AN INVESTIGATION OF POST-TRANSLATIONAL MODIFICATION OF HUMAN PAPILLOMAVIRUS E4 AND THE ROLE OF MODIFIED E4 PROTEINS DURING THE VIRUS LIFE CYCLE

By

ALICE GEORGIA PUGH

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Cancer Sciences

College of Medical and Dental Sciences

The University of Birmingham

2009

ABSTRACT

Infection with human papillomaviruses (HPV) is the etiological basis for cervical cancer. The HPV E1^E4 protein is proposed to be a major regulator of the HPV life cycle and the multiple and diverse activities associated with E1^E4 suggest that it is a multi-functional protein. This thesis has sought to address the hypothesis that phosphorylation and proteolysis contribute towards the pleiotrophic functions of the E1^E4 protein, and has investigated the functional significance of these E1^E4 post-translational modifications during the HPV life cycle. This study has uncovered the novel finding that the E1^E4 protein of HPV type 18 (HPV18) exists as a phospho-protein within cells and is a substrate for multiple cellular kinases *in vitro*. The phospho-acceptor residue for cyclin-dependent kinases (CDK) 1 and 2 has been identified as threonine 23, whilst serine 58 is phosphorylated by protein kinase A (PKA). Furthermore, a cyclin binding motif ([43]RRL[45]) within the HPV18 E1^E4 protein is required for association with active CDK complexes and this association may influence CDK activity since the activity of CDK2-cyclin A was shown to be reduced in the presence of HPV18 E1^E4.

This thesis has revealed that HPV18 E1^E4 is a target for N-terminal proteolysis, and this post-translational modification occurs during the HPV18 replication cycle. Key elements necessary for proteolysis have been mapped to a conserved leucine-rich sequence (LLXLL) present at the N-terminus of the HPV18 E1^E4 protein. Since E1^E4 expression is required for HPV18 genome amplification, N-terminally truncated E4 species may contribute towards its role in the replication cycle. To examine this hypothesis, mutations that attenuate E1^E4 proteolysis were introduced into HPV18 genomes and transfected into human foreskin keratinocytes (HFK). Mutation of the leucine-rich motif prevented efficient proteolysis of the E1^E4 protein during the HPV life cycle and resulted in reduced viral genome amplification within differentiating HFKs suggesting that efficient E1^E4 proteolysis may be required for this E4 function.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

CHAPTER 3

HPV18 E1^E4 IS PHOSPHORYLATED *IN VITRO* BY CYCLIN-DEPENDENT KINASES AND PROTEIN KINASE A

CHAPTER 4

HPV18 E1^E4 ASSOCIATION WITH CDK ACTIVITY AND THE EFFECT OF

E1^E4 ON CDK-DEPENDENT FUNCTIONS OF HPV PROTEIN

CHAPTER 5

N-TERMINAL PROTEOLYSIS OF HPV18 E1^E4 IS DEPENDENT UPON

SEQUENCES WITHIN AN N-TERMINAL LEUCINE-RICH MOTIF

CHAPTER 6

THE ROLE OF N-TERMINAL PROTEOLYSIS OF THE E1^E4 PROTEIN DURING

THE HPV18 VIRUS LIFE CYCLE

CHAPTER 7

OVERALL DISCUSSION AND FUTURE DIRECTIONS

APPENDICES

PUBLICATIONS AND PRESENTATIONS OF RESEARCH

REFERENCES

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 6

CHAPTER 7

APPENDICES

LIST OF TABLES

LIST OF ABBREVIATIONS

β-ME   Beta-mercaptoethanol

λ-PPase   Lambda phosphatase

AA   Amino acids

APS   Ammonium persulphate

ATP   Adenosine triphosphate

BPV   Bovine papillomavirus

BrdU   5-bromo-2-deoxyuridine

BS   Binding site

BSA   Bovine serum albumin

cAMP   Cyclic adenosine monophosphate

CCE   Cornified cell envelope

cDNA   Complementary DNA

CDK   Cyclin-dependent protein kinase

CIN   Cervical intraepithelial neoplasia

CK   Casein kinase

CO   Codon-optimised

CPE   Cytopathic effect

CR   Conserved region

DABCO   1,4-diazabicyclo[2.2.2]octane

DAPI   4', 6'-diamidino-2-phenylindole

Dlg   Discs-large

DMEM   Dulbecco's modified eagle medium

DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide tri-phosphate
DTT	Dithiothreitol
E6-AP	E6-associated protein
EBV	Epstein-Barr virus
EDTA	Ethylene diamine tetra-acetic acid
EGF	Epidermal growth factor
EGTA	Ethylene glycol tetra-acetic acid
ERK	Extracellular signal-regulated kinase
EtBr	Ethidium Bromide
FADD	Fas-associated death domain
FBS	Foetal bovine serum
G	G-force
GFP	Green fluorescent protein
Gy	Gray
H	Hour
HA	Haemagglutinin
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HFK	Human foreskin keratinocytes
HSIL	High grade squamous intraepithelial lesions
HHV	Human herpes virus
HINGS	Heat-inactivated goat serum

HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HTLV	Human T-cell leukaemia virus
IBV	Infectious bronchitis virus
IgG	Immunoglobulin G
IPTG	Isopropyl-ß-D-thiogalactopyranoside
Kb	Kilobase
kDa	Kilodalton
KGM	Keratinocyte growth medium
KSHV	Kaposi's sarcoma herpes virus
L	Litre
LB	Luria-Bertani
LCR	Long control region
LRR	Localisation regulatory region
LSIL	Low grade squamous intraepithelial lesions
MAb	Monoclonal antibody
mA	Milli-amp
MAPK	Mitogen-activated protein kinase
Min	Minute
MME	Minichromosome maintenance element
ND10	Nuclear domain 10
NES	Nuclear export sequence
NIKS	Normal immortalised keratinocytes
NLS	Nuclear localisation sequence

OC	Open circular
ORF	Open reading frame
Ori	Origin of replication
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDZ	PSD-95/discs large/ZO1
PI	Post infection
PKA	Protein kinase A
PKC	Protein kinase C
PML	Promyelocytic leukemia protein
pRB	Retinoblastoma protein
rAd	Recombinant adenovirus
Rb	Retinoblastoma
rRNA	Ribosomal RNA
RNA	Ribonucleic acid
RNase	Ribonuclease
rSV40	Recombinant SV40
SC	Supercoiled
SCC	Squamous cell carcinoma
SDS	Sodium Dodecyl Sulphate
Sec	Second
SRPK	Serine-arginine specific protein kinase
SSC	Saline sodium citrate

SV40	Simian virus 40
T₃	3,3',5-Triiodo-L-thyronine
TAg	Large T antigen
TE	Tris-EDTA
TEMED	N,N,N',N' tetramethylethylene-diamine
TBE	Tris/Borate/EDTA
TBP	TATA-binding protein
TBS-T	Tris-buffered saline – Tween20
TNF	Tumour necrosis factor
TNS	Trypsin-neutralising solution
TRADD	TNFR1-associated death domain
TRAIL	TNF-related apoptosis-inducing ligand
URR	Upstream regulatory region
UTR	Untranslated region
UV	Ultraviolet
V	Volt
VLP	Virus-like particle
vol/vol	Volume per volume
wt/vol	Weight per volume

