that a series of stop codons were introduced into the E4 ORF of HPV-16 and CRPV. Altogether, five different
E4 mutant virus genomes were constructed each for HPV-16 and CRPV. These mutant genomes should have the ability to express E4 mutants with variable C-terminal truncations.

Two different mutagenesis methods were used here, namely Kunkel site-directed mutagenesis and QuikChange site-directed mutagenesis (Stratagene Europe, UK). The reconstruction of the E4 mutant virus genomes for HPV-16 and CRPV is described.
7.2 Design of the HPV-16 and CRPV E4 mutants

The HPV-16 and CRPV (Washington b strain) viral genome sequences were downloaded from the HPV Sequence Database WWW Home Page (http://hpv_web.lanl.gov/), and the E4 sequences carefully studied for potential mutagenesis sites. E4 mutants were constructed by introducing in frame stop codons (TAG, TAA or TGA) along the length of the E4 ORF. This was possible because the E4 and E2 genes are located in different reading frames of the viral coding strand. The base changes in E4 were made at strategic positions that would result in the pre-mature termination of the E4 gene product, but created only silent mutational changes in E2. It is vital that the E2 amino acid sequence remains unaffected by the mutagenesis process.

7.2.1 HPV-16 E4 mutants

Eight different sites were found to be suitable for the introduction of the specific base changes in the HPV-16 E4 ORF (see Appendix 1/II). These were well-spaced within the E4 ORF and would allow the expression of five different C-terminal truncated E4 proteins. The first E4 mutant, labelled 16st15/6, was predicted to have the potential to express the first fourteen amino acids of the E1-E4 protein (Figure 7.1a). Therefore, we have regarded this mutant to be in effect a HPV-16 E4 knockout. The remaining E4 mutants were labelled as 16st20, 16st32/4, 16st67/3, and 16st84, and were predicted to have the ability to express different C-terminal truncated E4 proteins, ranging in length from nineteen, thirty-one, sixty-six and eighty-three amino acids respectively (Figure 7.1a). Each E4 mutant is labelled according to the position of the stop codon along the translated E1-E4 sequence. Interestingly, the position of each mutation allowed the examination of the role of characteristic motifs in the HPV-16 E4 protein (Figure 7.1a).
7.2.2 CRPV E4 mutants

Six stop codons were introduced within the CRPV E1\(^{\text{E4}}\) ORF (Appendix 1/II). These would allow the expression of five different lengths of C-terminal truncated CRPV E1\(^{\text{E4}}\) proteins (Figure 7.1b). The CRPV E4 mutants were named \text{CRst90, CRst41, CRst74, CRst82, and CRst89}.\)
Chapter 7 – Construction of HPV-16 and CRPV E4 Mutants

(a) HPV-16 E4 mutants

<table>
<thead>
<tr>
<th>Mutants</th>
<th>Amino acid length</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>16st15/6</td>
<td>14</td>
<td>Disrupts LLXLL motif (leucine cluster) which is important for keratin association.</td>
</tr>
<tr>
<td>16st20</td>
<td>19</td>
<td>Retains LLXLL but loses proline-rich region and other motifs.</td>
</tr>
<tr>
<td>16st32/4</td>
<td>31</td>
<td>Retains leucine cluster and part of proline-rich region.</td>
</tr>
<tr>
<td>16st67/3</td>
<td>66</td>
<td>Lost 1/3 of C-terminal sequences.</td>
</tr>
<tr>
<td>16st84</td>
<td>83</td>
<td>Lost last 9 amino acids of the C-terminus.</td>
</tr>
</tbody>
</table>

(b) CRPV E4 mutants

<table>
<thead>
<tr>
<th>Mutants</th>
<th>Amino acid length</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRst9/0</td>
<td>8</td>
<td>Contains only the first 5 amino acids of the E4 protein.</td>
</tr>
<tr>
<td>CRst41</td>
<td>40</td>
<td>Retains the N-terminal sequence up till a possible proline-rich region.</td>
</tr>
<tr>
<td>CRst74</td>
<td>73</td>
<td>C-terminal deletion.</td>
</tr>
<tr>
<td>CRst82</td>
<td>81</td>
<td>Loses last 22 amino acids of C-terminus.</td>
</tr>
<tr>
<td>CRst89</td>
<td>88</td>
<td>Loses last 15 amino acids of the C-terminus.</td>
</tr>
</tbody>
</table>

Fig 7.1. Descriptions of the (a) HPV-16 and (b) CRPV E4 mutants. (Number) - indicates the number of amino acids in the different translated regions of the E1^E4 protein. [LLXLL] - denotes the position of the leucine cluster. [Pro-rich] - denotes the proline-rich region.
Chapter 7 – Construction of HPV-16 and CRPV E4 Mutants

7.3 Construction of recombinant plasmids containing the E4 ORF

Before mutagenesis, the target E4 template sequence was cloned into a recombinant plasmid containing a phage origin of replication. This was necessary as an uracil-enriched single-stranded DNA template was required during the Kunkel mutagenesis method.

7.3.1 Vector and virus genome plasmids

pET-23b(+) (CN Biosciences Ltd, UK; Appendix 2) was chosen as it contains a multiple restriction enzyme region suitable for the cloning of HPV-16 and CRPV E4 fragments, a phage origin of replication and an ampicillin resistance gene.

The W12 strain of HPV-16 was used as the parental template for the E4 mutagenesis. This HPV-16 genome was originally extracted from a CIN-1 derived cell line (HCW12) which has been found to be able to propagate the virus episomally in keratinocyte growth medium in vitro even after several passages (Stanley et al., 1989). The genome was cloned into pSp64 vector (Appendix 2) at the BamHI site at position 6150 (plasmid pSpW12, a kind gift from Dr Margaret Stanley, Department of Pathology, Cambridge).

The CRPV DNA was cloned into pLA II (a derivative of pBR322 without the tox gene; Appendix 2) at the SalI site (position 4571), and propagated as the plasmid pLACRPV (a kind gift from Dr Janet Brandsma, Yale University). The Washington b CRPV genome strain was used in our experiments. This CRPV strain was originally extracted from a productive virus infected CT rabbit wart, and first described by Nasseri et al. (1989).

7.3.2 Cloning of the E4 mutagenesis template into pET-23b(+)

The HPV-16 and CRPV genome sequences were analysed using MapDraw (DNASTAR Incorporated, USA), a DNA sequence analysis program. A list of specific restriction enzymes that
recognise a single cut site in the virus genome were obtained for HPV-16 and CRPV. Table 7.2 shows two unique sites for HPV-16 and CRPV that were found to be suitable for the cloning of E4 into pET23b(+).

<table>
<thead>
<tr>
<th>Virus</th>
<th>E4 ORF</th>
<th>Unique restriction enzyme sites (positions in the viral genome; viral gene)</th>
<th>Fragment length (in bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV-16</td>
<td>3332-3616</td>
<td>Hind II (3210; E2) and BamH I (6150; L1)</td>
<td>2940</td>
</tr>
<tr>
<td>CRPV</td>
<td>3668-4015</td>
<td>EcoR I (1063; E7) and Sal I (4571; E5/L2)</td>
<td>3508</td>
</tr>
</tbody>
</table>

Table 7.2. Restriction enzyme sites chosen for the cloning of the HPV-16 and CRPV E4 fragments into the pET-23b(+) vector.

Recombinant plasmids were labelled as W12E4-pET (HPV-16) and CRE4-pET (CRPV). These were transformed into DH5α host bacteria, amplified and purified using a caesium chloride gradient before being used for site-directed mutagenesis.

7.4 Kunke site-directed mutagenesis

In this method, one sense primer containing the specific base change(s) was synthesised for each mutant (refer to Chapter 2 for mutant primers sequences). Uracil-enriched single-stranded (U-ssDNA) templates of W12E4-pET and CRE4-pET were synthesised (data not shown; refer to Chapter 2 for mutagenesis protocol).

7.4.1 Verification of the E4 mutant -pET clones

Following mutagenesis, the DNA was purified and transformed into E. Coli. Double stranded E4 mutant -pET plasmid DNA was purified and checked by appropriate restriction enzyme digestion. The HPV-16 clones were digested using BamH I and Hind II, and CRPV clones with Sal I and Bsm I (position 2390) (data not shown). Clones containing DNA fragments of the correct size were selected and sequenced in both directions (see Chapter 2 for
sequencing primers and protocol). The data from sequencing were checked for the presence of E4 mutants. Unfortunately, no E4 mutants were obtained using this method despite repeated efforts to optimise each step of the protocol. Table 7.3 shows the statistics for this mutagenesis method.

7.5 QuikChange site-directed mutagenesis

Following the failure to obtain the E4 mutants using the Kunkel mutagenesis method, the QuikChange site-directed mutagenesis protocol was used. Two complementary primers containing the E4 mutant sequences were required for this mutagenesis procedure (see Chapter 2 for complementary primers sequences). Double stranded W12E4-pET and CRE4-pET DNA were used as templates for the mutagenesis reactions. This ensures that subsequent enzyme digestion and sequencing checks of the mutant clones were consistent with the Kunkel mutagenesis method.

7.5.1 Verification of the E4 mutant -pET clones

Mutagenesis reactions were carried out in vitro according to the manufacturer's instructions. The double-stranded mutant plasmids were transformed into E. Coli and purified. The HPV-16 and CRPV E4 clones were digested with restriction enzymes as previously described (data not shown). Clones showing the expected digestion patterns were selected and sequenced in both directions across the mutated E4 regions. All the desired E4 mutations in HPV-16 and CRPV were obtained using this method (Table 7.3). The failure to recover any W12E4st20 mutants using the first two primers was probably related to their inability to anneal very well.
### Table 7.3. Mutagenesis statistics. Comparing the efficiency of the Kunkel and QuikChange site-directed mutagenesis methods.

The mutagenesis reaction for generating the W12E4st20 mutation using the QuikChange method was carried out using three different complementary primers: 16st20 P2, 16st20m P2, and 16st20A P2 (described in Chapter 2). * represents the third and final attempt of mutagenesis using the 16st20A P2 primer.

<table>
<thead>
<tr>
<th>Mutant clones</th>
<th><strong>Kunkel mutagenesis</strong></th>
<th><strong>QuikChange mutagenesis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Total no. of attempts = 3)</td>
<td>(Average no. of attempts = 1)</td>
</tr>
<tr>
<td></td>
<td>no. of clones</td>
<td>no. of +ve</td>
</tr>
<tr>
<td>W12E4st15/6-pET</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>W12E4st20-pET</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>W12E4st32/4-pET</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>W12E4st67/3-pET</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>W12E4st84-pET</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>CRE4st9/0-pET</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>CRE4st41-pET</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>CRE4st74-pET</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>CRE4st82-pET</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>CRE4st89-pET</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

7.6 Reconstruction of the HPV-16 E4 mutant viral genome

Mutated E4 ORFs were cloned into the parental pTZW12 viral genome plasmid. Each mutant clone was verified by extensive sequencing and restriction enzyme digestion (Figures 7.5a & b). Mutant and wildtype (wt) viral genomes were linearised by BamH I digestion, gel-purified and ligated to obtain re-circularised viral DNA (Figure 7.5c).
Chapter 7 – Construction of HPV-16 and CRPV E4 Mutants

7.6.1 Construction of pTZhW12 - a HPV-16 parental plasmid

Restriction enzyme sites Hind II (Table 7.2) and Apa I (position 4533; L2 gene) were used to clone the mutated E4 fragments from the W12E4-pET recombinant plasmids into the W12 parental genome. Since the original pSp64 vector plasmid contains multiple Hind II enzyme recognition sites, it was necessary to re-clone the W12 genome into a new vector plasmid. pTZh19U is a modified version of the vector plasmid pTZIQU (Appendix 2), created by treating Hind II linearised pTZ19U DNA with the modifying enzyme Klenow at 37°C for 1hr, followed by ligation (data not shown). The modified pTZh19U plasmid retained its ability to be digested by BamH I but failed to be linearised by Hind II incubation. The W12 genome was cloned into the pTZh19U vector at the BamH I site. Recombinant W12 genome plasmids (named pTZhW12) were digested with BamH I to check for the presence of linearised vector and viral genome DNA. A selected clone was sequenced across the BamH I clone site to ensure that the viral DNA sequence remains intact (digest and sequence data not shown).

7.6.2 Cloning of mutant E4 from W12st-pET plasmids into pTZhW12

Apa I and Hind II restriction enzyme digests of the W12st-pET E4 mutant plasmids and pTZhW12 were carried out in the same reaction buffer. DNA was first digested with Apa I at room temperature prior to Hind II digestion at 37°C (described in Chapter 2). These enzymes do not cut the vectors pET-23b(+) and pTZh19U, but recognise a single cut site in the W12 HPV-16 DNA. Each reaction was run in a separate lane of an agarose gel. A small DNA band of approximately 1.4kbp, containing the mutated E4 ORF, was gel-purified from each of the W12st-pET reactions (data not shown). The larger (approximately 9kbp) E4 ORF negative plasmid band from the pTZhW12 reaction was gel-purified (data not shown). The recombinant pTZhW12 containing the E4 mutant sequences were subsequently named with a general prefix (pTZh) before the mutant name, for example pTZhW12st15/6.
7.6.3 Verification of the pTZhW12st clones

Following *E. Coli* transformation, five colonies were selected for each pTZhW12st E4 mutant. Individual clones were verified by restriction enzyme digest with Apa I, followed by Hind II (Figure 7.5b). Positive clones showing the presence of the 1.4kbp E4 fragment were further digested with BamH I (Figure 7.5b). Clones showing two bands (approximately 3kb and 8kb) on the agarose gel were selected for sequencing (Figure 7.5a). At least two clones for each mutant were sequenced. Appropriate sequencing primers (described in Chapter 2) were used to check each E4 mutation site to ensure no other additional base changes had occurred.

7.6.4 Sequencing pTZhW12st E4 mutants

Following the first round of sequencing, at least one positive clone was obtained for each E4 mutation. A large scale preparation of the wt DNA and each E4 mutant was done. The pTZhW12 clones were sequenced in both directions across the E4 fragment, from positions 3094-4885 (in the coding strand) and 4537'-3210' (in the non-coding strand) (trace sequence data not shown). This was done to ensure that the viral sequence was not disrupted at the Apa I/Hind II restriction sites. Unexpectedly, the sequencing data revealed additional base changes in addition to the mutations introduced by the E4 mutant primers. Table 7.4 shows the nucleotide variations found from the sequence data. Although the E1 ORF of the W12 HPV-16 genome had previously been reported to be similar to the published HPV-16 sequence, it did not seem to be the case for all the viral ORFs. These sequence variations were later discovered to be similar to those identified by Flores *et al.* (1999), who published the sequence of the W12 HPV-16 genome (GenBank accession number AF125673) during the course of this study.
### Table 7.4. Nucleotide variations found in the W12 HPV-16 genome revealed by sequencing of the pTZhW12 clones between positions 3094-4885 (coding strand). These results were compared to the HPV-16 genome sequence (HPV-16R) previously downloaded from the [http://hpv_web.lanl.gov/](http://hpv_web.lanl.gov/) site. **E2** ORF: 2726-3853. **E1^E4** ORF: (865-880)^(3358-3620). **E5** ORF: 3808-4101. **L2** ORF: 4135-5658.