CHAPTER 1

INTRODUCTION

1.1 Infectious agents and cancer

Infectious agents have a significant role to play in the global burden of cancer, and it is estimated that viruses, bacteria and parasites contribute towards just over 20% of cancers worldwide (zur Hausen H., 2009). Epstein Barr virus (EBV) was the first virus identified to be associated with a human tumour and has been etiologically linked with Burkitt's lymphoma, nasopharyngeal carcinoma, T cell lymphomas and Hodgkin's disease (Rezk & Weiss, 2007). A number of other viruses have subsequently been associated with the development of human cancers including human papillomavirus (HPV) with cervical cancer, anogenital cancers and head and neck cancers; hepatitis B virus (HBV) and hepatitis C virus (HCV) with hepatocellular carcinoma; human T-cell leukaemia virus (HTLV-I) with adult T-cell leukemia, and Kaposi's sarcoma herpes virus (KSHV), also referred to as human herpes virus type 8 (HHV-8), associated with Kaposi's sarcoma (Carrillo-Infante et al., 2007). Furthermore, infections with parasites and bacteria are also known to promote development of a subset of cancers and infection with *Helicobacter pylori* is a major cause of gastric cancers (Correa & Houghton, 2007).

For many tumour viruses, cancer is an unintended consequence of a productive viral infection, and infection with the virus is merely a prerequisite for development of cancer, with further factors required for progression to carcinogenesis (Liao, 2006). Whilst many viruses act directly to promote carcinogenesis, for other viruses such as the human immunodeficiency virus (HIV) the mechanism of action is indirect. By inducing immunosuppression, HIV acts to permit secondary infections with other viruses including HHV-8 (Feller et al., 2007), HPV (Palefsky, 2009) and EBV (Tran et al., 2008), which can subsequently result in associated cancers. HCV and HBV have also been shown to cause cancer indirectly through causing chronic inflammation of the liver (Chemin & Zoulim, 2009).

Tumour viruses can be classified according to their genomic composition into DNA and RNA tumour viruses (Carrillo-Infante et al., 2007). For RNA tumour viruses such as HTLV-1 and HIV, reverse transcription converts the viral RNA genome into DNA and is often followed by integration into cellular DNA, although for HCV the RNA genome is maintained episomally (Carrillo-Infante et al., 2007). Similarly, genomes of DNA tumour viruses can be maintained as episomes or can become integrated into the host genome to exert carcinogenic effects. DNA tumour viruses are a diverse group of viruses and whilst the mechanisms by which they promote cellular transformation vary, viral oncogenes often target common cellular pathways promoting cellular growth and preventing apoptosis to enable replication of the virus within the host cell (Dayaram & Marriott, 2008). As a result of interference of cellular pathways, viral infections can result in abnormal cellular growth, genetic mutations and malignancy.

In addition to improving our knowledge of the mechanisms of carcinogenesis, studies of tumour viruses have also contributed greatly to our understanding of cellular biology (Carrillo-Infante et al., 2007). Studies of tumour viruses have led to the discovery of numerous cellular oncoproteins and tumour suppressors and widened our fundamental understanding of cellular pathways. The revelation that a subset of human cancers are caused by viruses presents a means by which the burden of cancer may be reduced and viruses are now attractive targets on which to base preventative therapeutics. It is proposed that future research is likely to uncover further links between human cancers and infectious agents and this will further increase our understanding of carcinogenesis (zur Hausen H., 2009).

1.2   Human papillomaviruses and cancer

Cervical cancer is the second most common cancer for women worldwide, accounting for 290,000 deaths per annum (World Health Organization, 2002). Cancers of the uterine cervix

are classified into two major histological forms, dependent upon the cell type from which the cancer originated, with squamous cell carcinoma (SSC) constituting around 80-85% of these malignancies, whilst adenocarcinomas, derived from adenomatous glandular cells are more infrequent, occurring in only 15-20% of cases (Green et al., 2003). Over 99% of squamous cell cervical carcinomas contain viral DNA from human papillomaviruses (HPV) and infection with this virus has been affirmed as the etiological basis for this cancer (Walboomers *et al.,* 1999). The prevalence of HPV DNA in adenocarcinomas of the cervix is also very high at 91% (Pirog et al., 2000) and HPV infection has also been shown to be a causative agent of other anogenital cancers. It is estimated that 40% of penile, vulva, and vaginal cancers are attributed to HPV infection whilst 90% of anal cancers are predicted to be a result of infections with this virus (Parkin & Bray, 2006). Infection with HPV has also been shown to be an aetiological factor for oral and pharyngeal cancers (Gillison & Shah, 2003), with 3 and 12% of cases respectively, attributable to HPV infection (Parkin & Bray, 2006).

Human papillomaviruses are a heterogeneous group of DNA viruses with a tropism for epithelial tissue, infecting cutaneous or mucosal squamous epithelium, dependent upon the viral type. Over 100 different HPV types have been identified and these can be categorised into different genera based upon sequence homology (de Villiers et al., 2004). Alpha papillomaviruses constitute the mucosal or genital species and can be further classified according to their oncogenic potential, into low risk and high risk subtypes dependent upon their propensity to cause cancer (Munoz et al., 2003). Low risk HPV types, of which HPV6 and 11 predominate, are associated with the formation of benign hyperproliferative lesions including condyloma acuminata, which clinically manifests as genital warts and rarely progress to malignancy (Longworth & Laimins, 2004). Fifteen HPV types have been classified as high risk, and infections with these subtypes of HPV, have a higher propensity to

progress to cervical carcinoma, with HPV16 and 18 being the most common types, and are together estimated to account for 70% of all cervical cancers (Clifford et al., 2006). Whilst HPV16 is the predominant HPV type of squamous cell cervical carcinomas, HPV18 predominates within adenocarcinomas (Bulk et al., 2006).

HPV is a common sexually transmitted virus and the risk of exposure to the virus within the sexually active population is high, with an 80% lifetime risk of contracting a genital HPV infection (Bekkers et al., 2004). However, very few infections result in progression to cervical cancer, and within 18 months 80% of infections are cleared (Baseman & Koutsky, 2005). Whilst the majority of infections are transient, persistent infection with a high risk HPV type is the greatest risk factor for cervical cancer (Ho et al., 1995). HPV infection is necessary but not sufficient for progression to malignancy, and cervical carcinogenesis is a multi-step process with a long period of latency often observed between the initial HPV infection and progression to cervical cancer. A number of cofactors for cervical cancer have been identified including: smoking, parity, and use of oral contraceptives, which can increase the propensity of cancer to arise (Castellsague & Munoz, 2003). Whilst the exact mechanisms are unknown, these cofactors are proposed to influence progression to carcinogenesis through suppression of the immune system, resulting in increased viral persistence, or through elevating levels of cellular DNA damage (Castellsague & Munoz, 2003).

1.3   HPV screening and prophylactic vaccination

Disparities exists in mortality rates of cervical cancer worldwide, and a significant burden is imparted upon the developing world, with 80% of deaths occurring within low income countries (World Health Organization, 2002). Within developed countries the incidence and

mortality rates of cervical squamous cell carcinoma have significantly reduced by as much as 80% over the past four decades, due largely to the widespread implementation of effective screening programmes which enable detection and early treatment of pre-cancerous changes within the cervix (Parkin & Bray, 2006). However whilst the incidence of squamous cell cervical carcinomas has been falling, an increased rate of adenocarcinomas has been reported, believed to be partly due to the reduced capability of screening techniques for detection of adenocarcinoma (Bray et al., 2005).

Whilst development of low grade lesions is frequent within young women, development of cervical cancer is more prevalent within older women, with an average age of around 45 years for diagnosis of cervical cancer (Schiffman et al., 2007). Often progression from the initial infection to malignancy can take many years and even decades. Therefore the use of screening programmes can be effective at detection of pre-cancerous lesions (Bekkers et al., 2004). The Papanicolaou test or Pap smear is used routinely to screen women for the detection of abnormal changes to the cervical epithelium resulting from a high risk HPV infection, which may have the propensity to progress to invasive cervical cancer (Bekkers et al., 2004). Cells from the surface of the cervix are examined for precancerous abnormalities, termed dysplasias, and the cervical intraepithelial neoplasia (CIN) system or the Bethesda system, which has largely replaced the former outside of the UK, can be used to classify aberrant cervical cytology into defined stages of squamous cell carcinoma precursors (Solomon et al., 2002). Low grade squamous intraepithelial lesions (LSIL) are mild dysplasias corresponding to CIN grade I and regression of this class is often observed as the lesion is cleared by the immune system. High grade squamous intraepithelial lesions (HSIL) encompass the moderate or severe dysplasias of CIN II and III and carcinomas in-situ, which refers to pre-invasive cancer (Solomon et al., 2002). Persistent infection of cervical

epithelium with high risk types, particularly HPV16 is predictive of HSIL formation. Progression from these high grade carcinoma precursors to invasive squamous cell carcinoma of the cervix may then result, in which invasion of the basement membrane occurs.

Early studies of prophylactic vaccines within animals paved the way for the development of HPV vaccines which are based on the ability of the capsid proteins to self-assemble into virus-like particles (VLP) (Suzich et al., 1995; Breitburd et al., 1995). To date, two prophylactic vaccines have been granted licences by the European Union, the bivalent Cervarix® vaccine (GlaxoSmithKline) and the quadravalent Gardasil® vaccine (Merck) which are prepared from purified VLPs of the major L1 capsid protein of HPV types 16 and 18, with the latter also comprised of VLPs of HPV types 6 and 11 L1 capsid proteins. Vaccination with VLPs elicits a virus neutralising antibody response to prevent infection, and both vaccines have been shown to be highly effective against type-specific HPV infections. Trials indicate that both vaccines have greater than 90% efficacy at preventing CIN2 and CIN3 caused by HPV types 16 and 18, in the per-protocol susceptible population, negative for those HPV types (Hendrix, 2008). However as the vaccines are preventative, women already infected with these HPV types are not protected and vaccination must therefore be received prior to the onset of sexual activity (Kahn, 2009). Although some evidence of cross protection of HPV types has been observed, there is a continual need for cervical screening alongside the programme of vaccination, incurring significant costs. The high cost of the vaccine, in addition to poor healthcare infrastructure, may prevent implementation of vaccination programmes within developing countries where the disease burden is greatest (Agosti & Goldie, 2007).

1.4   The HPV genome

HPVs are small non-enveloped viruses with an icosahedral capsid ~55 nm in diameter encasing the viral genome (Longworth & Laimins, 2004). Genomic organisation is well conserved between HPV types and viral genomes are composed of ~8 Kb of double-stranded circular DNA which can be divided into three main regions: the long control region (LCR), and early (E) and late (L) regions (Figure 1.1). The LCR, sometimes referred to as the upstream regulatory region (URR) is a non-coding region which contains transcriptional, post-transcriptional and replicative cis-regulatory elements (Graham, 2008; Thierry, 2009). Located within this region is the viral origin of replication (ori), the early promoter and overlapping binding sites for transcriptional activators and repressors (Thierry, 2009). The viral genome is composed of eight open reading frames (ORF) which encode for proteins which can be categorised into early and late proteins dependent upon their location within the genome (Longworth & Laimins, 2004). The early region of the viral genome consists of genes designated E1-E7 which are responsible for transcription, viral genome replication and transformation of the host cell, whilst the late region contains the L1 and L2 genes which encode the structural capsid protein.

E6 and E7 oncogenes encode for the major proteins involved in cellular transformation and immune modulation, whilst the E5 membrane signalling protein also plays a role in transformation (Stanley, 2001). The E4 ORF is expressed as an E1^E4 fusion protein with the N terminus of the E1 gene fused to the E4 gene, derived by RNA splicing of the first five codons of E1 to the ORF of E4 (Roberts, 2006). However, despite being classed as an early gene, the E1^E4 protein is the most abundantly expressed protein during the late phase of the virus life cycle and is thought to play a major role within the productive stage of the virus life cycle (Roberts, 2006). Located within the late region of viral genome are the L1 and L2
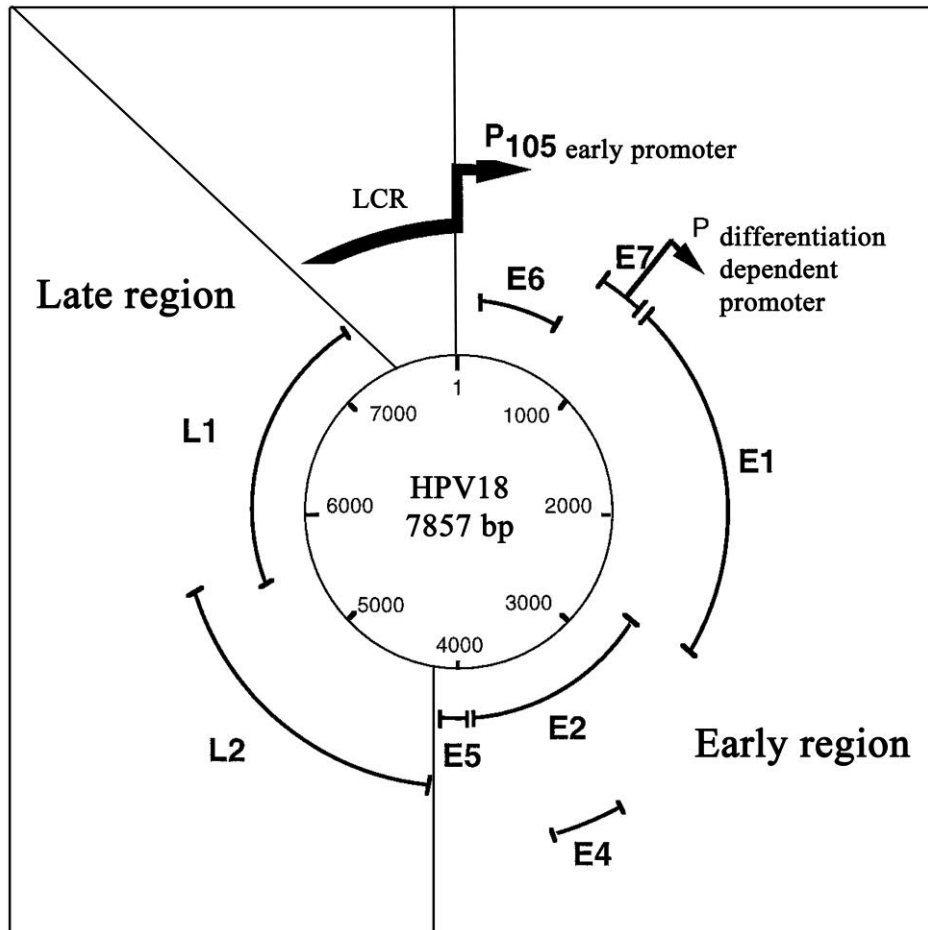
Figure 1.1 Genomic organisation of HPV18. The 7857 bp genome of HPV18 can be divided into three distinct regions: the long control region (LCR), the early and the late region. The eight open reading frames are divided between the early region, containing: E6, E7, E1, E2, E4 and E5 and the late region, containing the L1 and L2 ORF. The P_{105} early promoter lies upstream of the E6 ORF and the differentiation inducible promoter is located within the E7 ORF. Adapted from Desaintes and Demeret 1996.