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Nucleotide position</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>All pTZhW12st clones</td>
<td>3410</td>
<td>C -&gt; T</td>
<td>Pro -&gt; Ser</td>
<td>Silent mutation in <strong>E4</strong>. Substitutes Proline with Serine in the <strong>E2</strong> amino acid sequence at position 229.</td>
</tr>
<tr>
<td></td>
<td>3907</td>
<td>extra T</td>
<td></td>
<td>Missing base in original HPV-16 sequence.</td>
</tr>
<tr>
<td></td>
<td>3979</td>
<td>A -&gt; C</td>
<td>ATA -&gt; CTA</td>
<td>Change in <strong>E5</strong> amino acid sequence at position 58.</td>
</tr>
<tr>
<td></td>
<td>4042</td>
<td>A -&gt; G</td>
<td>ATA -&gt; GTA</td>
<td>Change in <strong>E5</strong> amino acid sequence at position 79.</td>
</tr>
<tr>
<td></td>
<td>4114</td>
<td>T -&gt; A</td>
<td></td>
<td>No effect.</td>
</tr>
<tr>
<td></td>
<td>4184</td>
<td>missing A</td>
<td></td>
<td>Frameshift in <strong>L2</strong>.</td>
</tr>
<tr>
<td></td>
<td>4185</td>
<td>missing T</td>
<td></td>
<td>Frameshift in <strong>L2</strong>.</td>
</tr>
<tr>
<td></td>
<td>4228</td>
<td>T -&gt; C</td>
<td>TAC -&gt; CAC</td>
<td>Change in <strong>L2</strong> amino acid sequence at position 32.</td>
</tr>
<tr>
<td></td>
<td>4365</td>
<td>A -&gt; T</td>
<td>GAA -&gt; GAT</td>
<td>Change in <strong>L2</strong> amino acid sequence at position 77.</td>
</tr>
<tr>
<td></td>
<td>4539</td>
<td>T -&gt; C</td>
<td>TCT -&gt; TCC</td>
<td>Change in <strong>L2</strong> aa sequence at position 135.</td>
</tr>
</tbody>
</table>

7.6.5 Re-circularisation of the wt and E4 mutant viral genome

E4 mutant and wt pTZhW12 plasmids were digested with BamH I and the 7.9kb band was gel purified and re-circularised (see Chapter 2). Ligation reactions were checked on a 0.8 % agarose gel which showed a single high molecular weight ligation product in all reactions (Figure 7.5c). This was thought to be the re-circularised HPV-16 genome.
Figure 7.5. Preparations of the re-circularised W12 HPV-16 wt and E4 mutants viral genome DNA.
(a) Sequence data of the pTZhW12st E4 mutants clones. Positions of the stop codons (TAA, TAG, TGA) are boxed in yellow. Mutations were introduced in positions 3384 and 3387 (W12st15/6), 3399 (W12st20), 3435 and 3441 (W12st32/4), 3540 and 3558 (W12st67/3) and 3591 (W12st84) along the W12 E4 ORF. Analysis of the translated mutated E2/E4 gene products confirmed the introduction of termination sequences in the E4 protein sequence, while the E2 protein sequence is not affected. (b) Midipreped DNA checked with enzyme digestions with Apa I/Hind II, and BamH I. pTZhW12st15/6 (lanes 1, 7), pTZhW12st20 (2, 8) pTZhW12st32/4 (lanes 3, 9), pTZhW12st67/3 (lanes 4, 10) and pTZhW12st84 (lanes 5, 11). ND denotes non-digested plasmid DNA. Apa I/Hind II double digest confirmed the presence of the E4 insert fragment following the cloning of the mutated E4 sequences into the parental pTZhW12 plasmid. BamH I digest produced the vector (pTZh19U) and viral genome (W12) bands. (c) Ligated W12 and W12st viral genomes. Lanes 13-19 represents W12, W12st15/6, W12st20, W12st32/4, W12st67/3 and W12st84 re-circularised viral DNA respectively. All ligation reactions showed a single major large molecular weight band, presumably representing single-copied ligated viral genomes.
Chapter 7 - Construction of HPV-16 and CRPV E4 Mutants

(a) Sequences

W12st15/6

(b) Restriction patterns of the pTZW12 clones

W12st20

Apa I/Hind II

W12st32/4

BamH I

(c) Ligation reactions of the pTZhW12 clones

W12st67/3

linear DNA

W12st84

ligation reactions
7.7 Reconstruction of the CRPV E4 mutant viral genome

Similarly, wt and mutated CRPV E4 ORFs were reconstructed back into the pLACRPV parental plasmid, verified by restriction enzyme digestions and sequencing, and re-circularised.

7.7.1 Cloning of mutant E4 fragments from CRst-pET plasmids into pLACRPV

Sac I and Sma I were used to clone the mutated E4 fragment from the CRstE4-pET clones into pLACRPV. Both restriction enzymes do not cut the vector plasmids (pLA-II and pET-23b(+)) but recognise a single cut site in the CRPV genome at positions 1882 (E1 gene) and 3984 (E4/E2 genes) respectively. When analysed by agarose gel electrophoresis, a 2.1kb DNA fragment, which contained the E4 ORF, was apparent in all the digestion reactions. The linearised E4 mutant fragments from the CRstE4-pET clones, and the linearised vector fragment (without the E4 ORF) from the pLACRPV reaction were gel purified. Approximately 7μg of each E4 mutant fragment and 1.5μg of the pLACRPV fragment were ligated overnight (data not shown). The recombinant E4 mutant plasmids were re-labelled with a general prefix (pLA) before the mutant name, for example pLACRst9/0.

7.7.2 Verification of pLACRst E4 mutants by restriction enzyme digestion

Following E. Coli transformation, four colonies from each mutant agar plate were selected. DNA was extracted and digested with Sal I and EcoR I (Figure 7.7b). Clones containing the vector and CRPV genome bands from the Sal I digest were incubated with Xba I, and Sac I/Sma I (Figure 7.7b). Interestingly, EcoR I and Xba I restriction digest patterns of the wt and E4 mutant pLACRPV clones revealed some unexpected nucleotide changes at positions 5031 (asparagine to glutamic acid change) and 5241 respectively (Table 7.6). Since these mutations were also found in the wt sequence, it was assumed to have originated from the particular CRPV strain used in this experiment. At least two positive clones were selected for each E4 mutant,
and subjected to sequencing. E4 mutations were found in all the sequenced clones, without additional nucleotide changes in the E4 ORF.

### 7.7.3 Sequencing pLACRst E4 mutants

One clone of each pLACRst E4 mutant was selected for a more extensive sequencing procedure. Six additional primers (described in Chapter 2) were used to verify the sequence of the cloned Sac I/Sma I fragment. Sequences were checked in both directions and revealed two sites with nucleotide variations (Table 7.6).

In the CRst41 clone, a single base change at position 2588 resulted in a silent mutation in the E1 protein sequence. An unexpected asparagine to serine amino acid mutation in the E1 protein sequence was discovered in all the mutant pLACRst clones as well as wt pLACRPV plasmid.

<table>
<thead>
<tr>
<th>Clones</th>
<th>Method</th>
<th>Nucleotide position</th>
<th>Base change</th>
<th>Amino acid change</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRst41</td>
<td>Sequencing</td>
<td>2588</td>
<td>T -&gt; G</td>
<td>Pro -&gt; Pro</td>
<td>Silent mutation in E1</td>
</tr>
<tr>
<td>All wt and CRst clones</td>
<td>Sequencing</td>
<td>2017</td>
<td>A -&gt; G</td>
<td>Asn -&gt; Ser</td>
<td>Change in primary sequence in E1. Both are polar amino acids.</td>
</tr>
<tr>
<td>All wt and CRst clones</td>
<td>EcoR I</td>
<td>5031</td>
<td>T -&gt; A</td>
<td>Asp -&gt; Glu</td>
<td>Generation of a new EcoR I site. Change in primary amino acid sequence in L2. Both have hydrophilic and acidic properties.</td>
</tr>
<tr>
<td>Xba I</td>
<td>not done</td>
<td>5241</td>
<td>not done</td>
<td>not done</td>
<td>Loss of the Xba I site at 5241-46.</td>
</tr>
</tbody>
</table>

Table 7.6. Spontaneous nucleotide base changes in wt and E4 mutant pLACRPV plasmids revealed by bi-directional automated sequencing and restriction enzyme digests. E1 ORF: 1362-3170. L2 ORF: 4378-5856.
7.7.4 Final plasmid verification and re-circularisation of CRPV viral genome

Large DNA preparations of the wt and mutated E4 pLACRPV plasmids were done. DNA was checked by restriction digestion with Sal I, EcoR I, Xba I and Sac I/Sma I and the patterns obtained were consistent with the known positions of restriction enzyme sites.

Wt and E4 mutant pLACRPV plasmids were digested with Sal I. Linearised viral genomes were gel purified and ligated overnight. All ligation reactions showed a single high molecular weight ligation product on a 0.8% agarose gel (Figure 7.7c). This was thought to be the re-circularised CRPV genome DNA.
Chapter 7 - Construction of HPV-16 and CRPV E4 Mutants

Figure 7.7. Preparation of the re-circularised CRPV wt and E4 mutants viral genome DNA. (a) Sequence data of the pLACRst E4 mutants clones. Positions of the stop codons (TAG, TGA, TAA) are boxed in yellow. Mutations were introduced in positions 3729 and 3732 (CRst9/0), 3825 (CRst41), 3924 (CRst74), 3948 (CRst82), and 3969 (CRst89) along the CRPV E4 ORF. Protein sequence analysis of the E4 and E2 gene products confirmed the introduction of termination codons into the E4 translated product, and only silent changes in the translated E2 sequence. (b) Maxiprepred DNA checked with enzyme digestions with Sal I, EcoR I, Xba I and SacI/Sma I. pLACRPV (lanes 1, 7, 15, 21), pLACRst9/0 (lanes 2, 8, 16, 22), pLACRst41 (lanes 3, 9, 17, 23), pLACRst74 (lanes 4, 10, 18, 24), pLACRst82 (lanes 5, 11, 19, 25) and pLACRst89 (lanes 6, 12, 20, 26). ND denotes non-digested pLACRPV or pLACRst plasmid DNA. Sal I produced 2 bands which represented the vector (3.5kb) and viral genome (8kb). Digestion with the EcoR I enzyme produced two bands instead of a single linearised band. Migration of the digested products and analysis of the CRPV sequence suggest that a possible base mutation (T->A) at position 5031 had occurred in the original wt clone, which generated a second EcoR I recognition site in the L2 gene. The Xba I digestion pattern also indicated a mutation at the 5241-5246 region (L2 gene) which abolished the Xba I site at this position. The double digestion of Sac I/Sma I produced the expected DNA fragments of 2.3kb and 9.2kb. (c) Ligated CRPV and CRst viral genomes. Lanes 29-34 represents CRPV, CRst9/0, CRst41, CRst74, pCRst82 and CRst89 respectively. All ligation reactions showed that majority of the DNA had ligated to give a single large molecular weight band, presumably re-circularised viral DNA.
Chapter 7 - Construction of HPV-16 and CRPV E4 Mutants

(a) Sequences

CRst9/0

CRst41

CRst74

CRst82

(c) Ligation reactions of the pLACR clones

CRst89

(b) Restriction patterns of the pLACR clones

<table>
<thead>
<tr>
<th>Restriction Pattern</th>
<th>Sal I</th>
<th>EcoR I</th>
<th>ND</th>
<th>Xba I</th>
<th>Sma I/Sac I</th>
<th>ND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6</td>
<td>3.5</td>
<td>7.5</td>
<td>1</td>
<td>2 3 4 5 6</td>
<td>7 8 9 10 11 12 13 14</td>
<td></td>
</tr>
<tr>
<td>15 16 17 18 19 20 21 22 23 24 25 26 27 28</td>
<td>1.1</td>
<td>2.3</td>
<td>9.2</td>
<td>10.4</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>

Linear DNA ligation reactions

219
7.8 Discussion

In this chapter, I have outlined the processes involved in the construction of E4 mutant viral genomes for HPV-16 and CRPV. Five E4 mutants, with C-terminal truncation sequences, were generated for each virus by QuikChange mutagenesis. It is predicted that each mutant would express various shortened versions of the E1^E4 proteins from the same start codon of the spliced E1^E4 transcript. Expression of these mutated E4 sequences in vivo was used to investigate the role of the E4 protein during the virus life cycle.

Careful attention was given to the planning of the E4 mutagenesis experiments. Firstly, the choice of PV had to be decided upon. Among the clinically important HPV types considered, HPV-16 was the most appropriate choice because of its frequent association with cervical cancer. In addition, it has been shown that the life cycle of several high-risk HPV types can be supported in differentiating raft cultures (Meyers et al., 1992, 1997; Flores et al., 1999; Thomas, 2001). This advance means that an in vitro system is now available for the analysis of the HPV-16 E4 mutants in the laboratory. In addition to HPV-16, CRPV was chosen not only because it is a widely used animal model, but also because of its precedence for the study of PV protein functions (Barbosa et al., 1988; Brandsma et al., 1991 & 1992; Wu et al., 1994). The availability of an experimental host (the NZW rabbit), which had been shown to be susceptible to CRPV infections, was an added advantage. When compared to the other widely used PV animal models at the time (such as BPV and COPV), CRPV was thought to be the most convenient and economical animal model system for our E4 mutant study.

The design of the E4 mutants was the next obstacle. The overlapping E4 and E2 genes in the PV genome impeded the scope of E4 mutagenesis, since one of the main objectives of the experiment was to prevent the mutagenesis of the E2 protein sequence. Complete or partial deletion mutants were thus impossible to make. The introduction of specific point mutations was
thought to be the only viable strategy for preventing E4 expression. Fortunately, at least five different positions were found within the HPV-16 and CRPV E4 sequences where such mutants could be made. These mutations did not alter the E2 protein sequence. The shortest E4 mutants were predicted to be expressed as peptides of eight amino acids (CRPV) or fourteen amino acids (HPV-16) E1^E4 proteins in vivo. It was assumed that these short peptides would be rapidly degraded by intracellular proteases or proteosomes in the cytoplasm, thus effectively providing an E4 knockout virus.

In terms of the mutagenesis strategies, QuikChange mutagenesis was clearly more efficient than the Kunkel mutagenesis method in my hands. The Kunkel method was laborious and inefficient. In comparison, QuikChange mutagenesis was less time consuming and involved a much simpler and straight-forward protocol. For the Kunkel method, the generation of a single-stranded uracil-enriched DNA template was crucial for success. A lot of time and effort was spent to ensure that every step during the preparation of this single-stranded DNA template was optimal. Despite multiple mutagenesis attempts, this method was unsuccessful. In my opinion, the Kunkel method should be considered out-dated as a tool for site-directed mutagenesis. On the other hand, QuikChange mutagenesis was easy to perform and had an average success rate of greater than 50%. This protocol did not require any complicated template preparation steps, nor the use of specific vector plasmids, thus significantly reducing the amount of time and labour required. The only starting materials required for this method was a double-stranded DNA template and two complementary mutant primers for each mutagenesis reaction.

Once the HPV-16 and CRPV E4 mutant viral genomes were generated, they were distributed to our collaborators for experimentation. The HPV-16 pTZhW12st15/6 E4 mutant and wt plasmid DNA were sent to Dr Paul Lambert who carried out the analysis of the E4 knockout viral genome in transfected human keratinocytes raft cultures. Similar studies of all HPV-16 E4 mutants are currently being carried out in our laboratory by Dr Shirley Southern.
CRPV E4 mutants were sent to Dr Janet Brandsma and Dr Neil Christensen who inoculated them into the skin of CT and NZW rabbits respectively. This mutant study is described in the next chapter.
8.1 Introduction ............................................................................................................. 224
8.2 E4 mutant experiment outline .............................................................................. 226
  8.2.1 Rabbits ............................................................................................................. 226
  8.2.2 CRPV DNA ................................................................................................. 226
  8.2.3 Particle-mediated inoculation apparatus and accessories (Genegun) ........... 226
  8.2.4 Genegun "bullets" ..................................................................................... 227
  8.2.5 Inoculation of DNA ................................................................................... 227
  8.2.6 Papilloma development .............................................................................. 228
8.3 Revised protocol for E4 mutant study in CT rabbits ............................................. 228
8.4 Loss of E4 did not affect the virus ability to induce papilloma development .... 229
  8.4.1 CT rabbit papillomas .................................................................................. 229
  8.4.2 NZW rabbit papillomas .............................................................................. 230
8.5 E4 knockout did not inhibit the development of rabbit papillomas ................. 235
  8.5.1 CT rabbit papillomas .................................................................................. 235
  8.5.2 NZW rabbit papillomas .............................................................................. 236
8.6 Papilloma morphology was not affected by the E4 knockout ......................... 240
8.7 The E4 knockout did not disrupt the early stage of the virus life cycle ............ 242
  8.7.1 PCNA staining patterns .............................................................................. 242
  8.7.2 Cyclin E staining patterns .......................................................................... 242
8.8 The E4 knockout abolished the viral late events .............................................. 245
  8.8.1 Expression of the E4 protein ...................................................................... 245
  8.8.2 Onset of viral DNA amplification ............................................................. 246
  8.8.3 Production of the L1 capsid protein .......................................................... 246
8.9 Discussion ............................................................................................................. 249
Chapter 8 – Analysis of the CRPV E4 Mutants in NZW and CT Rabbits

8.1 Introduction

The function(s) of the PV E4 protein is still unclear. Expression of HPV-16 E4 in cultured cell lines has shown that this protein is able to associate to cytokeratins but not other cytoskeletal filaments, such as the tubulin and actin networks (Doorbar et al., 1991). However, the effect of this association varies between different HPV types. Some E4 proteins can cause the collapse of the cytokeratin network while others do not (Rogel-Gaillard et al., 1993). Recently, our lab has shown that the expression of HPV-16 E4 causes a G2-arrest phenotype in the fission yeast (S. pombe) as well as in HPV carcinoma-derived immortalised human cell lines (HeLa and SiHa), and in non-HPV immortalised primate cell lines (COS-1) (Davy, 2002).

As described in Chapter 5, E4 expression and viral DNA amplification are closely linked events during the PV life cycle. This observation, together with the G2-arrest phenotype in E4 expressing cells, strongly suggests a role for this viral protein during amplification of the viral DNA by affecting the cycling state of the infected cells. Although earlier works in vitro had also proposed roles for E4 in virus maturation and release by affecting cell differentiation, preservation and fragility, no evidence had been found in vivo to support a role for E4 in these processes (Doorbar et al., 1986 & 1991; Rogel-Gaillard et al., 1992).