ORFs which encodes for structural proteins, constituting the major and minor capsid proteins respectively and together form the icosahedral capsid (Longworth & Laimins, 2004).

1.5　　　The HPV life cycle

Papillomaviruses replicate within stratified epithelia and the life cycle of the virus is tightly linked to the differentiation programme of the host keratinocyte, and can be divided into early and late stages (Figure 1.2) (Stubenrauch & Laimins, 1999). Within the early stages of the virus life cycle following the initial infection, viral genomes are amplified to around 50 copies per cell and are subsequently maintained at a stable copy number within dividing basal cells (Kadaja et al., 2009). The productive phase then commences with differentiation-dependent late stages in which viral genomes are amplified to several thousand copies per cell before packaging into infectious virions which are released from the cell (Howley & Lowy, 2009).

1.5.1　　　Early stages of the virus life cycle

Initial papillomavirus infection occurs through microlesions, or mild abrasions of the epithelium and although the receptor for HPV entry has not been identified, heparin sulphate receptors have been shown to mediate initial attachment of the virus to the cell (Shafti-Keramat et al., 2003). Mechanisms for internalisation by receptor-mediated endocytosis appear to differ between papillomaviruses, with HPV16 infecting cells through clathrin-dependent endocytosis (Day et al., 2003) whilst HPV31 virions have been shown to be internalised by caveolae (Bousarghin et al., 2003). Disassembly of HPV particles occurswithin late endosomes or lysosomes and the viral DNA is transferred to the nucleus, chaperoned by the minor capsid protein L2 (Day et al., 2004).
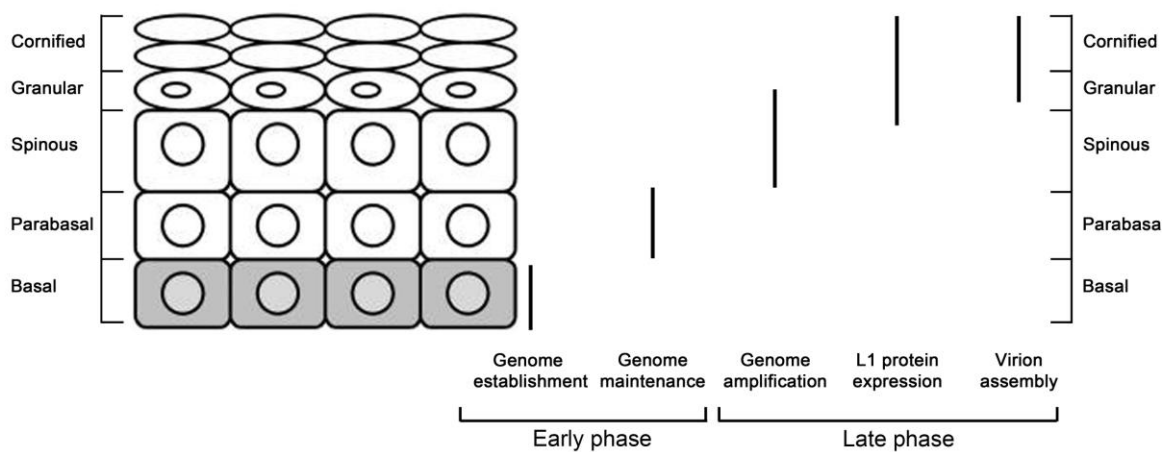
Figure 1.2 The HPV life cycle. Key stages of the HPV life cycle are shown in relation to the stage of keratinocyte differentiation in which they occur. Following HPV infection of the basal layer of the epithelium the early stage of the life cycle commences and viral genomes are established with a copy number of around 50 genomes per cell. As cells migrate to the parabasal layer the copy number of viral genomes are maintained, before the late stage of the virus life cycle begins as viral genomes undergo amplification within the spinous layer to several thousand viral genomes per cell. Following expression of capsid proteins, viral genomes are packaged into virions within the upper layers of the epithelium. Adapted from Knight et al 2004.

During early stages of the virus life cycle transcription from the early promoter, located immediately upstream of the E6 ORF (Figure 1.1), results in expression of the early proteins: E6, E7, E1 and E2 (Longworth & Laimins, 2004). The E1 and E2 proteins are required for replication of HPV genomes, however aside from these viral replicative proteins, the virus is also dependent upon the host cellular replication machinery (Chow & Broker, 2006). Within the nucleus of basal cells the viral genome is established as a multi-copy extrachromosomal element, with ~50 genome copies per cell (Stubenrauch & Laimins, 1999). The lower layer of the epithelium is composed of stem cells which divide, with the viral genome replicated concurrently with the cellular DNA and partitioned into daughter transit-amplifying cells. To ensure maintenance of genomes within daughter cells upon cell division, the E2 protein associates with the viral genome and tethers it to mitotic chromosomes, either through association with the cellular bromodomain-containing Brd4 protein, as shown with BPV1 (You et al., 2004) or directly to repetitive ribosomal sequences within the viral DNA as demonstrated with HPV8 (Poddar et al., 2009).