Mutant E4 proteins have previously been analysed by others by expression in cell lines or yeast. These included deletion and truncation mutants, and mutants with single to multiple base changes. In Chapter 7, five CRPV E4 mutants are described which are predicted to express E4 proteins with variable lengths of C-terminal truncation. In order to investigate the effects of the E4 mutants in the life cycle of the virus in its natural (CT) and experimental (NZW) rabbit hosts, close collaborations were established with Dr Janet Brandsma (Yale University) and Dr Neil Christensen (Hershey Medical Centre, Pennsylvania). An experimental plan was drawn out beforehand for the testing of the E4 mutants so that results obtained from each laboratory
might be comparable. Re-circularised and plasmid CRPV DNA were sent to our collaborators, which were then experimentally inoculated into the cutaneous skin of the rabbit hosts.

This chapter describes the experimental design and results obtained from the CT and NZW rabbit inoculations comparing papilloma development and particular aspects of the viral life cycle between the wt and E4 knockout mutant DNA infections.
8.2 E4 Mutant Experiment Outline

Collaborations with Dr Brandsma and Dr Christensen were initiated to study the effect of CRPV E4 mutants in rabbits. Both collaborators are experts in their own fields of research, especially in the use of the rabbit model system for PV research. In brief, the mutant and wt CRPV DNA genomes were prepared in our laboratory as described in the previous chapter, and sent to our collaborators to be inoculated into the shaven backs of the rabbit hosts. The parameters of the collaboration were drawn up before the start of the experiments and were as follows.

8.2.1 Rabbits

Both NZW and CT rabbits were used for the experiment. Four NZW rabbits were obtained commercially by Dr Brandsma, and three wild CT rabbits were trapped by Dr Christensen. All rabbits were allowed to become accustomed to their new environment before they were inoculated with DNA.

8.2.2 CRPV DNA

Re-circularised and gel-purified CRPV (Washington B strain) DNA were prepared in deionised water to a working concentration of 1μg/μL. The DNA preparations included the wt and five E4 C-terminal truncation mutants, CRst9/0, CRst41, CRst74, CRst82 and CRst89 (See Chapter 7).

8.2.3 Particle-mediated inoculation apparatus and accessories (Genegun)

Dr Brandsma used a prototype genegun, manufactured by Powderject Vaccines (USA), for the inoculation of the NZW rabbits. For the CT rabbits, Dr Christensen used a Helios-Genegun (Bio-Rad Laboratories, USA) for his inoculations. To avoid variations in the inoculation procedure, the same Tefzel tubing was shared between the laboratories and used to make the
DNA “bullets” for the geneguns. Likewise, the same batch of gold particles was used for coating the DNA used in all experiments.

8.2.4 Genegun “bullets”

The DNA “bullets” were prepared by the individual laboratories according to their standard preparation protocol. However, the amount of materials used were standardised. 2μg of DNA was mixed with 1mg of gold. The tubing was coated with 1mg of DNA/gold per inch (1inch = 2.54 cm), and each “bullet” was prepared as 0.5 inch of particle coated tubing. Therefore, each “bullet” would deliver approximately 1μg of DNA per inoculation. Prior to the experiment, the amount of coated DNA per “bullet” was checked for each DNA type by elution and gel electrophoresis. Approximately the same amount of DNA was eluted from each “bullet” of the various DNA types (personal communications).

8.2.5 Inoculation of DNA

Three areas on the dorsal of the NZW rabbits, namely the right, left and rear regions, were shaved for inoculation. Each NZW rabbit received five “shots” per DNA type (five mutants and one wt). The right and left sides of each rabbit were divided into three rows (top, middle and bottom). Each rabbit was inoculated with the wildtype (top), CRst9/0 (middle) and CRst41 (bottom) E4 mutants on the right side. While on the left side, they were inoculated with the CRst74 (top), CRst82 (middle) and CRst89 (bottom) E4 mutant DNA. As an additional experimental control, pLACRPV plasmid DNA was inoculated at the rear of the rabbits. Thus, each NZW rabbit received a total of thirty-three “shots” of seven different types of CRPV DNA. The DNA was delivered at a previously optimised genegun pressure of 350 psi into the epidermis. Unfortunately, no line boundaries were drawn on the backs of the rabbits prior to, or at the time of inoculation, to indicate the respective rows described on each side of the rabbits.
CT rabbits are much smaller in size than the NZW rabbits. The dorsal of each CT rabbit was shaven where a grid consisting of eighteen squares was drawn. Three "shots" of the wt and E4 mutants DNA were delivered per rabbit. A total of eighteen "shots" were given at a genegun pressure of 300 psi.

8.2.6 Papilloma development

A time course for the development of the rabbit papillomas was proposed. Papilloma growth and regression were recorded at weekly intervals beginning on the fourth week post-inoculation.

8.3 Revised protocol for E4 mutant study in CT rabbits

Unfortunately, the first attempt using the CT rabbit DNA for infection did not produce any papilloma even after nine weeks post-inoculation. The reasons for this failure were not pursued but were suspected to be related to the failure to optimise the experiment due to limited rabbit numbers. A revised protocol for the re-inoculation of the CT rabbits was drawn.

In the second attempt, the E4 mutant study was carried out in conjunction with another CRPV mutant (E8 mutant) experiment in Dr Christensen's laboratory. The DNA "bullets" were prepared as described previously. Another three CT rabbits were prepared and inoculated by genegun with bacterially expressed, supercoiled plasmid DNA of the wt CRPV and CRst9/0 E4 mutant in pLA I I vector at the respective marked sites. Recombinant pLACRPV DNA had previously been shown to be able to induce typical epithelial papillomas on NZW rabbit hosts using a jet injector delivery method (Brandsma et al., 1991). Thus, CRPV plasmid DNA was able to induce papillomas much in the same way as re-circularised viral genome DNA. The inoculation area on the dorsal epidermal skin was shaved and divided into a grid of two columns by four
rows. Each rabbit was inoculated with two "shots" of wt DNA on the first row (R1 and L1) and three shots of the CRst9/0 mutant DNA on the second, third and fourth rows of the right column (R2, R3, R4). The second to fourth rows of the left column (L2-L4) were used for the E8 mutant study. Each CT rabbit received a total of eight "shots" delivered at a gene gun pressure of 300 psi. Results obtained from this experiment were analysed.

8.4 Loss of E4 did not affect the virus ability to induce papilloma development

Papillomas were collected from four NZW (Dr Brandsma) and three CT (Dr Christensen) rabbits at the end of the experiments. Whenever possible, wart tissues were halved, and one half fixed overnight in buffered neutral formalin and paraffin embedded, while the other half was cryopreserved. As commented on by Dr Brandsma, the usual take rate for papilloma formation after gene gun inoculations on rabbit skin is usually assumed to be approximately 50%.

All E4 mutant DNA induced papilloma development in both the CT and NZW rabbits. The ability of the E4 mutant DNA to induce papilloma growth was not significantly different from the wt DNA prepared in the same way.

8.4.1 CT rabbit papillomas

Papillomas from the three CT rabbits (labelled A0343, A0399 and A0429) were harvested. Five papillomas were obtained from the wt pLACRPV sites, nine from the CRst9/0 DNA challenged sites, and four from the E8 mutant sites. Table 8.1 shows the take rate of the gene gun inoculation as well as the number of papillomas received from Dr Christensen at the end of the experiment.
Chapter 8 – Analysis of the CRPV E4 Mutants in NZW and CT Rabbits

<table>
<thead>
<tr>
<th></th>
<th>Paps from Rb A0343</th>
<th>Paps from Rb A0399</th>
<th>Paps from Rb A0429</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control wtCRPV</td>
<td>2/2 (R1, L1)</td>
<td>2/2 (R1, L1)</td>
<td>0/2</td>
</tr>
<tr>
<td>CRst9/0</td>
<td>3/3 (R2, R3, R4)</td>
<td>2/4 (R3, R4)</td>
<td>4/3 (R2, R3B, R4A, R4B)</td>
</tr>
<tr>
<td>E4 mutant</td>
<td>NA</td>
<td>1/4 (L4)</td>
<td>3/3 (L2A, L2B, L2C)</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 8.1. CT rabbit papillomas collected at the end of the experiment.
Rb - rabbit. Numbers in bold - the number of papilloma harvested. /n - the number of inoculations delivered. ( ) - the name given to each papilloma.

8.4.2 NZW rabbit papillomas

A large number of papillomas were harvested from four NZW rabbits, numbered #1, #2, #3 and #4 (Table 8.2 & Figure 8.3). To identify the DNA responsible for the induction of each papilloma growth, DNA was extracted from each wart and analysed using the SSCP (Single-Strand Conformation Polymorphism) method by Dr Brandsma and co-workers.

<table>
<thead>
<tr>
<th>CRPV DNA</th>
<th>Paps from Rb</th>
<th>Paps from Rb</th>
<th>Paps from Rb</th>
<th>Paps from Rb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#1</td>
<td>#2</td>
<td>#3</td>
<td>#4</td>
</tr>
<tr>
<td>wt</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CRst9/0</td>
<td>18</td>
<td>7</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>CRst41</td>
<td>12</td>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>CRst74</td>
<td>12</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CRst82</td>
<td>13</td>
<td>16</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>CRst89</td>
<td>1*</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pLACRPV</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>65</td>
<td>40</td>
<td>18</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 8.2. Frequency data of papilloma formation twelve weeks post-inoculation of NZW rabbits collected by Dr Brandsma and co-workers. The CRPV DNA were also referred to as DNA 1 (wt), DNA 2 (CRst9/0), DNA 3 (CRst41), DNA 4 (CRst74), DNA 5 (CRst82), and DNA 6 (CRst89) (Figure 8.3). * - revertant DNA from st89 to wt E4 sequence.

On week thirteen post-inoculation, numerous papillomas had developed in all the inoculated areas on Rbs #1 and #2, whereas very few papillomas had developed on Rb #3. The frequency of papilloma formation for each DNA type on Rbs #1 and #2 were similar. However, a curious difference was observed between Rbs #1 and #2. Only tiny papillomas developed from
the pLACRPV recombinant DNA challenged sites on Rb #2 compared to the large confluent papillomas found at the same sites on Rb #1. In addition, the volume of the papillomas induced by re-circularised wt CRPV DNA were significantly smaller on Rb #2 than Rb #1. It was thought that this papilloma phenotype difference was not due to technical variations during the inoculation of the rabbits, but might be related to differences in the genetic makeup of the hosts. Sequencing of the extracted DNA from some papillomas collected from Rb #1 and #2 were done by Dr Brandsma and co-workers (see Figure 8.3). From the data, DNA reversion was found in one papilloma (number 27) harvested from the lower region on the left side of Rb #1. Papilloma 27 was found to have a wt E4 gene sequence instead of the E4 sequence of DNAs 4-6 (CRst74, CRst82 and CRst89). From its position on the rabbit, it was thought that this isolated papilloma might have originated from DNA 6 and is a wt revertant.

Rabbits #3 and #4 did not support the growth of the papillomas as well as Rbs #1 and #2. Genetic variations between the hosts may have influenced the differences in the outcome of the experiment, since it has previously been shown that the MHC class II genotype of rabbits plays a significant role in determining the course of CRPV infections (Han et al., 1992 & 1994). Thirteen papillomas were harvested from the wt and E4 mutant DNA inoculated sites on Rb #3. Six papillomas developed from the pLACRPV control sites. Unfortunately, the wt re-circularised DNA did not produce any papilloma growth on Rb #3, although the pLACRPV DNA control showed a normal growth pattern. No warts were collected from Rb #4 at the end of the experiment because all papillomas, including those induced by the pLACRPV control, remained small throughout the course of the study and completely regressed at week 8.

Notably, the take rate following CRst89 (last E4 mutant) DNA inoculation was close to zero in all four NZW rabbit hosts. This was most likely due to the poor quality of the DNA preparations used for inoculation. This could be attributed to the deterioration of the DNA
sample during the package delivery process, or the contamination of the DNA during preparation by chemical impurities.
Figure 8.3. Papilloma development following DNA inoculation on NZW rabbits, Rb#1, Rb#2, Rb#3 and Rb#4. DNA “shots” were delivered onto the shaven areas on the right, left and rear of the NZW rabbit hosts. Photographs of papilloma growth on Rbs #1-#3 were taken twelve weeks post-inoculation, while Rb#4 was photographed four weeks post-inoculation. Rbs #3 and #4 were poor hosts for the CRPV DNA infections. Papillomas on Rb#4 completely regressed after seven weeks. The papilloma tissues were harvested, fixed and embedded in paraffin wax in groups as indicated by the thin black boundary line around each group. Each group of tissue paraffin block was labelled alphabetically (i.e. A-Q). Extracted DNA from each papilloma was analysed by SSCP by Dr Brandsma and her co-workers. The SSCP data are represented as \( \text{wtCRPV} \) (for wtCRPV), \( \text{CRst9/0} \) (for CRst9/0), \( \text{CRst41} \) (for CRst41), \( \text{CRst74} \) (for CRst74), \( \text{CRst82} \) (for CRst82) and \( \text{CRst89} \) (for CRst89) on the back of the rabbits. DNA from selected papillomas from Rb#1 were further sequenced (denoted by s) to confirm the SSCP data. Papilloma 27 harvested from the left side of Rb#1 was found to have wt CRPV E4 sequence (black circle), and was thought to be caused by a revertant viral genome originally containing the CRst89 E4 mutant sequence. The low take-rate following CRst89 inoculations in Rbs#1-#3 might be related to the quality of the DNA preparations. Surprisingly, pLACRPV recombinant viral genome induced papillomas (additional positive control) showed variable extents of papilloma growth at the rear of the rabbits. The heterogeneity in the ability of the CRPV DNA (re-circularised or recombinant plasmids) to support papilloma development in the rabbit hosts indirectly indicated the significance of host genetics in determining the course of CRPV infections in the NZW rabbit hosts.
8.5 E4 knockout did not inhibit the development of rabbit papillomas

Papilloma growth was monitored in terms of the timing of wart formation and regression, as well as the three-dimensional sizes of the papillomas with time. Papilloma volume were recorded weekly for the NZW but not the CT rabbits.

The sizes of the papillomas from the wt and E4 mutant DNA challenged sites were variable, even among the same DNA types from the same host, in both CT and NZW rabbits. Papilloma development and regression patterns for each DNA were similar within each rabbit host. From the data provided, it was concluded that the development of the rabbit papillomas was not affected by the expression of truncated forms of CRPV E4.

8.5.1 CT rabbit papillomas

From Dr Christensen’s description, Rb A0399 produced two wt papillomas (one small) and several small E4 mutant papillomas, and Rb A0429 produced one wt and three E4 mutant papillomas. These papillomas were similarly sized when they were harvested approximately ten weeks post-inoculation. It was reported that papilloma regression was observed on Rb A0399 at the time of harvest. Papillomas on Rb A0343 were left for another few weeks before they were harvested. From the growth pattern of the wt and E4 mutant DNA induced papillomas, the E4 knockout DNA did not seem to have an effect on the development of the CT rabbit papillomas.
8.5.2 NZW rabbit warts

The timing of wart formation and regression, as well as the volume of the rabbit papillomas was recorded for each DNA type from weeks four to thirteen by Dr Brandsma and co-workers (Figure 8.4). In general, the average growth rate of the papillomas induced by the different E4 mutant DNA were relatively constant for each rabbit, although the final mean papilloma volume varied between lesions caused by the different DNA types (Figure 8.4b). Inconsistencies in the growth patterns of the papillomas induced by re-circularised wt CRPV DNA made it difficult to evaluate the influence of E4 mutations in the growth rate of the papillomas, even though papilloma formation was clearly unaffected by the E4 mutations.

In terms of papilloma formation and regression, Rbs #1 and #2 showed a consistent course of papilloma development without significant signs of wart regression in the areas inoculated with the re-circularised DNA. In Rb #1, wart formation was first observed at week five at sites inoculated with DNA 1-5, as well as the pLACRPV control sites. The number of wt DNA papillomas was less than half the number found with each E4 mutant DNA type, except DNA 6, which did not produce any papillomas until week seven. Interestingly, a second wave of papilloma formation was seen with Rb #1 as new papillomas appeared on the right side of the rabbit at eight to nine weeks post-inoculation. This observation is unusual and the reason(s) for this is not fully understood. Rb #1 showed a quick regression pattern for the pLACRPV induced papillomas starting from week eleven, while the papillomas from the re-circularised DNA challenged sites were not affected.

Rb #2 developed warts even before the start of the first recorded time point at week four. As with Rb #1, the average number of wt papillomas was lower than that found with each E4 mutant, except DNA 6 which did not produce any papillomas at all. It was not clear why the number of the DNA 5 induced papillomas could be almost four times that of other DNA types at certain time points during the course of the experiment. One probable explanation might be the
uneven coating of the DNA “bullets”, resulting in an increased in the amount of DNA used for the inoculation of Rb #2. A boost in papilloma formation was also seen during weeks eight to nine when the number of papillomas increased by 50% in the DNA 2, and almost 300% in the pLACRPV inoculated sites.

Rabbit #3 showed a long latent period and quick regression during the thirteen-week time course. Papillomas appeared at week five at the DNA 2 and pLACRPV inoculated sites, but appeared after a further two to four weeks on other inoculated sites. Strangely, DNA 1 did not induce any papilloma formation on Rb #3. Papilloma regression was found in all the non-pLACRPV sites on Rb #3 during weeks eight and nine, except one small papilloma induced by DNA 2. A second wave of papilloma formation was also noticed at week nine. This phenomenon is probably comparable to that also seen with Rbs #1 and #2 at approximately the same time point.

Rb #4 showed a very short papilloma growth time course. Papillomas appeared at weeks four and five but quickly regressed after a maximum of three weeks of development. No new papillomas were found after the last regression at week eight.
Chapter 8 - Analysis of the CRPV E4 Mutants in NZW and CT Rabbits

(a)

I) DNA 1

II) Rab #1

II) Rab #2

II) Rab #3

II) Rab #4

CRPV-pLAI II

Time after DNA inoculation (days)

Time after DNA inoculation (days)

Time after DNA inoculation (days)

Time after DNA inoculation (days)
Figure 8.4. Time course of papilloma development on NZW rabbits.
(a) Shows the formation and regression of the papillomas (I) by each DNA type, and (II) by each rabbit. Results showed here are based on the SSCP data of each papilloma. DNA 1-6 represent the re-circularised viral genome with wt, CRst9/0, CRst41, CRst74, CRst82 and CRst89 E4 gene sequences respectively. CRPV-pLAII represents the recombinant viral genome plasmid. Rabbits #1 and #2 were comparable in their ability to support papilloma development after inoculation, whereas Rbs #3 and #4 did not produce any significant numbers of papillomas during as well as at the end of the experiment. Rapid regression also contributed to the lack of papilloma growths in Rbs #3 and #4. The difference in the number of papilloma formation at the CRPV-pLAII sites between Rbs #1 and #2 was a surprising observation. All these observations strongly suggest that the genetic makeup of the rabbit hosts can determine the outcome of the CRPV infections. Comparing the data from Rbs #1 and #2, the E4 mutant DNA 2-5 did not inhibit the formation of papillomas on the rabbits. The overall poor take rate of DNA 6 in the rabbits was most probably related to the sub-optimal quality of the DNA preparation. (b) Shows the mean volume of the rabbit papillomas by each DNA type. These data indicate that there is little variation in the rate of papilloma growth induced by different DNA types on the same rabbits. The main exception is the development of significantly larger papillomas wt CRPV DNA (re-circularised or plasmid) induced papillomas on Rb#1.
8.6 Papilloma morphology was not affected by the E4 knockout

Tissue sections were stained with hematoxylin and eosin counterstains and the histology of the rabbit papillomas were examined for typical pathological effects commonly observed in CRPV infected tissues.

All papillomas (CT and NZW) collected from the study exhibited similar patterns of histological transformation. However, the extents of cytopathic effect in different tissue samples were variable, as the tissues were probably in different stages of papilloma formation at the time of harvest. When wt and E4 mutant papillomas of similar sizes from the same rabbit species were compared, no significant difference in papilloma morphology was found.

Tissue samples from CT Rbs A0429 and A0343, as well as NZW Rbs #1 and #2 were examined for the histological features of CRPV infection. All papilloma tissue sections from the rabbits exhibited acanthosis (thickening of the epidermis), papillae formation, presence of koilocytes, increase in the number of dark granules in the granular layers, preservation of the cellular structures in the differentiated layers and nuclear retention in the cornified layers (parakeratosis). These features showed variation in different papilloma samples. Heterogeneity was found in the extent of papillomatosis and acanthosis, in the number of koilocytes present, and in the number and size of cytoplasmic granules. These differences were probably due to variations in the time of lesion development and in the length of the viral latency period in the different regions of a tissue or different tissues. From the examination of tissue histology, it was apparent that the development of typical pathological features as a result of CRPV infection was similar between the papillomas induced by wt and CRst9/0 (DNA1) DNA (Figure 8.5).
Figure 8.5. Histology of the wt and E4 mutant (CRst9/0) DNA induced CT and NZW rabbit papillomas.

Wt and mutant papillomas from the same rabbit host were compared. Sections were stained with hematoxylin and eosin to show the morphological features of the papillomatous growths. The normal epidermis of the CT and NZW rabbit skin are denoted by N in CTst9/0/10x and NZWst9/0/10x respectively. The CRst9/0 DNA showed an equal ability to induce similar morphological features that were also found in the wt papillomas. Some of these included parakeratosis (thickening of the epidermis), koiocytes formation in the spinous and granular layers, increase in pigmented or dark granules in the granular layer and nuclei retention in the thickened conified layers.
8.7 The E4 knockout did not disrupt the early stage of the virus life cycle

Surrogate markers were used to detect the presence of the early stage of the viral life cycle. As described in earlier chapters, low levels of the early viral proteins E6, E7, E1 and E2, are expressed in the basal cells following virus infection. The presence of E7 could be indirectly shown by the detection of PCNA and cyclin E in the infected tissues.

Similar PCNA and cyclin E staining patterns were found between the wt and E4 knockout (CRst9/0) DNA induced papillomas. This showed that the early stage of the virus life cycle was maintained in the CRst9/0 induced as in the wt induced papilloma.

8.7.1 PCNA staining patterns

In CT and NZW CRPV infected rabbit papillomas, high levels of PCNA expression could be detected from the basal layer to the intermediate or top layers of the epidermis (see Figure 6.2). The up-regulated expression of PCNA detectable by immuno-staining was observed in both the wt and CRst9/0 induced papillomas from both the CT and NZW rabbits (Figure 8.6).

8.7.2 Cyclin E staining patterns

During a productive infection, cyclin E levels in the CRPV infected CT rabbit papilloma were shown to be up-regulated. Cyclin E expression was detected in the parabasal to intermediate layers, but the protein became undetectable in the higher layers, disappearing before viral DNA amplification was detected by FISH (see Figure 6.3). In the CRst9/0 induced CT and NZW papillomas, the cyclin E immuno-staining pattern was similar to that observed in the wt papillomas. However, it was apparent from tissue immuno-stained for cyclin E and viral DNA amplification, that viral DNA was not detected in the CRst9/0 induced papillomas (Figure 8.6).
Figure 8.6. Immuno-detection of surrogate markers of E7 expression (PCNA and cycE) and viral DNA amplification in CRPV wt and E4 knockout (CRst9/0) DNA induced CT and NZW rabbit papillomas.

There was no difference in the staining patterns of PCNA and cyclin E between the wt and CRst9/0 DNA induced papillomas in CT and NZW rabbits. (a) shows similar patterns of up-regulated PCNA expression in wt and CRst9/0 induced papillomas in both CT and NZW rabbits. PCNA (red) was detected from the basal to differentiated layers of the epithelia. Cell nuclei were counter-stained with DAPI (blue). (b) shows a double stain of cyclin E (red) and viral DNA (blue), where cell nuclei were counter-stained with Sytox Green (green). Up-regulated cyclin E expression was detected in the lower half of the epithelium in both the wt and st9/0 DNA induced CT and NZW rabbit papillomas. High levels of viral DNA were detected in cell layers just above the cyclin E positive zones in the wt but not the st9/0 induced papillomas. Dotted lines denote the basal layers.
8.8 The E4 knockout abolished the viral late events

The late stage of the virus life cycle is dependent on the activation of the viral differentiation-dependent promoter. In CRPV infections, this starts from the granular to superficial layers of the epidermis, where viral DNA amplification, E4 and capsid protein expression occurs.

Unlike in the early stage, the loss of the viral E4 protein had a significant effect on the late stage of the viral life cycle. In CT rabbits, CRst9/0 DNA induced papillomas failed to initiate viral DNA amplification and L1 expression. Similarly, E4 knockouts in the NZW rabbit papillomas resulted in the absence of the viral late events altogether.

8.8.1 Expression of the E4 protein

Two different CRPV E4 specific antibodies were used on the rabbit papilloma tissue sections, namely the anti-CRPT2 antisera and anti-MBP-CRE1^E4 polyclonal antisera (see Chapter 4).

E4 was detected in abundance in the cytoplasm and nuclei of sporadic cells in the wt DNA induced CT rabbit papillomas. The staining pattern was typical of CRPV E4 staining patterns observed in other virus infected CT warts (see Figure 5.2). Using the same immuno-staining protocol, E4 could not be detected in any of the CRst9/0 DNA induced CT papillomas, using both the anti-CRPT2 and MBP-CRE1^E4 antisera. This was thought to be due to the rapid degradation of the unstable E4 peptide (eight amino acids long) in vivo.

In the NZW rabbit tissues, immuno-detection of E4 was more difficult and inconsistent, and required signal amplification using ABCComplex and tyramide substrate (see Chapter 2). Fortunately, E4 expression was detected in some wt NZW papillomas. In these tissue sections,
the E4 staining patterns were similar to that seen in the wt CT rabbit papillomas. As with the CT rabbit experiment, E4 was not detected in any of the CRst9/0 DNA induced papillomas (a total of fourteen from Rbs #1 and #2).

8.8.2 Onset of viral DNA amplification

Unfortunately, a double immuno-stain of E4 and viral DNA could not be successfully carried out. Therefore, E4 and viral DNA staining on adjacent tissue sections was used to compare the expression patterns of these two viral events. In the wt CT rabbit tissues, viral DNA amplification and E4 expression were detected in the same areas of the same papilloma tissues (Figure 8.7). This indicated that a close correlation in the occurrence of the two viral events might be present in these tissues. In contrast, viral DNA amplification could not be detected using FISH in the CRst9/0 DNA induced CT rabbit papillomas previously shown to lack E4 expression.

Viral DNA amplification was detected in all the wt NZW papillomas from Rbs #1 and #2. In the sections where E4 expression was detected, viral DNA amplification was found to occur in the same regions of the rabbit tissue sections. Similarly, viral DNA amplification was not detectable in any of the CRst9/0 DNA induced NZW rabbit papillomas from Rbs #1 and #2.

8.8.3 Production of the L1 capsid protein

As expected, L1 was detected in the wt but not the CRst9/0 DNA induced CT rabbit papillomas (Figure 8.7). L1 was expressed in the highly differentiated layers in the same regions previously shown to be supporting viral DNA amplification and E4 expression in the wt CT rabbit papillomas. At this point, our results had showed that the complete productive life cycle of CRPV was supported in the CT rabbit hosts following the inoculation of recombinant wt CRPV DNA. The loss of the full length E4 gene product severely disrupted the initiation of the late stage of the CRPV life cycle in the CT rabbits.
Although the early stage of the viral life cycle has been shown to be similar between the CT and NZW rabbits, the late events in the CRPV life cycle are not fully supported by the NZW rabbit hosts (Shope, 1933). This could be due to the lack of any major transcripts that contained the L1 and L2 ORFs in NZW rabbit papillomas (Phelps et al., 1985). The failure to support the full life cycle of CRPV indicated that the NZW rabbits are secondary hosts to the virus infections (Shope, 1993 & 1935; Ginder, 1952). Since viral infections in these hosts usually resulted in non-productive life cycles, the comparison of L1 expression in the wt and E4 mutant papillomas did not seem appropriate for the NZW rabbit model system.
Figure 8.7. Detection of the viral major late events markers in the wt and E4 knockout (CRst9/0) DNA induced CT and NZW rabbit papillomas.
Viral DNA amplification and L1 expression were abolished in the absence of E4 expression in the CT rabbit papillomas. Similarly, viral DNA could not be detected in the E4 knockout induced NZW rabbit papillomas. (a) shows the complete life cycle of CRPV was supported in the CT papillomas induced by the wt CRPV DNA, while viral DNA and L1 expression were not detectable in the E4 knockout CT rabbit papilloma. (b) shows viral DNA amplification in the NZW rabbit papilloma induced by wt CRPV but not CRst9/0 DNA. Immuno-detection of E4 expression in the NZW rabbit tissues showed low levels of staining in areas that had been found to be supporting DNA amplification. Due to the low immunofluorescent signals and high background levels, enhancement of the digital images was required before E4 staining patterns could be properly visualised in the wt NZW rabbit papillomas. Cell nuclei are represented in blue and epithelial basal layers by dotted lines.
8.9 Discussion

Using an E4 mutant CRPV genome (CRst9/0), we were able to show a direct effect on the late stage of the CRPV life cycle in vivo. The failure of the CRst9/0 induced papillomas to support viral DNA amplification was seen in both the CT and NZW rabbit hosts.

The CRPV-rabbit model is a convenient system for the testing of E4 mutant genomes. The successful use of this system to investigate the role of the E7 gene in the induction of papilloma growth in rabbits had set a precedence for its use in the analysis of PV protein functions (Brandsma et al., 1991). In addition, the use of gene-gun DNA delivery into rabbit epithelium has proven to be useful where CRPV DNA was used in experimental infections of rabbits (Sundaram et al., 1996; Xiao & Brandsma, 1997). The need for an animal model system can be justified by the lack of better alternatives to study the full life cycle of PVs in a differentiated host environment. Although the analysis of aspects of the HPV life cycle can now be studied in vitro using HPV DNA immortalised human keratinocyte cell lines, or cells derived from natural HPV infected carcinoma biopsies, crucial host elements, that may affect the course of a natural virus infection, are lacking in these cell culture systems. These host elements include the activated immune response, and the role of the dermis.

In our experiments, the failure to infect the CT rabbits in the first attempt with re-circularised DNA inoculations could be explained by the following possibilities. Firstly, deterioration of the re-circularised DNA during the shipment to our collaborators may have affected the quality of the DNA “bullets” and hence the efficiency of the DNA infection. Secondly, differences in the structure of the epithelial tissue of CT and NZW skin suggest that separate experimental parameters may be required for the infection of each rabbit species. Therefore, the use of the 2.1 micron gold particles, which had been shown to be optimal for NZW rabbit skin inoculation, might not be suitable for use on the more delicate CT rabbit
epidermis, which is known to be thinner than that of the NZW rabbit (personal communication with Dr Breitburd, Pasteur Institute, France). Since the localised delivery of the DNA-gold particles to the basal layer of the epidermis is crucial for the initiation of the viral infection, appropriate adjustments should have been made to the materials used for the inoculation of the CT rabbit skin.

The course of CRPV infection in CT and NZW rabbits varies significantly between the two rabbit species. For instance, the progression to carcinogenesis is 75% in NZW rabbits, and only 25% in CT rabbits (Syverton, 1952). RNA transcript mapping in CT and NZW papillomas and carcinomas has also revealed significant differences in CRPV transcription patterns in the different rabbit species. In NZW rabbit infections, E6 and E7 transcript levels were found to be expressed in almost equal amounts in the papillomas, with levels gradually decreasing with epithelial differentiation (Zeltner et al., 1994). L2 transcripts were undetectable in the NZW rabbit papillomas, and only isolated cells contained L1 transcripts (Zeltner et al., 1994). In contrast, in CT rabbit papillomas, the levels of E7 transcripts were higher than those of E6, and transcript levels generally increased with differentiation. In addition, L1 and L2 transcripts were expressed in abundance in the granular layer in these papillomas (Zeltner et al., 1994). Despite the variations in viral gene transcription patterns, immuno-detection of surrogate markers of E7 expression in the CT and NZW rabbit papillomas showed that the early stage of the CRPV life cycle is conserved. A major difference in CRPV infections between the two rabbit species is the poor productivity of the NZW rabbit papillomas. However, the use of the two rabbit model systems demonstrated a universal role of CRPV E4 in the late stage of the CRPV life cycle. Viral DNA amplification was abolished in the differentiated layers when the expression of the full-length E4 protein is substituted with an eight amino acid E1^E4 peptide.

The early stage of the CRPV life cycle does not seem to be affected by loss of full-length E4 protein expression. This was demonstrated by the immuno-staining patterns of PCNA
and cyclin E in the wt and CRst9/0 DNA induced CT and NZW rabbit papillomas. Despite the detection of the E4 ORF in early viral transcripts, in undifferentiated CRPV and HPV DNA immortalised cell lines, the immuno-detection of E4 expression was found only in the differentiated cells in rafts and tissues (Danes et al., 1985; Dostatni et al., 1988; Rohls et al., 1991, Hummel et al., 1992; Pray & Laimins, 1995; Doorbar et al., 1997; Chapter 5). However, the expression of very low levels of E4 during the early stage of the virus life cycle, beyond the sensitivity of the immuno-detection methods, cannot be dismissed. Nonetheless, our data here support the fact that the CRPV E4 protein is not required for papilloma formation, as well as the early stage of the virus life cycle in vivo.

Our E4 mutant analysis also showed a close correlation between CRPV E4 expression and viral DNA amplification, and suggests that E4 may have a crucial role in the initiation of CRPV DNA amplification. Since the two viral events coincide exactly in human and animal PV infected tissues (Peh et al., in press; Chapter 5), the role of E4 during the late stage of the virus life cycle may be conserved among different PV types. Interestingly, previous efforts to isolate a major late transcript containing the E1^E4 ORF from CRPV infected tissues have not been successful (Nasseri & Wettstein, 1984; Wettstein et al., 1987; Zeltner, 1994). Instead, CRPV E4 was shown to be expressed in abundance in the granular layer of CRPV infected CT rabbit papillomas using CRPV E4 specific antibodies (Chapters 3 & 5). We are unable to determine the exact role of E4 in the virus life cycle and further analyses of different CRPV E4 mutants induced papillomas may be required.