Following cell division of basal keratinocytes, one of the daughter cells continues to divide within the basal layer, whilst the other migrates upwards towards the suprabasal layers of the stratified epithelium, with the onset of terminal differentiation in the resultant progeny cells (Stubenrauch & Laimins, 1999). Paradoxically, HPV replication occurs within cells that have committed to terminal differentiation and have therefore exited from the cell cycle. To overcome this predicament, viral genome replication is supported within post-mitotic cells by the actions of the E7 oncoprotein which, through inactivation of the host tumour suppressor retinoblastoma protein (pRB), drives quiescent cells to re-enter S phase in which the host cellular replication machinery can be exploited (Cheng et al., 1995). As a result of unscheduled DNA replication cellular apoptosis is triggered within these cells, which is

counteracted by the E6 oncoprotein through inactivation of the host tumour suppressor protein p53 (Scheffner et al., 1990).

1.5.2   The differentiation-dependent late stages of the virus life cycle

The productive phase of the virus life cycle is dependent upon cellular differentiation, commencing within differentiating cells of the epithelium. For low risk HPV types including type 1, this stage initiates within parabasal layers, whilst for high risk HPV types including 16 this occurs later, commencing within the spinous layers of the epithelium (Peh et al., 2004). During the productive phase of the virus life cycle transcription occurs from the differentiation-dependent late promoter located within the E7 ORF resulting in expression of E1, E2, E1^E4, E5 and the late proteins L1 and L2 proteins (Graham, 2006). Differentiation triggers activation of productive replication and viral genomes are amplified to ~1000 copies per cell. High level expression of viral proteins occurs within the upper spinous and granular layers, with packaging of viral DNA into virions formed by the structural capsid proteins L1 and L2 (Graham, 2006). Infectious HPV particles are subsequently shed from differentiated epithelia.

1.6   Natural history of cervical cancer

Oncogenic HPVs are imperative for the pathogenesis of cervical cancer and the natural history of cervical carcinogenesis commences with infection of the cervical epithelium with HPV (Stanley, 2001). Whilst most papillomavirus infections are transient and rapidly cleared by the immune system, the presence of DNA from high risk HPV types can lead to the onset of cancer, although progression to malignancy occurs as a low incidence event (Stubenrauch & Laimins, 1999).   In low-grade lesions, genomes of high risk HPVs are maintained episomally and a productive infection results (Longworth & Laimins, 2004). However during

formation of a high-grade lesion, or progression to carcinoma, the genome becomes integrated into the host DNA with episomal integration ordinarily occurring within the E2 region of the HPV genome resulting in loss of expression of both E2 and E4, the latter of which is contained within the E2 ORF (Longworth & Laimins, 2004). This critical event terminates the productive life cycle due to a loss of expression of viral DNA sequences including the E2 replication protein which disables viral DNA replication, and the production of infectious particles therefore ceases. Persistent infection with oncogenic HPV increases the likelihood of a genome integration event (Ho et al., 1995) which often occurs early in the natural history of a HPV infection within CIN3 or in invasive carcinoma and is a strong indicator of malignancy (Collins et al., 2009).

As a result of viral genome integration, expression of E2 from integrated viral DNA is abrogated however, continued presence of episomal DNA within cells following an integration event results in expression of E2 which functions as a transcriptional regulator and inhibits expression of the E6 and E7 oncoproteins from integrated viral DNA (Pett et al., 2006). Subsequent loss of inhibitory episomal DNA from these cells however provides a selective growth advantage as loss of E2 expression releases repression of genes transcribed from the integrated viral DNA, enabling unregulated expression of the E6 and E7 oncoproteins (Pett et al., 2006). It is proposed that episomal loss occurs as a result of activation of antiviral genes which generate an environment in which episomal loss accompanied by integration of viral DNA confers a selective growth advantage (Pett et al., 2006).

During the normal infectious life cycle of the virus, expression of the viral E6 and E7 oncogenes is tightly regulated, however upon viral genome integration, and loss of episomal

DNA, deregulated expression of these proteins occurs, allowing inappropriate expression within dividing cells where their oncogenic potential can be exerted. Deregulated expression of E6 and E7 viral oncogenes acts during the malignant pathway to achieve abrogation of cell cycle controls and uncontrolled cellular proliferation ensues, resulting in cellular immortalisation and an increased probability of accumulation of secondary genetic changes which are required for progression to malignancy (Stanley 2001). .

1.7 Regulation of viral transcription during the HPV life cycle

Viral genes are transcribed from a single DNA strand into polycistronic mRNA from two main promoters: an early and late promoter, with usage dependent upon the differentiation state of the cell which divides the viral life cycle into early and late stages (Kalantari & Bernard, 2006). Aside from these two main promoters, further promoters are also likely to be utilised, and may play important roles in the virus life cycle (Ozbun & Meyers, 1998; Kim & Taylor, 2003; Milligan et al., 2007). Alternative splicing mechanisms act to regulate the combination of proteins expressed (Longworth & Laimins, 2004).