Immuno-detection of E4 and L1 in PV infected tissues has shown that the expression of E4 always proceeds L1 (Chapter 5; Peh et al., in press). The lack of L1 detection in the CRst9/0 CT papillomas suggests a link between the expression of E4 and L1. L1 expression has been found to be regulated at the post-transcriptional level by inhibitory RNA sequences and RNA processing (Schwartz, 1998). Different regulatory mechanisms are likely to be involved in the
regulation of L1 expression in different PV infections. This was demonstrated by the premature induction of L1 expression in differentiated cells via different mechanisms, such as changes in codon usage and removal of RNA inhibitory sequences (Zhou et al., 1999; Collier et al., 2002). It is difficult to speculate a role for E4 in the regulation of L1 expression during a productive CRPV infection, since very little is known about the mechanism which governs the expression of the capsid proteins in rabbit papillomas.

From the analyses of the wt and CRst9/0 DNA induced papillomas, we have showed that the loss of E4 did not affect papilloma growth and morphology, as well as in the early stage of the CRPV life cycle in CT and NZW rabbits. CRPV E4 plays a crucial role in the initiation of viral DNA amplification, and may be involved in the regulation of L1 expression. However, the transcription patterns of the wt and CRst9/0 DNA induced papillomas must be analysed to ensure that the observations reported here are due to the direct effect of loss of the E4 protein, and are not a result of a failure to transcribe the viral late mRNAs.
CHAPTER 9 – FINAL DISCUSSION
PVs infect and replicate only in differentiated epithelial cells. They are highly host specific and target specific tissue sites to cause infections. This specific tropism of the PVs for particular epithelial tissues indicates a closely dependent relationship between the virus life cycle and the host cell. From the initial infection of the basal keratinocyte to the development of a papilloma, the expression of viral transcripts and proteins are strictly regulated events which enable the completion of the viral life cycle. The events in the viral life cycle can be divided into the early and late stages, in which the early and differentiation-dependent viral promoters are activated respectively. Activation of the early viral promoter(s) takes place in the lower layers of the differentiated epithelium following the infection of basal keratinocytes, and leads to the expression of low levels of E6, E7, E1 and E2 (reviewed in Howley & Lowy, 2001). Activation of the differentiation-dependent promoter is less well-understood. Besides the requirement for keratinocyte differentiation, the exact events that lead to its activation are still unclear. However, it is certain that viral-host interactions must occur at the appropriate times in order to create a suitable cellular environment for genome amplification and virus synthesis.

One way in which we can better understand the viral life cycle is to study the relationship between individual viral events, as well as the host cell changes that occur during an infection. Altered patterns of host protein expression or protein localisation can be used to predict the altered cellular states created in the virus infected cell. By studying virus and host protein expression patterns, it is possible in some instances to observe the consequence of protein associations by immunofluorescence. Examples of these include the association of E7/Rb binding which can be visualised by the up-regulation of E2F activated gene products, such as PCNA, cyclin A and cyclin E.

*Different PV types exhibit the same sequence of events during productive infection*

The expression of E6 and E7 during the early stage of the virus life cycle disrupts the cell cycle of the infected cells by interfering with the functions of regulatory proteins such as p53.
and pRB (Dyson et al., 1989; Munger et al., 1998b; Werness et al., 1990; Kessis et al., 1993; Slebos et al., 1994). E6-p53 binding targets p53 for rapid degradation in a ubiquitin-dependent pathway (Scheffner et al., 1990). The loss of p53 blocks the cell's ability to undergo apoptosis and growth-arrest via p53-dependent pathways, thus disabling a major mechanism that guards against cell transformation (Kessis et al., 1993). At the same time, E7 binds pRB which leads to the up-regulation of E2F transcription factor responsive S-phase genes (Martin et al., 1998; Harbour & Dean, 2000). The combination of E6 and E7 expression alters keratinocyte differentiation by maintaining cells in a replication competent state and by disabling the apoptosis and growth-arrest checkpoint responses (Munger et al., 1989a; Hawley-Nelson et al., 1989).

It is possible to monitor PV induced cell transformation through the immuno-detection of host cell markers of viral oncogene activity such as PCNA and cyclins E and A. Increased levels of these G1/S-phase related proteins, which result from the expression of E7, are apparent in human and animal PV infections, as seen in this other studies (Zerfass et al., 1995; Lu et al., 1996; Martin et al., 1998; Southern & Herrington, 1998; Southern et al., 2000; Chapter 6). The up-regulated expression of these G1/S-phase proteins in different PV infected tissues suggests similar virus induced cellular gene expression during the early stage of the viral life cycle. However, the extent of alteration in the PCNA and cyclin E expression patterns and resultant keratinocyte hyper-proliferation are different in the different tissue types (Figures 6.2 & 6.3). The effects of CRPV E6 and E7 expression in the cutaneous rabbit skin, which causes an enormous amount of keratinocyte proliferation and the formation of deep folding papillae, is far more dramatic than the development of the small oral lesions caused by ROPV on the under-surface of the rabbit tongue. CRPV infected tissues show considerable epithelial thickening in the differentiated layers, as well as the up-regulation of PCNA expression which is apparent in cells in the uppermost layers of the papillomas (Figures 8.4 & 6.2b). Lesions induced by ROPV showed a more subtle effect with regard to tissue thickening, although high levels of PCNA
expression was apparent (Figure 6.2a). These differences may be the result of differences in the potency of the viral oncoproteins in the different rabbit PV types. A difference in the Rb binding affinity of the HPV-6 (low-risk) and HPV-16 (high-risk) E7 proteins has been previously demonstrated in vitro (Gage et al., 1990; Munger et al., 1989). The weaker association of HPV-6 E7 protein and pRB is thought to reduce its ability to disrupt the cell cycle and to cause cellular transformation (Demers et al., 1994 & 1996; Halbert et al., 1992). The dramatic keratinocyte hyper-proliferation in CRPV infected tissues may be due to the more potent transforming ability of its oncoproteins.

Keratinocyte proliferation during the early stage of the virus life cycle, is generally followed by activation of the viral late events in the differentiated layers of the epithelium. By immuno-staining different PV infected tissue samples, the late events were also found to exhibit conserved features with regards to the timing of viral DNA amplification, and the expression of E4 and L1. Consistent with previous findings in HPV-16 and HPV-1 induced tissues (Doorbar et al., 1997), E4 expression and viral DNA amplification were found to correlate exactly in all the animal and human infected material examined (Peh et al., in press; Figure 5.3). The late stage of the PV life cycle can be subdivided into two phases based on the time at which the different viral events are triggered. The first phase begins when E4 is expressed and viral DNA amplification begins. The second phase is marked by the expression of the viral capsid proteins (L1 and L2). Exactly why E4 expression coincides with the onset of vegetative viral genome amplification remains unclear.

In Chapter 8 of this thesis, experimental infections of rabbits with a re-constructed E4 knockout CRPV genome suggest a crucial role for E4 in the late stage of the CRPV life cycle. Loss of E4 resulted in the loss of viral DNA amplification and L1 expression in both the CT and NZW rabbit infections. L1 expression is regulated by molecular mechanisms that are still poorly understood. Recent work has shown the presence of inhibitory sequences (negative regulatory
elements; NREs) in HPV-16 L1 transcripts that prevent L1 protein expression in the lower epithelial layers (Collier et al., 2002). Despite the discovery of similar inhibitory sequences in multiple HPV types, the same mechanism may not be used by all PV types (Collier et al., 2002). In addition to being regulated by NREs, the expression of BPV-1 L1 was shown to be influenced by the change in mRNA codon usage in the terminally differentiated epithelial cells (Zhou et al., 1999). Even though different mechanisms may be involved in the control of L1 expression, the observation that E4 expression and viral DNA amplification precede L1 expression holds true for all the animal and human PV infections studied here. The link between the abundant expression of viral DNA and E4 in cells that will eventually trigger the expression of the viral capsid proteins still remains to be investigated. Our analysis of the CRPV E4 knockout mutant suggest that E4 may be indirectly involved in the initiation of L1 expression, in addition to having a role in viral DNA amplification.

An overlap between the occurrence of the loss of early markers and the first appearance of late markers was apparent in all the PV infections studied. Detection of PCNA and viral late markers by double staining revealed a region where a subset of PCNA-positive cells were supporting viral DNA amplification and E4 expression (Figure 6.2). This overlap region suggests the presence of a phase during the viral life cycle when the virus can amplify its DNA in cells that contains PCNA-dependent replication machinery. The extent of the overlap region varied between different PV infections however. The productive ROPV infected lesions had the most extensive overlap, whereas the HPV-11 infected lesions contained a considerably smaller population of PCNA/viral DNA double-positive cells (Figure 6.2). ROPV L1 expression is triggered in the intermediate layers whereas HPV-11 L1 is expressed in the superficial epithelial layers (Figure 5.5). We speculate that the larger PCNA/viral DNA(E4) double positive cell population and greater number of cells expressing L1 indicates that ROPV is well-adapted to its host environment.
Animal PVs and experimental animal models are useful tools for the study of HPV infections

HPV infections cause epithelial diseases which range from benign cutaneous warts (common warts) to cervical cancers in women. The clinical importance of HPV infections is recognised in diseases caused by the high-risk type mucosal-tropic HPVs, such as HPV-16, HPV-18 and HPV-31. High-risk HPV types are thought to cause over 99% of all cervical cancers. Although not as frequently associated with development of cancer, low-risk HPV types, such as HPV-11, HPV-6 and HPV-44 are important for causing genital papillomas. The outcome of PV infection is affected by the immune response of the infected host (Schiffman, 1994). This is most clearly shown by the increased risk of HPV infection and cancer progression in renal transplant recipients (Alloub et al., 1985; Arends et al., 1997; Sillman et al., 1997) and HIV-infected individuals (Palefsky et al., 1999; Sun et al., 1997; Vermund et al., 1991; Wright et al., 1994). The immune system is also crucial in protecting against re-infection, in controlling the development of diseases and in the regression of existing HPV infections (Han et al., 1992; Brieitburd et al., 1996; Frazer, 1996; Malejczyk et al., 1997). For this reason, animal PVs and experimental animal model systems have become necessary tools for monitoring disease development, especially during immunisation studies as well as therapeutic and prophylactic treatment assessments for PV infections. Rabbits, dogs and transgenic mice are three of the most widely used animal models for PV vaccine development (Brietburd et al., 1997; Stanley et al., 1997). Immunisation studies using virus-like particles (VLPs), DNA and viral proteins have proved successful in protecting animals against infection or re-infection by animal and human PVs (Ghim et al., 1995; Jochmus et al., 1999; Marais et al., 1999; Han et al., 1999; Christensen et al., 2000; Stanley et al., 2001; Yuan et al., 2001). These studies are important for the development of treatments for clinically significant HPV infections.

From the study of infected tissues from dogs, rabbits, ox and murine xenografts, a general pattern is apparent in the way the virus life cycle is organised, that is similar to the pattern
observed in different HPV infections (Doorbar et al., 1997, Peh et al., in press). Although the order of life cycle events is conserved, the timing of events appear to differ in lesions caused by different PV types. Consistency amongst the virus life cycle patterns is probably related to the specificity of the virus-host interactions and mechanisms involved in generating a viable environment in which the virus life cycle can be supported (Chien et al., 2002). When choosing an appropriate animal model for the study of HPV infections therefore, tissue specificity and the regulation of life cycle events should be considered. Comparing life cycle events in mucosal tropic animal and human PVs suggests that ROPV is an appropriate model for HPV-11 infections. Both PV types can infect oral and genital sites, which usually results in limiting productive infections in their hosts. They also have comparable PCNA and cyclin E staining patterns, as well as similar patterns of viral late markers (Chapters 5 and 6). Recent immunisation studies have shown that COPV is a potentially useful animal model for the development of vaccines against mucosal PV infections (Ghim et al., 1995; Stanley et al., 2001; Yuan et al., 2001). The main drawbacks with the COPV model however, are the unusual genomic organisation of COPV, and the fact that its viral late events are initiated in the basal layer (Delius et al., 1994; Nicholls et al., 2001). Interestingly, COPV shows homology at the protein and DNA sequence levels with the cutaneous human viruses, HPV-1 and HPV-63 (Delius et al., 1994; Chapter 3). The onset of viral late events in the lower layers of the epithelium and the similarities in the life cycle organisation support the observation that COPV is evolutionarily related to HPV-1 and HPV-63 (Peh et al., in press). The most extensively studied animal model systems are the cutaneous viruses CRPV and BPV-1/BPV-2. The high frequency of progression to cancer in NZW rabbit hosts makes the CRPV-NZW rabbit system a practical model for the study of carcinogenesis and for the identification of therapeutic strategies (Breitburd et al., 1997). CRPV is also widely used in vaccine development and for viral protein function studies. Although the life cycle of the virus does not appear to resemble any cutaneous HPV types studied so far, this model remains extremely useful.
As an alternative, the transplantation of infected human and animal tissue into murine models has provided a different approach to the study of PV infections (reviewed in Bonnez, 1998; Pawellek et al., 2002). Developments in the kidney capsule and skin transplantation methods in NUDE and SCID mice respectively, have enabled the full life cycle of several HPV types to be supported in these animals. Immuno-staining of such lesions revealed that the order of viral events are preserved, and are similar to events that take place following infection of their natural hosts (Chapter 5). Examination of tissue histology showed evidence of papillae formation and neoplastic characteristics in both the NUDE and SCID xenograft models. These characteristics were most apparent in the SCID model. Furthermore, the SCID mouse model enables the outgrowth of the papillomas to be constantly monitored, and allows the topical application of putative antivirals at regular intervals during therapeutic studies. These aspects cannot be achieved in the internal implantations, unless the NUDE mice are sacrificed at regular time points.

**Hypothesis regarding the mechanisms involved in the regulation of the PV late events by E4**

Work from our laboratory and others have shown that E4 can affect progression through the cell cycle and cause cell cycle arrest. Expression of HPV-16 E4 in fission yeast (*Schizosaccharomyces pombe*) and in mammalian cells in monolayer culture inhibits mitotic entry and results in a G2 phase arrest phenotype (Davy, 2002). As discussed by Davy (2002), it has been suggested that the expression of HPV-1 E4 can cause an arrest in the G1 phase. Although the exact mechanisms by which the proteins work to affect the cell cycle is unclear, both HPV-16 and HPV-1 E4 are shown to be able to bind cyclin B and mcm (mini chromosome maintenance) *in vivo* respectively. So how do the cell cycle arrest events in E4 expressing cells relate to the viral life cycle?
From E4 mutant studies of CRPV in CT and NZW rabbits, we have shown that E4 has an essential role in the initiation of viral DNA amplification and L1 expression during the viral life cycle (Chapter 9). Recently, work in our laboratory has also shown that the in vitro expression of CRPV E4 in monolayer cell culture can cause cells to arrest in the G2 phase of the cell cycle, similar to that observed following the expression of the HPV-16 E4 protein in cells in culture (Davy, personal communication). A possible mechanism by which HPV-16 and CRPV E4 may work to establish a block in the G2/M transition phase is by sequestering cyclin B in the cytoplasm. The sub-cellular localisation of the cyclin B/cdc2 complex is a crucial regulatory event in the control of mitotic entry (Pines & Hunter, 1991). Phosphorylation of cyclin B by activated cdc2 stimulates the translocation of the protein complex from the cytoplasm into the nucleus, thus initiating mitotic events (Kishimoto & Okumura, 1997; Ferrel, 1998). As a predominantly cytoplasmic protein, E4 may prevent cyclin B/cdc2 nuclear import by binding to the cyclin B protein complex and tethering it in the cytoplasm. In support of this hypothesis, several HPV E4 proteins, including that of HPV-16, have been shown to be able associate with the cytokeratin network in monolayer cultured cells and to be predominantly cytoplasmic (Doorbar et al., 1991; Roberts et al., 1993). In fact, co-localisation between E4, keratin, and cyclin B has recently been shown (Davy, 2002). Although this triple localisation pattern has not yet been shown in natural HPV-16 infected lesions, both E4 and cyclin B can be found together in HPV-16, COPV, ROPV and CRPV infected lesions (Chapter 6). This suggests that E4 may act to arrest cells in the G2 phase in vivo, possibly by sequestering cyclin B in the cytoplasm during the late stage of the viral life cycle. Cytoplasmic retention sequences such as that which anchors E4 to the cytokeratin network, may also allow E4 to retain the cyclin B/cdc2 protein complex in the cytoplasm. In addition to the effects of E4, the expression of E6 and E7 enables cells to undergo endoreplication without completing mitosis by abrogating the mitotic spindle checkpoint and increasing the levels of the cellular oncoprotein MDM2, a cellular oncoprotein (Thomas & Laimins, 1998). It appears that the viral proteins may act to ensure cell cycle arrest in cells where the viral late events are being initiated.
Chapter 9 – Final Discussion

By Immuno-staining for surrogate markers of E7 expression in the rabbit papillomas, we conclude that the early stage of the CRPV life cycle is not affected by the loss of E4 expression (Chapter 8). It is thought that the mutant CRPV still retains its ability to replicate its genome at low levels during the late stages of the viral life cycle in the presence of E1 and E2. However, the failure to detect viral DNA by FISH in the differentiated layers indicates an absence of viral DNA amplification as a consequence of the loss of CRPV E4 expression. This suggests that CRPV E4 may have a role in affecting the host DNA replication machinery. The replication licensing system (reviewed by Tada & Blow, 1998) acts to ensure that the host genome is replicated only once in a single cell cycle, and is made up of two main factions, namely the replication licensing factors (RLFs) -M and -B. The RLF-M consists of complexes containing all six members of the MCM/P1 protein family, while the components of RLF-B are still unknown. Other factors required for replication licensing include the origin recognition complex (ORC), cdc6/cdc18 and cdk activities (Tada & Blow, 1998). Each licensing component is dependent on another to bring about DNA replication, and the regulation of many of these proteins are regulated according to their phosphorylated or dephosphorylated states, as well as by their subcellular localisation. Although it is speculative to suggest that E4 may have the potential to affect cdk activities and subcellular localisation through its cyclin binding capability, this possibility cannot be ruled out without further investigation. E4 may also interfere directly with the host replication system by associating with replication proteins, such as mcm proteins, so that components of host replication machinery may be made available for viral rather than cellular DNA replication.

*E4 is likely to have evolved to adapt to different viral life cycles*

Based on E4 protein sequence analysis of several PV types (Chapter 3), the E4 protein appears to share little primary sequence homology between the different virus types. Even so, some protein characteristics are still conserved. The data in Chapter 3 showed that the E4 proteins of different viruses have multiple phosphorylation sites, and that viruses from the same
phylogenetic classification groups share similar motif sequences at similar positions. Moreover, features such as the proline-rich domains and charged regions are universal in all E4 proteins as previously published (Doorbar et al., 1989; Doorbar & Myers, 1996). Preservation of these features suggests that E4 may have a common function in the different virus life cycles. This assumption is unlikely to be totally true since physical properties, such as cytokeratin association and protein binding ability, as well as the intracellular expression patterns and cytopathic effects of E4 (Chapter 5), have been shown to be variable between viruses. The presence of various post-translational modifications, including hyperphosphorylation, proteolysis and multimerisation, have been proposed to influence the role of E4 during the viral life cycle (Grand et al., 1989; Rogel-Gaillard et al., 1992; Ashmole et al., 1998). HPV-1 E4 proteins differing in their N-terminal sequences were shown to localise to different cellular compartments, and this was suggested by the authors to be a possible way in which E4 might play several roles during the viral life cycle (Rogel-Gaillard et al., 1992). The importance of phosphorylation in regulating protein activity, including protein translocation and the overall charge of the protein, is widely recognised (Jans & Hubner, 1996). The presence of putative phosphorylation sites at the potential NLS and NES within E4 suggests that phosphorylation may be a mechanism by which the cellular localisation of E4 is regulated (Jans & Hubner, 1996; Kaffman and O'Shea, 1999; Chapter 3). However, the heterogeneity in the intracellular localisation patterns of E4 in lesions caused by different virus types (Figure 5.6) suggests that the movement of E4 within the cell may be regulated to varying extents by the different viruses. It is still not known if the role(s) of E4 are governed by its subcellular localisation. Translocation of E4 between the nucleus and cytoplasm may occur since mutant studies of HPV-1 and HPV-16 E1^E4 showed that different E4 mutants can exhibit varying extents of nuclear and cytoplasmic localisation in vitro (Rogel-Gaillard et al., 1992; Roberts et al., 1994 & 1997; Doorbar et al., 2000). Immuno-detection of E4 in CRPV and COPV induced lesions also showed that the proteins can exist at different levels (based on the staining levels of the proteins) in the nucleus and cytoplasm (Chapter 5). It is possible that the properties of E4 may have evolved according to the tissue type and host
specificity of the different PV types. The non-crossreactivity of specific E4 antibodies to the E4 proteins of different PV types (Chapter 4) also indicates that significant modification in the tertiary structure and antigenicity of the proteins have occurred during the evolution of these viruses. Together, these points may suggest that the E4 protein of different PV types may be regulated (or even function) via different molecular mechanisms during the viral life cycle of different viruses. It appears from the data obtained in this thesis and other work being carried out in the lab that E4 may interfere with cell cycle progression in order to allow the completion of the virus life cycle.
## APPENDIX 1 –SEQUENCES

<table>
<thead>
<tr>
<th>Appendix 1/I</th>
<th>BLAST2.0 E1^E4 protein alignment</th>
<th>266</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 1/II</td>
<td>A) Nucleotide changes introduced into the HPV-16 E4 mutant genomes [E1^E4 ORF: (865 - 880)^{(3357 - 3619)}]</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>C) Nucleotide changes introduced into the CRPV E4 mutant genomes [E1^E4 ORF: (1362-1371)^{(3714-4015)}]</td>
<td>269</td>
</tr>
</tbody>
</table>
Appendix 1/1

Selected list of proteins obtained from the basic local alignment sequence (BLAST2.0) entry of (a) BPV-1, (b) CRPV, (c) ROPV, (d) COPV, (e) HPV-11 E1'-E4 protein sequences.
### Appendix 1/1

#### 1) HPV-21 complete genome (E4) (136aa) (U17797|AA79398)
- Identities: 33/117 (28%), Positives: 43/117 (36%), Gaps: 26/117 (22%)
- **Query**

<table>
<thead>
<tr>
<th>6</th>
<th>VPPEP</th>
<th>R</th>
<th>P</th>
<th>PP R</th>
<th>P</th>
<th>PP R</th>
<th>P</th>
<th>PP R</th>
<th>G</th>
<th>G</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>LHP</td>
<td>W + N *</td>
<td>P</td>
<td>PP P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
</tr>
</tbody>
</table>

#### 2) HPV-5 E4 protein.
- Identities: 40/142 (28%), Positives: 52/142 (36%), Gaps: 41/142 (28%)
- **Query**

<table>
<thead>
<tr>
<th>7</th>
<th>PP P</th>
<th>R</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>LHP</td>
<td>W + N *</td>
<td>P</td>
<td>PP P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 3) Pseudorabies virus ORFl, ORF2, AND ORF3. (1958aa) (M57505|Q69340)
- Identities: 31/91 (34%), Positives: 36/91 (39%), Gaps: 24/91 (26%)
- **Query**

<table>
<thead>
<tr>
<th>7</th>
<th>PP P</th>
<th>R</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>LHP</td>
<td>W + N *</td>
<td>P</td>
<td>PP P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 4) Nucleolin (Protein C23) [Gallus gallus] (694aa) (P15771|NUCL_CHICK)
- Identities: 31/91 (34%), Positives: 36/91 (39%), Gaps: 24/91 (26%)
- **Query**

<table>
<thead>
<tr>
<th>19</th>
<th>PP P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>LHP</td>
<td>W + N *</td>
<td>P</td>
<td>PP P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
</tr>
</tbody>
</table>

#### 5) Common chimpanzee papillomavirus type 1 complete genome (E4).
- Identities: 47/94 (50%), Positives: 61/94 (64%), Gaps: 11/94 (11%)
- **Query**

<table>
<thead>
<tr>
<th>7</th>
<th>PP P</th>
<th>R</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>LHP</td>
<td>W + N *</td>
<td>P</td>
<td>PP P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 6) Pygmy chimpanzee papillomavirus type 1 E4 protein.
- Identities: 43/94 (45%), Positives: 56/94 (38%), Gaps: 10/94 (10%)
- **Query**

<table>
<thead>
<tr>
<th>7</th>
<th>PP P</th>
<th>R</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>LHP</td>
<td>W + N *</td>
<td>P</td>
<td>PP P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td>G</td>
<td>G + + P</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 1/A. Nucleotide changes introduced into the HPV-16 E4 mutant genomes [E1^E4 ORF: (865 - 880)^3357 - 3619]).

- represent the location and sequence of the stop codons introduced into each E4 mutant. **Bold and underlined codons** represent the original sequence of the wt E4 ORF. * represents the original stop codon of the E4 ORF.
Appendix 1/II/B. Nucleotide changes introduced into the CRPV E4 mutant genomes [E1^E4 ORF: (1362-1371)^3714-4015].

() - represent the location and sequence of the stop codons introduced into each E4 mutant. **Bold and underlined codons** - represent the original sequence of the wt E4 ORF. * - represents the original stop codon of the E4 ORF.
APPENDIX 2 – DNA PLASMID MAPS

Appendix 2

A) pBR322 plasmid map ................................................................. 271
B) pSp64 plasmid map ................................................................. 272
C) pET-23b(+) plasmid map .......................................................... 273
D) pTZ19U plasmid map ............................................................... 274
E) pGex-4T-1 plasmid map ............................................................ 275
F) pMal-c plasmid map ................................................................. 276
Appendix 2/A. Plasmid map of pBR322 cloning vector (DNA map obtained from http://www.fermentas.com).
The pLA II plasmid was a derivative of pBR322, without the tox gene. CRPV was cloned into pLA II at the Sal I site.
Appendix 2/B. Plasmid map of pSp64 cloning map (Promega Ltd, UK). The HPV-16 genome was originally cloned into pSp64 vector at the BamHI site.
Appendix 2/C. Plasmid map of pET-23b(+) cloning vector (ON Biosciences Ltd, UK). pET-23b(+) was used to synthesis parental templates for Kunkel and QuikChange site-directed mutagenesis of HPV-16 and CRPV E4. The HPV-16 E4 ORF was cloned between the BamH I and Hind II sites, while CRPV E4 ORF was cloned between the Sal I and EcoR I sites. The helper phage f1 ori was required for the synthesis of single-stranded DNA templates for Kunkel site-directed mutagenesis.
Appendix 2/D. Plasmid map of pTZ19U cloning vector (DNA map obtained from http://www.fermentas.com).

The W12 strain HPV-16 genome was re-cloned into pTZ19U. pTZ19U is a modified version of pTZ19U, which lack the Hind II enzyme recognition site.
Appendix 2/E. Plasmid map of pGex4T-1 expression vector (Amersham-Pharmacia Biotech, UK).

This vector contained an EcoR I and a Sal I recognition sites in the multiple cloning region which were used to clone E1-E4 PCR amplified products (Figure 4.1). Bacteria transformants were selected for ampicillin resistance (Amp⁵). Expression of the GST fusion gene transcripts were triggered from the IPTG inducible tac promoter.

pMal-c2 vectors are identical to the pMal-p2 vectors except for a deletion of the malE signal sequence. This vector was used similarly to pGex4T-1 but for the expression of MBP-E4 proteins. The EcoR I and a Sal I recognition sites in the multiple cloning region were used to clone EcoR I/Sal I digested E1/E4 PCR amplified products (Figure 4.1). Bacteria transformants were selected for ampicillin resistance (Ap'), and transcription of the MBP fusion gene sequences were triggered from the IPTG inducible 'lac' promoter (positions 1046-1433).
APPENDIX 3 – SDS-PAGE

<table>
<thead>
<tr>
<th></th>
<th>Running gel (10mls)</th>
<th>Stacking gel (2mls)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8%</td>
<td>10%</td>
</tr>
<tr>
<td>Deionised water</td>
<td>4.6</td>
<td>4.0</td>
</tr>
<tr>
<td>30% Acrylamide solution</td>
<td>2.7</td>
<td>3.3</td>
</tr>
<tr>
<td>1.5M Tris (pH 8.8)</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>1M Tris (pH 6.8)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10% SDS</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>10% APS</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>TEMED</td>
<td>0.006</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Appendix 3. Components of SDS-polyacrylamide gel solutions.
REFERENCES


References


279


References


Neutralise Canine Oral Papillomavirus. *16th International Papillomavirus Conference, University of Siena, Italy.* p322.


145. Han R., Cladel N. M., Reed C. A., Peng X., and Christensen N. D. (1999b) Protection of Rabbits from Viral Challenge by Gene Gun-Based Intracutaneous Vaccination with a
Combination of Cottontail Rabbit Papillomavirus E1, E2, E6 and E7 Genes. *Journal of Virology.* 73:7039-43.


References


References

204. Lechner M. S., Mack D. H., Finicle A. B., Crook T., Vou


References


References


References


References


297


Life Cycle Heterogeneity in Animal Models of Human Papillomavirus-Associated Disease

Woei Ling Peh,1 Kate Middleton,1 Neil Christensen,2 Philip Nicholls,3 Kiyofumi Egawa,4 Karl Sotlar,5 Janet Brandsma,4 Alan Percival,7 Jon Lewis,8 Wen Jun Liu,9 and John Doorbar1*


ABSTRACT

Animal papillomaviruses are widely used as models to study papillomavirus infection in humans despite differences in genome organization and tissue tropism. Here, we have investigated the extent to which animal models of papillomavirus infection resemble human disease by comparing the life cycles of 10 different papillomavirus types. Three phases in the life cycles of all viruses were apparent using antibodies that distinguish between early and late events, the onset of viral genome amplification, and the expression of capsid proteins. The initiation of these phases follows a highly ordered pattern that appears important for the production of virus particles. The viruses examined included canine oral papillomavirus, rabbit oral papillomavirus (ROPV), cottontail rabbit papillomavirus (CRPV), bovine papillomavirus type 1, and human papillomavirus types 1, 2, 11, and 16. Each papillomavirus type showed a distinctive gene expression pattern that could be explained in part by differences in tissue tropism, transmission route, and persistence. As the timing of life cycle events affects the accessibility of viral antigens to the immune system, the model ideal model system should resemble human mucosal infection if vaccine design is to be effective. Of the model systems examined here, only ROPV had a tissue tropism and a life cycle organization that resembled those of the human mucosal types. ROPV appears most appropriate for studies of the life cycles of mucosal papillomavirus types and for the development of prophylactic vaccines. The persistence of abortive infections caused by CRPV offers advantages for the development of therapeutic vaccines.

Received 25 March 2002/Accepted 3 July 2002

Human papillomaviruses (HPVs) are associated with epithelial lesions that can progress to high-grade neoplasia and cancer (22, 103). Cancer of the cervix is caused by high-risk papillomavirus types, such as HPV type 16 (HPV16), and is the second-most-common female cancer worldwide. Low-risk human papillomavirus types, such as HPV11, cause genital warts (88). Genital warts afflict ~1 in 200 young adults and in many countries are the most common sexually transmitted disease (51). They often recur and are difficult to treat and often recur.

The prevalence of HPV infections, and the severity to which some lesions can progress, has led to the development of model systems in which to study virus infection. Papillomaviruses infect epithelial tissue, and their life cycle is regulated as the infected epithelial cell differentiates. This leads to a highly ordered pattern that appears important for the production of virus particles. The viruses examined included canine oral papillomavirus, rabbit oral papillomavirus (ROPV), cottontail rabbit papillomavirus (CRPV), bovine papillomavirus type 1, and human papillomavirus types 1, 2, 11, and 16. Each papillomavirus type showed a distinctive gene expression pattern that could be explained in part by differences in tissue tropism, transmission route, and persistence. As the timing of life cycle events affects the accessibility of viral antigens to the immune system, the model ideal model system should resemble human mucosal infection if vaccine design is to be effective. Of the model systems examined here, only ROPV had a tissue tropism and a life cycle organization that resembled those of the human mucosal types. ROPV appears most appropriate for studies of the life cycles of mucosal papillomavirus types and for the development of prophylactic vaccines. The persistence of abortive infections caused by CRPV offers advantages for the development of therapeutic vaccines.
cycle to compare lesions caused by the main animal papillo-
avirus model systems to compare lesions caused by the main animal papillo-

cycle to compare lesions caused by the main animal papillo-

such as those for PCNA and cyclin A. Such proteins are sur-

layers of infected tissue, expression of E6 and E7 (and maybe

among the viruses. Late events begin in the basal layer in

amplification and marks the initiation of late events (42). Only

results of RNA-mapping experiments (IIPV III, CRPV, and bovine papillomavi-

sham, Little Chalfont, United Kingdom) or pMale (New England Biolabs) using

(AOG ATC CGA ATT CGC AAA CGA TAA AGA GAT CGC CCA GAC

GGA) and the reverse primer was BPVE14P2 (GCA ACC CGG GTC GAC

COE14P2 (GCA ACC CGG GTC GAC CTA CTC GTT CTT CTT CTG TG). For CRPV, the primer

CR51E1 (AGG ATC CGA ATT CGC GGC TAG AAA AGT GCC GCC

CGA ATT CAC CGA CAG AGCT CCC CAC TAC GGA CTT CTG GG) and

II65E14P1 (AGG ATC CGA ATT CGC TGA AGC TCA ACA AAC

GTC GAC CTA CTA AAA AAC CGA CTT CTT CCT GCT TG). For HPV11, the forward primer

I165E14P1 (AGG ATC CGA ATT CGC TGA AGC TCA ACA AAC

GTC GAC CTA CTA AAA AAC CGA CTT CTT CCT GCT TG). For HPV11, the forward primer

H65E14P2 (GCA ACC CGG GTC GAC CTA CTC GTT CTT CTT CTG TG). For CRPV, the primer

H51E14P2 (GCA ACC CGG GTC GAC CTA CTC GTT CTT CTT CTG TG). For CRPV, the primer

FIG. 1. Heterogeneity in the timing of events during the life cycles of human and animal papillomaviruses. (A) Tissue sections from a mucosal lesion caused by COPV (top row) and a cutaneous lesion caused by HPV2 (middle row) were double stained using antibodies to E4 (green) and PCNA (red) before being counterstained with DAPI (blue) to visualize cell nuclei. Surrogate markers of E7 expression (PCNA) were confirmed to sporadic cells in the lowest parabasal layers. E4 (green) was first detected in these cells as they migrated through the lower layers of the epithelium. E4 was expressed in the basal layer in lesions caused by COPV and in a few cells above the basal layer in lesions caused by HPV2. Surrogate markers of E7 were lost as E4 appeared in the merged images (m) are shown at the left. A section through a cutaneous lesion caused by CRPV is shown in the bottom row after immunostaining to detect PCNA (red) and in situ hybridization to detect viral genome amplification (blue). The nuclei were counterstained using Sytox green. Surrogate markers of E7 expression (PCNA) did not persist to the epithelial surface and were lost soon after the onset of viral genome amplification (blue). Genome amplification began in the mid-spinous layers in lesions caused by CRPV. The broken lines indicate the positions of the basal layers. The images were taken using a 20× objective.