1.7.1 Regulation of transcription from the early promoter

During the early phase of the HPV life cycle the viral DNA is transcribed from the early promoter referred to as P_{97} in HPV types 16 and 31 and P_{105} in HPV18, lying immediately upstream of the E6 ORF within the LCR, with transcripts terminating at polyadenylation sites within E5 and L1 (Stubenrauch & Laimins, 1999). Polycistronic transcripts encoding the E1, E2, E6 and E7 proteins are derived from the early viral promoter with coding potential for the E5 and E1^E4 proteins (Figure 1.3). However these latter proteins are not detectable within undifferentiated cells and are likely to be expressed at low levels or not at all due to their
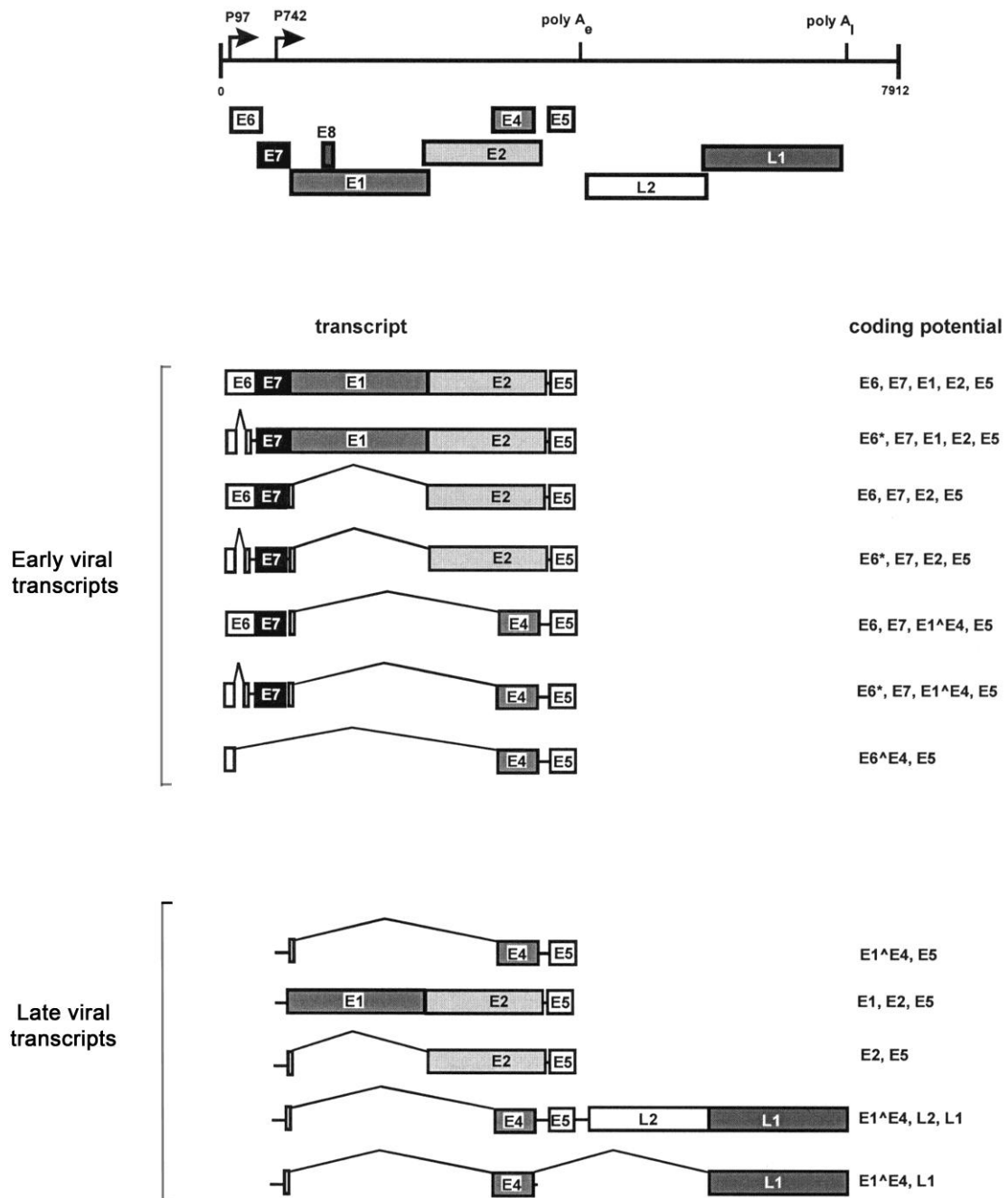
Figure 1.3 HPV31 early and late transcripts. A schematic showing a linearised representation of the HPV31 genome with early and late transcripts, which initiate from the early (P97) and late (P742) promoters respectively. Early (polyA0) and late (polyAl) polyadenylation sites are indicated. Adapted from Stubenrauch and Laimins 1999.

positioning as third or fourth ORFs within many early polycistronic RNA transcripts (Longworth & Laimins, 2004).

HPV transcription is regulated by cis-responsive elements within the LCR and involves the E1 and E2 viral proteins in addition to a number of cellular factors (Longworth & Laimins, 2004). The LCR contains binding sites for a number of cellular transcription factors including: SP1, AP-1 and -2, YY1 TFIID, YY1, OCT-1, NF-1, EPOC-1, KRF-1, TEF-1, TEF-2, glucocorticoid receptor (GRE) and the TATAAA sequence of the TATA box (Longworth & Laimins, 2004).

The E2 protein is well characterised as a viral transcription factor and has been shown to function as both a transcriptional activator or repressor dependent upon the levels of this protein within the cell (Bouvard et al., 1994). The organisation of the four E2 binding sites (E2BS) within the LCR is conserved between HPV types, with three E2BS located in close proximity with the promoter sequence, with the fourth binding site 400-500 bp upstream from the early promoter and separated by a tissue specific enhancer element (Morgan & Donaldson, 2006; Kalantari & Bernard, 2006). At low levels E2 binds only to the E2BS#4, for which the protein has the highest affinity and activates transcription (Steger & Corbach, 1997). However, as the levels of E2 accumulate within the cell as a result of enhanced transcription, the other E2BS become occupied by E2. As a result of binding of the E2 protein to the E2BS#1 and 2 most proximal to the TATA box, the TATA-binding protein (TBP) is unable to bind thereby preventing binding of the transcription initiation complex and inhibiting transcription (Dostatni et al., 1991). In addition, binding of E2 at E2BS#2 has also been shown to displace SP1 and TFIID transcription factors (Tan et al., 1994). HPV transcription is therefore regulated by a negative feedback loop whereby levels of

transcription are regulated by the abundance of E2 within the cell (Steger & Corbach, 1997). Regulation of HPV transcription from the early promoter acts to modulate expression of the early proteins.

1.7.2 Regulation of transcription from the late promoter

Within the stratum spinosum of the epithelium a higher level of viral transcription is induced from the late promoter sequence located within the E7 ORF, designated P_{742} in HPV31 or P_{670} in HPV16 (Graham, 2006). Utilisation of this promoter is dependent upon differentiation-specific cellular factors thereby limiting the expression of late transcripts to the terminally differentiating layers of the epithelium (Longworth & Laimins, 2004). Late viral transcripts from the differentiation-dependent promoters of HPV16 and 31 terminate at polyadenylation sites either at the end of the E5 or L1 ORF and therefore encode for the capsid proteins L1 and L2 in addition to the E1, E2, E1^E4 and E5 proteins (Figure 1.3).

Late gene expression is also regulated post-transcriptionally due to the presence of cis-acting RNA elements within late viral transcripts (Graham, 2006). A late regulatory element (LRE), overlapping with the 3' end of the L1 gene, represses late gene expression within undifferentiated cells, ensuring that the viral capsid proteins are not expressed until the cells reach the upper layers of the epithelium (Cumming et al., 2008).

1.8 Papillomavirus replication

1.8.1 The E1 and E2 replication proteins

E1 is expressed from the largest of the HPV ORFs, is well conserved amongst HPV types and has been shown to play an essential role in papillomavirus replication as an origin (ori) recognition protein initiator (Wilson et al., 2002; Chow & Broker, 2006). E1 proteins possess

ATPase activity and function as site specific DNA helicases which assemble at origins of replication and recruit host replication complexes (Hughes & Romanos, 1993; Chow & Broker, 2006). The localisation regulatory region (LRR) of the E1 protein contains a dominant nuclear export sequence (NES) and a nuclear localisation sequence (NLS) which are regulated by cyclin-dependent kinase (CDK) and mitogen-activated protein kinases (MAPK) phosphorylation respectively, with phosphorylation modulating the nuclear localisation of the protein (Deng et al., 2004; Yu et al., 2007). The LRR is highly conserved between HPV E1 proteins suggesting a common mechanism is utilised to regulate nuclear localisation of the protein.

The E2 protein, of 40-45 kDa, is integral to the regulation of HPV replication and transcription (Morgan & Donaldson, 2006). E2 is the primary ori binding HPV protein and consists of three main domains: an N-terminal transactivation (TA) domain, a central flexible hinge region and a C-terminal dimerisation and DNA binding domain (DBD) (Chow & Broker, 2006). The hinge region of E2 overlaps with the E4 ORF and is a highly divergent region of the protein, with variable sequence and length between E2 proteins of different HPV types, and is thought to provide a flexible linker between the two other domains. The hinge region of HPV11 E2 also contains the NLS required for nuclear localisation of the protein (Zou et al., 2000).

1.8.2 E1 and E2 directed replication

Replication of HPV DNA requires the E1 and E2 replication proteins, however due to the limited coding capacity of the HPV genome, the virus is also dependent upon the host cellular replication machinery for viral replication (Chow & Broker, 2006). The origin of replication (ori) of the HPV genome is an AT-rich region located within the LCR. The ori contains

several binding sites (BS) for the E2 protein (E2BS) which are comprised of the pallindromic sequence: $ACCGN_4CGGT$ to which E2 binds as a dimer, where N is AT rich for mucosal HPV types (Morgan & Donaldson, 2006). Flanked by E2BS is an E1BS which is composed of a series of overlapping variants of the hexameric sequence: AACAAT, which is present in six copies within the E1BS and it is thought that binding of E1 to these multiple sites will result in oligomeric E1 complex formation (Chen & Stenlund, 2001).

Both E1 and E2 proteins can bind independently to the origin, however the E1 protein alone displays low affinity and specificity, and binding of E2 enhances E1 binding, increasing the affinity of the protein for the AT rich sequences of the ori (Frattini & Laimins, 1994). The E1 initiator protein contains two distinct regions of DNA binding, the highly specific DNA binding domain (DBD) and non-specific binding activity of the helicase domain, which confers overall low specificity of the E1 protein for the ori (Stenlund, 2003). Formation of an $E1_2E2_2$-ori complex is essential for initiation of the ori and is mediated by association between E2 and the non-specific binding domain of the E1 protein, with this interaction inhibiting the ability of E1 to bind non-specifically to DNA and therefore restoring E1 binding specificity (Stenlund, 2003). E2 then disassociates from the $E1_2E2_2$ precursor complex resulting in an E1-ori complex, followed by sequential binding of further E1 proteins to form a di-hexameric complex which induces changes to the structure of the template DNA which are thought to open the DNA duplex (Chen & Stenlund, 2002). The di-hexameric E1 helicase complex unwinds the double stranded DNA at the replication fork and recruits a number of cellular replication proteins including replication protein A (RPA), which binds single stranded DNA preventing reassociation of double-stranded DNA (Loo & Melendy, 2004) and DNA polymerase α/primase, which is required for replication of DNA (Park et al., 1994; Masterson et al., 1998). Both E1 and E2 proteins have been shown to associate with

and activate topoisomerase I, which moves ahead of the replication fork breaking and rejoining the strands to prevent torsional strain (Clower et al., 2006a; Clower et al., 2006b). Whilst both E1 and E2 are required to initiate replication, only E1 is required for the elongation stage (Chow & Broker, 2006).

1.8.3 Differentiation-dependent replication of viral genomes

During the non-productive phase of the HPV life cycle within basal keratinocytes, E2 regulates its own expression and that of the E1 protein from the early promoter by an E2 auto-regulation mechanism (Morgan & Donaldson, 2006). When levels of E2 become high, transcription from the early promoter is inhibited, and the level of viral replication is stabilised with viral copy numbers maintained. Within dividing basal cells, the E2 protein acts to ensure equal segregation of viral genomes into daughter cells by localisation of genomes to the mitotic spindle during cell division (Van Tine et al., 2004). Upon a differentiation-dependent change in promoter usage, increased expression of E1 and E2 proteins occurs, due to a lack of E2 autoregulation mechanisms acting upon the late promoter (Klumpp & Laimins, 1999). Increased levels of the viral replication proteins results in increased replication within differentiated cells and the numbers of viral genomes are amplified from approximately 50 copies per cell to ~1000 copies during this productive phase of the HPV life cycle (Stubenrauch & Laimins, 1999)

A switch in the mechanism of replication of viral genomes is thought to occur upon cellular differentiation, from a theta structure in which replication occurs bi-directionally with two replication forks progressing in opposite directions, to a rolling circle method of replication whereby replication occurs in a uni-directional manner (Flores & Lambert, 1997).

1.9   HPV oncoproteins

1.9.1   Cellular transformation and immortalisation by HPV oncogenes

High risk HPV genomic DNA can extend the lifespan of primary human foreskin keratinocytes indefinitely and this function has been shown to be dependent upon the E6 and E7 oncoproteins which cooperate to immortalise these cell lines (Hawley-Nelson et al., 1989). The ability to immortalise primary keratinocytes is restricted to the high risk HPV types and low risk types are not in possession of transforming activity, with the ability to immortalise cells correlating with the oncogenic potential of the virus (Munger & Howley, 2002). Immortalised keratinocytes are not fully transformed and conversion to malignancy is dependent upon factors in addition to HPV. Co-transfection assays of primary baby rat kidney epithelial cells with HPV DNA with the activated *ras* oncogene demonstrated that high risk HPV types but not those of the low risk category are able to cooperate with *ras* to transform primary cells (Storey et al., 1988). The HPV E7 gene encodes for the major transforming function of high risk HPV types and expression of E7 is necessary and sufficient for transformation, whilst E6 contributes towards transformation within these cells.

1.9.2   The HPV E7 oncoprotein

HPV E7 is a small nuclear protein of around 100 aa which shares functional similarities with the E1A adenovirus protein and the SV40 large T antigen (TAg), and conserved regions (CR) shared between these viral oncoproteins are important for their transformation function (Howley & Lowy, 2009). E7 proteins contain three main domains: N-terminal CR1, centrally located CR2 and a C-terminal zinc binding domain located within CR3. CR2 contains two functional domains, a conserved LXCXE motif which forms a core binding site for association with pRB and the related pocket proteins (Munger et al., 1989) and an adjacent

casein kinase II (CKII) consensus site (Barbosa et al., 1990). CR3 contains a bi-partite sequence of two Cys-X-X-Cys elements which constitutes a zinc-binding domain at the C-terminus of the E7 protein. The C-terminal region of the E7 protein is required to mediate dimerisation of the protein and deletions within this region have also been shown to affect protein stability (Howley & Lowy, 2009). Sequences within CR3 of the E7 protein are also required in addition to the core pRB binding site to mediate association with pRB.