(B) Tissue sections from mucosal lesions caused by ROPV (top row), HPV11 (middle row), and HPV16 (bottom row) were double stained with antibodies to E4 (green) and PCNA (red) before being counterstained with DAPI (blue) to visualize cell nuclei. The merged images are shown at the left. In such lesions, surrogate markers of E7 expression (PCNA; red) usually persisted into the upper epithelial layers but were lost following the expression of E4. E4 was rarely detected in the lower epithelial layers in lesions caused by HPV11 and -16. Cells expressing both E4 and PCNA are indicated by arrows in the HPV11 and HPV16 images. The broken lines indicate the positions of the basal layers.
lomas (four were examined) were obtained from diagnostic material submitted to Philip Nicholls. Naturally occurring BPV-induced fibropapillomas from cutaneous sites on the head and neck (four total) were provided by Wen Jun Liu. Human warts caused by HPV1 (22 were examined for E1/E4 expression), HPV6 (3 were examined), and HPV65 (6 were examined) were provided by Kiyofumi Egawa or were obtained from previously described sources (38, 42, 43). Cutaneous warts caused by HPV2 (25 were examined for E1/E4 expression) or healed in a domestic pressure cooker in the same time (at 650 W) in citrate buffer (pH 16.0 for all E4 detection except CRPV, which was carried out according to established protocols (9, 28).

Paraffin-embedded sections were dewaxed in xylene (one time for 10 min followed by one time for 5 min) and rehydrated by passage through graded alcohols (twice for 3 min each time in 100% ethanol and then for 2 min each in 80, 50, and 30% ethanol) before being placed in PBS (twice for 5 min each time). For the detection of E1/E4 in lesions caused by BPV, CRPV, ROPV, HPV11, and HPV16, formalin-fixed tissue sections were either microwaved three times for 5 min each time (at 650 W) in citrate buffer (pH 6.0) or heated in a domestic pressure cooker in the same buffer for 3 to 5 min. The citrate buffer was prepared by adding 9 mL of 0.1 M citric acid and 45 mL of 0.1 M sodium citrate to 500 mL of water. The sections were subsequently blocked for 1 h at room temperature using 10% fetal bovine serum or 10% goat serum in PBS before being incubated with the monoclonal or polyclonal antibodies. Epitope exposure was not necessary in order to visualize E1/E4 in lesions caused by COPV, HPV11, and HPV2. Polyclonal antibodies to the E4 protein of BPV were used at dilutions of 1:10 (anti-GST-E1/E4), while those to the E4 protein of COPV and HPV11 were used at a dilution of 1:500. Polyclonal antibodies to the E4 proteins of CRPV and HPV11 were used at 1:100 dilution (anti-peptide and anti-GST-E1/E4), while those to the E4 protein of ROPV were used at a dilution of 1:400 (anti-GST-E1/E4). Preimmune serum was used as a control.

Double staining was carried out following the incubation of tissue sections with two different antibodies. In lesions caused by BPV, COPV, and HPV11, E4 was detected using E4-specific rabbit polyclonal serum and LI was detected using the mouse monoclonal antibody CAMVIR-1 (1:100 dilution; Pharmingen). Rabbit polyclonal antibodies raised against BPV particles (1:100 dilution; DAKO) were used to detect LI in lesions caused by ROPV, while the LI protein of CRPV was detected using a monoclonal antibody to CRPV (Neomarkers). The sections were labeled with either biotin (BioProbe Random Primed DNA Labeling system; Enzo) or digoxigenin (DIG DNA Labeling and Detection kit; Roche Molecular Biochemicals) according to the manufacturer’s instructions. Full-length linearized papillomavirus DNA was used as a template for random primer labeling. In situ hybridization reactions were performed on paraffin-embedded formalin-fixed tissue sections after dewaxing and rehydration (as described above). The sections were then digested for 15 min at 37°C using proteinase K (50 μg/ml) before being washed in PBS (twice for 2 min each time) and air dried at room temperature. The labeled DNA probes were diluted 1:25 in hybridization buffer (10% deionized formaldehyde, 1X Denhardt’s, 5% dextran sulfate, 200 μg of salmon sperm DNA/ml, 4X SSC [1X SSC is 0.15 M NaCl plus 0.015 M sodium citrate]) before being added to the sections. The sections were then covered with glass coverslips, and sealed with cow gum (Cow Proceedings Ltd.), and heated at 90°C for 5 min. The slides were immediately quenched in ice for 5 min and were then incubated overnight at 37°C in a humid chamber. After removal of the cow gum, the coverslips were shaken off in 0.5X SSC and the slides were washed in stringent wash buffers (20% formamide, 2X SSC, 0.05% Tween 20 [twice for 5 min each time] and 2X SSC [twice for 5 min each time]) at 42°C. DIG-labeled DNA was detected using a Tyramide Signal Amplification-Direct system (blueFISH or redFISH) (NEN Life Science Products) according to the manufacturer’s instructions. Endogenous peroxidase activity was removed by incubation in 3% hydrogen peroxide at room temperature. In situ hybridization of each virus type (green), and the nuclei were counterstained using DAPI (blue). The images were taken using a 40X objective.
ing E6, E7, and E5 in human and animal tissues. Antibodies to all three proteins gave similar patterns of staining, but that to PCNA (PC10; Neomarkers) showed good cross-reactivity between species and could detect the protein in bovine, canine, and rabbit tissue. PCNA was present in the basal layers of infected and uninfected epithelium but extended into the upper epithelial layers in papilloma tissue (Fig. 1). Early gene activity, as revealed by the presence of surrogate markers, never persisted to the epithelial surface in areas of productive viral infection but showed a pattern that was characteristic of each infecting papillomavirus type (Fig. 1). Oral lesions caused by COPV typically expressed PCNA in basal cells and in sporadic cells immediately above them—a pattern of expression similar to that seen in human cutaneous lesions (verrucas) caused by HPV1 (data not shown and reference 42) and HPV2 (Fig. 1A). Oral and penile lesions caused by ROPV (17) usually showed more extensive early gene activity, with PCNA being detected throughout the epithelium (an oral papilloma is shown in Fig. 1B). This pattern is similar to that seen in human mucosal lesions caused by HPV11 and HPV16 (an LSIL is shown in Fig. 1B). Lesions caused by HPV16 do not always support the productive stages of the virus life cycle, however, and under these circumstances, markers of early gene expression can be found in cells at the epithelial surface (data not shown and K. Middleton, L. Morris, W. Peh, A. El-Sherif, K. Sotlar, D. Jenkins, R. Seth, H. Griffin, H. Hibma, R. Laskey, N. Coleman, and J. Doorbar, 19th Int. Papillomavirus Conf., abstract P-21, 2001). This is a characteristic of high-grade squamous intraepithelial lesions (HSIL) caused by HPV16 and was not seen during productive infection.

CRPV and BPV infect cutaneous epithelium, making comparison with human mucosal infections more difficult. BPV-induced papillomas from cattle and CRPV-induced papillomas from cottontail (natural host) and New Zealand White (artificial host) rabbits express markers of viral oncogene expression above the basal layer (a cottontail rabbit papilloma is shown in Fig. 1A). The distributions of PCNA were broadly similar in papillomas induced in both rabbit hosts, even though New Zealand White rabbits do not support the full CRPV life cycle. In Fig. 1, DNA in situ hybridization was used to identify cells supporting late events during the life cycle of CRPV, as the CRPV E4 antibodies gave relatively high levels of background staining (Fig. 2B). The distribution of surrogate markers of early gene expression in CRPV-induced papillomas in their natural host resembled that seen in LSIL caused by HPV11 and -16 and in oral papillomas caused by ROPV (Fig. 1). The expression of such markers without the subsequent onset of late events was seen only in lesions produced in New Zealand White rabbits (artificial host) and HSIL caused by HPV16 (data not shown).

Heterogeneity in the timing of late-stage onset in lesions caused by animal and human papillomaviruses. The late stage of the virus life cycle is triggered only as the infected cell migrates towards the epithelial surface. Although E6 and E7 are difficult to detect by immunostaining, E4 is abundantly expressed and can be detected during epithelial cell differentiation by using type-specific antibodies. The expression of E4 has previously been shown to coincide precisely with the onset of vegetative viral genome amplification in lesions caused by HPV1 and HPV16 (42), and in the first instance we sought to establish that this was a common event in the life cycles of all papillomavirus types. To establish this, antibodies were generated to the E1E4 major late proteins of ROPV, CRPV, COPV, and BPV after expression as fusion proteins in bacteria. Antibodies to the E1E4 proteins of HPV2, HPV11, HPV63, and HPV65 were also generated for comparison. In all cases, high-affinity antibodies were obtained which allowed detection of the appropriate E4 protein in biopsy material. With the exception of antibodies to the BPV1 and HPV11 E4 proteins, which cross-reacted with the E4 proteins of BPV2 and HPV6, respectively, the antibodies did not cross-react between virus types. Antibodies to the E1E4 proteins of HPV1 and HPV16 have been described previously (39, 41, 42). No staining was apparent using preimmune serum.

In all the human and animal papillomaviruses examined, genome amplification coincided with the presence of detectable levels of E4, although as with early events, the life cycles of the different viruses showed variation. The most unexpected observation was that viral genome amplification and the expression of E4 begins in cells of the basal layer in oral lesions caused by COPV (Fig. 2A). Most basal cells express PCNA (Fig. 1A), but in lesions caused by COPV, a subset of these express E4 (Fig. 1A) and support viral genome amplification (Fig. 2A) (73). Viral genome amplification in the basal layer was not observed in lesions caused by other papillomavirus types, although HPV1 and -63 (to which COPV is closely related [24, 34]) and HPV65 do initiate genome amplification in cells of the lowest parabasal layers (Fig. 2A) (42). These viruses cause verrucas rather than oral papillomas. By contrast, ROPV, HPV11, and HPV16 (LSIL) showed similarity in the timing of their late events (Fig. 2B). In both cases, genome amplification and E4 expression coincided in the mid-spinous layers, with E4 being detectable to the epithelial surface (Fig. 2B). A correlation between the detection of E4 and the onset of genome amplification was also observed in lesions caused by BPV1 (data not shown) and CRPV (Fig. 2B). The timing of late events in CRPV-induced lesions generated in cottontail rabbits resembled that seen in productive human lesions caused by HPV11 and HPV16 (LSIL) and was distinct from the "sporadic" pattern seen in human cutaneous papillomas caused by related viruses, such as HPV63 (24) (Fig. 2).

Although the association between the appearance of detectable levels of E4 and the onset of viral genome amplification was a common feature of all the lesions examined here, high levels of viral DNA were occasionally seen in the absence of E4 in lesions caused by COPV, as well as in lesions caused by CRPV (Fig. 2B and 3). In CRPV-induced lesions (Fig. 2B), such cells appear to arise from E4-expressing cells (present in the spinous layer) that are supporting viral genome amplification, and we assume that E4 does not persist to the epithelial surface or that the antibody is not able to detect it in the upper epithelial layers. Although this may also be true of the cells that contain amplified viral genomes (but not E4) in lesions caused by COPV, the sporadic distribution of E4-containing cells in these lesions makes this less evident. We do not know for certain whether the in situ signals seen in the E4-negative cells in COPV-induced lesions represent true genome amplification or a partial relaxation of the copy number control that operates in the basal layer. Cells containing amplified viral
FIG. 3. E4 expression and the onset of genome amplification are not necessarily coincident. (A and B) Tissue sections taken from two different COPV-infected oral lesions. Tissue sections were double stained for the presence of viral genome amplification (DNA) and for E4 expression (E4). Viral DNA was occasionally detected in the absence of E4 (cells indicated by arrows). The nuclei were counterstained using Sytox green (nuclei). (B) E4 expression in cells (indicated by arrowheads) that are not supporting viral genome amplification. A cell (indicated by an arrow) supporting viral genome amplification but lacking E4 is shown in the same field of view. The images were taken using a 40X objective.

...genomes in the absence of E4 were not seen in lesions caused by other papillomavirus types. The onset of late events is closely linked to the completion of early events during the papillomavirus life cycle. Although our data show that the timing of the late-event activation is divergent among papillomavirus types, a high degree of consistency was apparent when lesions caused by the same virus type were compared. This was most evident in papillomas induced by the evolutionarily related viruses (24) COPV, HPV1, and HPV63, which trigger late gene expression in the basal and parabasal cell layers (Fig. 1 and 4). Over 60% of the cells in the upper epithelial layers of such lesions express E4, with such cells usually appearing in clusters—a distribution that is also seen in lesions caused by BPV1 (Fig. 4). No variation in this pattern was apparent in >8 COPV-induced oral papillomas and 25 HPV1-HPV63-induced verruca lesions examined. Lesions caused by CRPV in its natural host, ROPV, HPV11, and HPV16 (LSIL) also showed consistency in their patterns of late gene expression (which began in the intermediate layers), but these were distinct from the patterns seen in oral lesions caused by COPV and in cutaneous lesions caused by HPV1, HPV2, or HPV63 (Fig. 4). The expression pattern characteristic of each papillomavirus type is shown in Fig. 4. Despite the heterogeneity among viruses, the disappearance of early markers always occurred after the onset of late events, suggesting that the two stages in the virus life cycle are closely coordinated (Fig. 1). Cells expressing both E4 and viral oncogenes, such as E7 (as determined by the presence of surrogate markers), were found in a narrow region of overlap, which could be as little as one cell layer thick in LSIL caused by HPV16 (Fig. 1B) to over five cell layers thick in COPV-induced oral papillomas or cutaneous warts caused by HPV2 (Fig. 1A). The overlap between surrogate markers of viral early gene activity (e.g., PCNA and cyclin A) and late markers (e.g., E4 and genome amplification) shown in Fig. 1 was a common feature of all the productive papillomavirus infections examined here. Only during abortive infection, such as in HPV16-associated HSIL, or in lesions produced by CRPV in New Zealand White rabbits (artificial host), is this pattern of expression disrupted.

Intracellular distribution of E4 in lesions caused by animal papillomaviruses. The E4 proteins of many papillomaviruses form intracellular structures that are characteristic of the infecting virus type. While this is particularly evident in lesions caused by HPV1 and HPV4 (38), it has been suggested that most cutaneous viruses form some sort of cytoplasmic inclusion (32, 46, 47, 70). By contrast, the E4 proteins of human mucosal viruses do not produce inclusions (19, 21, 40, 42, 81). HPV16 E4 associates with the cellular intermediate filament network, whereas HPV11 E4 localizes to the cornified envelope of the cell. E4 is one of the few viral proteins that can easily be detected during productive infection, and the generation of antibodies allows its intracellular distribution to be examined. Although the role of E4 is uncertain, similarities in intracellular localization may suggest functional conservation. Of the viruses examined here, only ROPV E4 had an in vivo intracellular distribution that broadly resembled that of the human mucosal viruses. The E4 proteins of HPV16 and HPV11 are predominantly cytoplasmic, and this was also true of ROPV E4 (Fig. 5A). Unlike the E4 proteins of HPV16 and HPV11, however, ROPV E4 assembled into cytoplasmic inclusion granules similar to those seen in warts caused by HPV1 (82, 83), and these increased in size as the cells approached the epithelial surface (Fig. 5A). Granule formation is an inherent property of the E4 proteins of HPV1 and ROPV and also...
Timing of late events is highly conserved among lesions caused by the same papillomavirus type. Similarities in the patterns of E4 expression and genome amplification confirmed that E4 was an effective marker of the late stages of the papillomavirus life cycle. Tissue biopsies infected by different papillomaviruses were sectioned and stained using E4-specific antibodies (green) before being counterstained with DAPI (blue). The staining patterns illustrated were typical of those seen following the examination of five or more independent biopsy specimens infected by each virus type (except for HPV63, for which three lesions were examined [see Materials and Methods]). The timing of late-stage onset was a characteristic feature of each infecting virus type. All images were taken using a 4x objective except for the HPV16 and CRPV images, which were taken using a 10x objective lens. The basal layers are indicated by the broken lines.

Although E4 is predominantly cytoplasmic in all the human papillomavirus infections examined to date (17, 38, 42), in lesions caused by COPV and CRPV the protein was distributed throughout the cell (Fig. 5). Cells expressing COPV E4 contained cytoplasmic structures that were similar in appearance to the inclusions produced by ROPV (Fig. 5A and C). COPV E4 protein was also associated with the nuclei (Fig. 5D) and with cell membranes (Fig. 5C). Although CRPV E4 was not associated with any obvious structures, lesions caused by CRPV and COPV contained abundant keratohyalin granules. In CRPV-induced warts, these were found in the cells expressing E4 (Fig. 5B). Mucosal epithelium, such as that infected by COPV and ROPV, does not normally contain abundant keratohyalin granules (Fig. 5C, inset), making their presence in COPV-infected tissue surprising. Keratohyalin granules were not, however, present in cells expressing COPV E4. This is similar to the situation seen in lesions caused by HPV1, where E4-expressing cells lack filaggrin and do not produce visible keratohyalin (42). Although ROPV and COPV infect similar epithelial sites, keratohyalin granules were never detected in lesions caused by ROPV and were only occasionally present in human mucosal infections caused by HPV11 and -16 (data not shown).

Expression of E4 always precedes expression of the L1 capsid protein. The major virus coat protein was expressed in the nuclei of terminally differentiating cells near the surface of the epidermis. In all the lesions examined here, L1 expression followed that of E4, and a gap was apparent between the first appearance of E4 and the first appearance of L1 (Fig. 6). Although L1 expression was first detected in E4-containing cells, L1 expression was not always supported; and cells expressing only E4 could occasionally be found at the epithelial surface. This was sometimes apparent in productive infections (e.g., HPV11 [Fig. 6]) but was seen more regularly in HSIL caused by HPV16. Such lesions resemble those produced in New Zealand White rabbits by CRPV and do not support the full life cycle of the virus (data not shown and Middleton et al., 19th Int. Papillomavirus Conf.). While the coordinated pattern

![Image of papillomavirus lesions](image-url)
of E4 and L1 expression was conserved in all productive lesions examined, the gap between the first appearance of E4 and the onset of L1 expression varied considerably. Viruses that initiate their late events in the lower epithelial layers, such as COPV, HPV1, and HPV63 (which are evolutionarily related [24]), typically showed a greater number of L1-expressing cells than those that initiate late gene expression in the upper epithelial layers (compare COPV and HPV11 in Fig. 6). LSIL caused by HPV16 and cutaneous papillomas caused by CRPV (in the cottontail rabbit host [Fig. 6]) had intervals as small as one or two cell layers between the site of E4 expression and that of L1. Cutaneous lesions caused by HPV1 and HPV2 (Fig. 6) usually initiated L1 expression four or five cell layers after the first appearance of E4. By contrast, L1 expression in HPV63- or COPV-induced lesions may be separated from the onset of E4 expression by as many as 20 cell layers (42) (Fig. 6). Of the lesions examined here, those caused by HPV11 and ROPV showed the greatest similarity in the expression patterns of their E4 and L1 proteins, although E4 expression in ROPV lesions usually began earlier than in lesions caused by HPV11 (Fig. 6).

Viral gene expression in xenografts resembles that seen during infection of the natural host. Several papillomaviruses, including ROPV (28), CRPV (30), and HPV11 (36, 63), can be propagated in epithelial tissue implanted under the renal capsule. This approach has been used to generate stocks of animal (28) and human papillomaviruses and to study the HPV life cycle in vivo (18, 20). Renal-capsole xenografts infected with CRPV or ROPV were compared to experimental infections in their natural hosts with regard to the timing of initiation of late events described above. The lesions produced in both systems were broadly similar, and the timing of E4 expression, genome amplification, and virus synthesis (described here) was preserved (Fig. 7). Detectable E4 expression began in the intermediate and spinous cell layers and coincided with the onset of viral genome amplification (ROPV is shown in Fig. 7). Although genome amplification was not examined in CRPV-infected xenografts, the expression of CRPV E4 followed a pattern that resembled that seen during natural infection (Fig. 4 and 7). In all xenografts examined, irrespective of the infecting papillomavirus type, the L1 protein was first expressed in only a subset of E4-positive cells in the upper layers of the epidermis (HPV11 is shown in Fig. 7). Xenografts propagated under the kidney capsule, however, did show a lower degree of papillomatosis than lesions produced at the natural site of infection, as reported previously (8). It appears that L1 expression always follows that of E4, in contrast to previous reports, which have indicated that the expression of E4 and L1 is coincident (18, 20). HPV11 xenografts propagated on the skin (in SCID mice) showed similarity to naturally occurring human genital lesions both in their morphology and in their pattern of late gene expression (compare Fig. 1, 2, 4, and 6 and HPV11SCID/E4 and -L1 in Fig. 7).

DISCUSSION

Papillomaviruses are highly host specific, making the use of animal models a valuable strategy for the in vivo analysis of infection. CRPV has been extensively used for the development of vaccines (16, 53, 54, 65, 87, 94) and has provided insight into the roles of viral proteins during papilloma formation (13, 14, 33, 72, 104). Vaccination studies have also been carried out using COPV, which, unlike CRPV, infects oral epithelium (5, 91, 95, 105). Although the mucosal tropism of COPV has led to its use as a model of genital HPV infections, the COPV life cycle differs in several ways from those of HPV11 and HPV16. At the DNA and protein sequence levels, COPV shows homology with HPV types that cause plantar and palmar warts, and this is reflected in similarities among lesions caused by these viruses (24). Of the papillomaviruses examined here, the COPV E4 protein resembles the E4 proteins of HPV4 and HPV65 most closely, while its E6 protein shows the greatest sequence homology with the E6 proteins of HPV1 and -63. The initiation of late events in basal and parabasal cell layers is a characteristic of these related virus types and is an obvious difference between lesions caused by COPV and those caused by CRPV, ROPV, HPV11, and HPV16 (summarized in Fig. 8). Interestingly, COPV-infected cells contained inclusions that were similar in appearance to those seen in lesions caused by HPV1 and ROPV. The E4 proteins of COPV and ROPV contain variants of the C-terminal DLXDW motif that is found in cutaneous human papillomaviruses and which is involved in E4 multimerization (2). This motif has not previously been identified in the E4 proteins of mucosal papillomavirus types (44). If vaccination studies using model systems are to be relevant in the prevention and treatment of human disease, the chosen model should closely mimic infection in humans. COPV and HPV11 have similar tissue tropisms but differ in the timing of their life cycle events (Fig. 8). By contrast, the life cycles of HPV11, HPV16, and ROPV (which infect genital tissue) are organized similarly. The ability of New Zealand White rabbits to support the full ROPV life cycle makes this system useful for the study of late gene function, but the short duration of the experimental ROPV-induced papillomas may limit its use to the development of prophylactic vaccines. Although experimental papillomas caused by CRPV do not support late events, they are persistent and offer some advantages in the development of therapeutic vaccines and antivirals.

It has been suggested that papillomaviruses evade immune detection, at least in part, by expressing their abundant late proteins only in the upper epithelial layers (86, 89, 92). Viral early proteins, such as E6 and E7, are expressed at levels below those necessary to stimulate an effective immune response. While this hypothesis is plausible for human mucosal infections, it does not explain the results obtained here from the analysis of lesions caused by HPV1, HPV65, or COPV, where high levels of E4 are apparent in the basal and parabasal layers (Fig. 4 and 8). Unlike those caused by HPV16 or HPV11, naturally occurring lesions induced by COPV are rarely persistent (74) and usually undergo spontaneous regression within 12 weeks of infection (23). Perhaps significantly, plantar warts caused by HPV1 (myrmecia) are also often short-lived (52). Such lesions usually respond to treatment better than plantar warts caused by HPV2 (mosaic warts) and rarely recur, due to the presence of neutralizing antibodies (64, 80). Little is known of the natural history of verrucas caused by HPV63 and HPV65, but it is possible that wart persistence may be linked to an ability to retard the expression of E4 and capsid proteins until the cell has cleared the basal layer. If this is the case, then infections caused by HPV11 and HPV16 may be more persis-
tent than infections caused by COPV or HPV1 (Fig. 8). The rapid regression of genital lesions caused by HPV6 or -11 is relatively uncommon and usually occurs following a period of wart proliferation, as can often occur during pregnancy (7, 76, 96). HPV16 infections vary greatly in duration (57, 99) and also show variation in the timing of their late gene expression. Persistent infections caused by HPV16 (at cervical epithelial sites) and CRPV (in New Zealand White rabbits) are associated with an increased risk of malignant progression (31, 62, 101). Although there is as yet no firm evidence to link the duration of infection with the extent of virus synthesis, many reports have revealed the importance of the host immune system in restricting papilloma growth and in stimulating regression (reviewed in references 56 and 89). The exposure of viral late proteins to the immune system can prevent papillomavirus spread and can stimulate the early regression of existing warts (reviewed in references 15 and 59). Viruses that are transmitted by intimate physical contact, such as HPV65 [verrucas] and COPV [canine oral papillomas], where transmission by intimate physical contact is unlikely, the chance of initiating a new infection is enhanced. Papillomaviruses that have been examined to date, E4 is expressed from a differentiation-dependent promoter late in infection (37). In human papillomaviruses, this promoter is positioned in front of the E1 helicase, which is known to have a direct role in the replication of viral genomes (reviewed in references 15 and 59). The E1 and E4 genes share the same initiation codon, and the relative abundances of the two messages are controlled at the level of splice site selection. This apparent link between the synthesis of full-length E1 transcripts and the production of the primary E1/E4 message suggests that the two proteins may be required together during the virus life cycle. In fact, E4 appears to be expressed as part of a single transcriptional unit that also includes E1, E2, and E5. E5 reactivates DNA synthesis in quiescent cells (1, 10, 93, 98, 100), while E2 is necessary for the efficient recruitment of E1 to viral origins (49, 67). Interestingly, a role for HPV16 E4 in genome amplification has recently been proposed (C. Day, D. Jackson, K. Raj, P. Master- son, J. Millar, and J. Doorbar, 19th Int. Papillomavirus Conf., abstr. O-174, 2001), and it has been shown that loss of E4 prevents amplification of the CRPV genome in domestic and cottontail rabbits (W. Peh, J. Brandsma, N. Cladel, N. Chris- tensen, and J. Doorbar, 19th Int. Papillomavirus Conf., abstr. O-146, 2001).

E4 expression and the onset of genome amplification were always found to begin in cells expressing surrogate markers of early gene activity, such as PCNA or cyclin A (Fig. 1 and 8). The viral E6, E7, and E5 proteins drive cells into S phase and stimulate the synthesis of the cellular proteins necessary for viral genome amplification. The region where E4-PCNA double-positive cells were found varied greatly in thickness in

FIG. 5. Intracellular distribution of the E4 proteins of animal papillomaviruses. (A) The E4 protein of ROPV (red) is predominantly cytoplasmic and is associated with inclusion granules (center image, g) similar to those seen in cutaneous lesions caused by HPV1. The nuclei were counterstained with DAPI (blue) and are visible in the merged image at the left (ROPV/H+E). The images were taken using a 40× objective. (B) The E4 protein of CRPV (green) is distributed throughout the nucleus and the cytoplasm. The nuclei were counterstained with DAPI (blue) and are visible in the merged image shown on the left (CRPV/H+E). The cytoplasmic structures that did not stain with antibodies to E4 are keratohyalin granules (center image, arrows). They are also shown (K) in the hematoxylin- and eosin-stained image (CRPV/H+E) on the right. (C) The E4 protein of COPV (green) is cytoplasmic and nuclear but was also associated with the nuclear and cellular periphery (upper center image, arrow). The nuclei were counterstained with either propidium iodide (upper left image, red) or DAPI (lower left image, blue). A granular pattern was apparent in the cytoplasm of some cells (granule-like structures [arrows] in lower center image). Keratohyalin granules (KH) are abundant in COPV-induced warts and are shown in the hematoxylin- and eosin-stained image (COPV/H+E; upper right). The surrounding mucosal epithelium is devoid of keratohyalin (inset). Cells expressing COPV E4 had a characteristic morphology that may result from the presence of E4 inclusion granules. The permissive cells (P) that express E4 and the nonpermissive cells (NP) that do not express E4 are shown in the hematoxylin- and eosin-stained image on the lower right. (D) In the lower epithelial layers of experimental warts caused by COPV, nuclear E4 protein (green) was found associated with the nuclei. At the left (COPV/m), the E4 and DAPI (blue) stain is shown as an overlay of the phase-contrast image. The phase-contrast image is shown in the center panel to indicate the presence of the nuclei (NUC). The presence of the permissive "granular" cells (P) and the nuclei is clearly visualized in the immunoperoxidase stain (brown) shown on the right [COPV/E4 (DAB)].
FIG. 6. Expression of capsid proteins follows expression of E4 in lesions caused by different papillomavirus types. Tissue sections of lesions caused by COPV, ROPV, HPV11, CRPV, and HPV2 were double stained using antibodies to the L1 capsid protein (red) and E4 (green) before being counterstained with DAPI (blue). The merged images (m) are shown on the left. Although E4 expression always precedes the expression of L1, the distance between the first appearance of E4 and the first appearance of L1 varied considerably. The positions of the basal layers are indicated by broken lines. The images were taken using a 10× (HPV2, HPV11, and COPV) or 20× (CRPV) objective.
FIG. 7. The timing of late gene expression in xenografts resembles that seen during natural infection. Epithelial tissue infected by CRPV or ROPV was propagated under the kidney capsule of nude mice (images labeled CRPVNUDE and ROPVNUDE) before being stained for E4 and (for ROPV) genome amplification. The nuclei were counterstained with DAPI and are shown on the left. In all cases, the timing of late-stage activation (as determined by E4 expression) was similar to that seen following natural and experimental infection of the natural host (Fig. 4). Genome amplification and E4 expression coincided closely (images labeled ROPVNUDE/DNA and ROPVNUDE/E4). Epithelial tissue infected by HPV11 was propagated under the kidney capsule of nude mice or as a skin graft on SCID mice. L1 expression followed that of E4 in renal capsule xenografts infected with HPV11 (images labeled HPV11NUDE/E4 and HPV11NUDE/L1). This pattern of expression was also seen in skin grafts propagated on SCID mice (HPV11SCID/E4 and HPV11SCID/L1). The morphology of the skin xenografts closely resembled that seen in HPV11-induced genital lesions (Fig. 2 and 4). The images were taken using a 10× (ROPV and HPV11) or 20× objective. The broken lines indicate the positions of the epithelial basal layers.

different lesions. In HPV1- or HPV2-induced warts, double-positive cells were abundant, while in lesions caused by HPV16 they were scarce. The increase in the levels of the viral replication proteins E1 and E2 that result from the activation of the differentiation-dependent promoter are thought to contribute to genome amplification (6, 61, 78, 79). Our observation that markers of E6 and E7 expression are lost soon after the appearance of E4 can be readily explained if the levels of E2 and E4 rise concomitantly. Both E2 and E4 are expressed from the differentiation-dependent promoter late in infection (26, 27, 42), with E2 showing an expression pattern similar to that of E4 (66). High-level expression of E2 leads to down-regulation of the viral early promoter and inhibition of E6 and E7 expression. Although E2 is necessary for the amplification of viral genomes, its accumulation may eventually restrict the duration of vegetative viral DNA replication by down-regulating the expression of E7.

Our comparison of the life cycles of 10 viruses reveals differ-
FIG. 8. The regulation of early and late events in lesions caused by different papillomavirus types. The timing of life cycle events in lesions caused by animal and human papillomaviruses is indicated by the bars. The shaded bars show the presence of amplified viral DNA, while the stippled bars show the extent of E4 expression. The expression pattern of surrogate markers of E7 is shown by the solid bars, whereas the L1 expression pattern is indicated by the hatched bars. The darker region at the bottom of the shaded bars indicates the region where vegetative viral genome amplification is thought to occur. Mucosal epithelial tissue infected by COPV, ROPV, HPV11, and HPV16 is divided into four layers, shown on the left. Cutaneous tissue infected by HPV1, HPV63, BPV1, HPV65, HPV2, and CRPV is divided into five layers, shown on the right. COPV triggers late events in the lowest epithelial layers. Lesions caused by HPV11 and HPV16 (LSIL) usually support late gene expression only in the upper half of the epidermis, whereas late events are not always supported in HSIL. In all instances, the loss of surrogate markers of E7 does not occur until after E4 has accumulated to detectable levels.