1.9.2.1   E7 and the G1/S transition

The progressive maturation of the keratinocyte is characterised by exit from the cell cycle and in normal epithelia, as the cell migrates from the basal layer, exit from the cell cycle occurs as the cell undergoes the process of differentiation. However in HPV-infected basal cells, E7 promotes direct entry of post-mitotic cells into S-phase where viral DNA replication can be supported by deregulation of the G1/S transition (Cheng et al., 1995). Regulated passage through the G1/S checkpoint ensures that environmental conditions are favourable and that the cellular DNA is undamaged prior to commitment of the cell to DNA replication. Stimulation of cells to divide by extracellular signals induces transcription of cyclin D which forms a complex with the CDK-4 or -6 catalytic subunits which then phosphorylate downstream target proteins of which the best characterised are those that constitute the retinoblastoma pathway (Sherr & McCormick, 2002). In its hypo-phosphorylated state the retinoblastoma protein forms an inhibitory complex with members of the E2F family of transcription factors, repressing transcription of their target genes which are essential for S-phase of the cell-cycle. Phosphorylation of pRB by the CDK4/6-cyclin D complex facilitates entry into S phase by inducing a conformational change in pRB which results in release of E2F. Up-regulation of the expression of genes required for DNA replication follows,

including S-phase cyclin A and the G1/S cyclin E, the latter of which will complex with CDK4/6 to phosphorylate pRB, thus acting to amplify the signal.

Deregulation of the RB pathway is a common event in carcinogenesis, with pRB thus defined as a tumour suppressor protein (Sherr & McCormick, 2002). Loss of pRB function results in constitutive E2F activity with entry into S phase no longer prevented by withdrawal of external stimuli, enabling inappropriate and uncontrolled cellular proliferation. The HPV E7 protein of high risk types can bind hypo-phosphorylated pRB (Dyson et al., 1989), disrupting the interaction between pRB and E2F and thereby inducing unregulated entry into S phase (McMurray *et al.,* 2000). Different HPV types have been shown to display varying degrees of affinity for pRB and a correlation with oncogenic potential has been proposed, with high risk E7 proteins possessing a significantly greater affinity for pRB than E7 proteins of low risk types (Munger et al., 1989). Whilst E7 proteins of low risk HPV types are able to associate with pRB, this interaction is not sufficient to enable activation of E2F responsive genes (Longworth & Laimins, 2004). High risk HPV E7 proteins induce degradation of the RB protein by the proteosome and this has been shown to be dependent upon the cellular protease, calpain (Boyer et al., 1996; Darnell et al., 2007). In addition to pRB, E7 has also been shown to associate with other proteins of the pocket protein family including p107 and p130 which also regulate E2F transcription (Dyson et al., 1989; Davies et al., 1993; Hu et al., 1995).

Properties aside from the ability of E7 to associate with pRB are however thought to mediate cellular transformation. E7 mutants have been identified which are able to associate with pRB but which are not able to effectively transform cells indicating that another function of the E7 protein is required for cellular transformation (Banks et al., 1990). Mutation of a

single cysteine residue of one of the zinc binding elements CXXC, completely abrogates the ability of the high risk HPV16 genome to immortalise cells suggesting that further E7 functions are involved in cellular transformation (Chesters et al., 1990). Furthermore, using genetically engineered mice it has been shown that disruption of the interaction of E7 with pRB does not prevent E7-dependent formation of cervical cancers within mice (Balsitis et al., 2006). Together these studies demonstrate that pRB independent activities of E7 contribute towards carcinogenesis. E7 has been shown to associate with cyclins A and E which, in complex with CDK2, drive progression through the cell cycle through phosphorylation of pRb (Longworth and Laimins 2004). Interaction with E7 has been shown to increase the levels of these proteins, with the resultant effect being to enhance G1 to S progression. E7 also binds to and inactivates the CDK inhibitors p21 and p27, releasing their inhibition of the G1 CDK-cyclin complexes and thereby acting to abrogate the G1/S checkpoint (Cho *et al.,* 2002). The ability of E7 to inactivate p21 has been shown to contribute towards carcinogenesis (Shin et al., 2009). Expression of a mutant E7 protein which is unable to inactivate p21 resulted in reduced formation of E7 induced cervical cancers within mice in comparison with those which expressed the wild type E7 protein (Shin et al., 2009).

An interaction between E7 and histone deacetylases (HDACs) has also been characterised whereby the association displaces HDACs from the Rb protein, preventing their recruitment to E2F-inducible promoters at which they ordinarily act to repress promoter expression, resulting in activation of E2F transcription (Longworth et al., 2005). E7 can also associate with histone acetyl transferases (HAT) including p300 and pCAF and inactivate the transcriptional co-activator function of these proteins (Avvakumov et al., 2003; Bernat et al., 2003). This may represent another mechanism through which E7 may activate E2F-mediated transcription and thus promote entry into S-phase.

1.9.2.2   Further functions of HPV E7

In addition to regulators of the cell cycle E7 has been shown to associate with numerous cellular proteins, and notably a role for E7 has been demonstrated in regulation of transcription. HPV16 E7 protein has been shown to associate with TBP and inhibit binding of the protein to DNA. Phosphorylation of E7 by CKII modulates this E7 inhibitory function by enhancing binding of the protein to TBP (Massimi et al., 1996; Maldonado et al., 2002). A correlation has also been observed between the rate at which E7 proteins undergo CKII phosphorylation and their oncogenic potential, with high risk HPV types phosphorylated by the kinase to a greater extent (Barbosa et al., 1990). Furthermore, phosphorylation of the CKII motif of E7 has been shown to enhance the ability of the E7 protein to promote degradation of the p130 pocket protein within differentiated keratinocytes and to drive entry of cells into S phase (Genovese et al., 2008).

A role for E7 has also been proposed in induction of genomic instability. E7 can induce aberrant duplication of centrioles which coordinate chromosome segregation during cell division, resulting in aneuploidy, whereby cells contain an abnormal number of chromosomes (Duensing et al., 2001). Aneuploidy is often detected in pre-malignant lesions and contributes towards genomic instability and tumourigenesis (Laughlin-Drubin & Munger, 2009).

1.9.3   The HPV E6 oncoprotein

HPV E6 oncoproteins are approximately 150 amino acids in length and share functional similarity with the adenovirus E1B and SV40 TAg viral oncoproteins (Thomas et al., 2006). E6 proteins contain four highly conserved zinc binding Cys-X-X-Cys domains which enable formation of two zinc fingers (Howie et al., 2009). The presence of a PSD-95/discs large/ZO1 (PDZ) binding domain at the extreme C-terminus is restricted to high risk mucosal

HPV types, and mediates binding of E6 proteins to a number of cellular proteins containing the PDZ domain (Thomas et al., 2006).

1.9.3.1   E6 and the G1/S transition

Within normal cells the integrity of DNA at the G1 to S transition is assessed by the p53 tumour suppressor, a cellular transcription factor upon which several DNA damage pathways converge (Sherr & McCormick, 2002). The p53 protein functions as a regulator of cell growth and is induced in response to cellular stress resulting in transcriptional upregulation of cell cycle arrest and pro-apoptotic genes, activating a cascade which terminates in apoptosis, the self-destruction of the cell. As a consequence of unscheduled DNA replication directed by the HPV E7 protein, the p53 pathway is activated as virus-infected cells activate this apoptotic pathway and prevent spread of the virus to the surrounding cells and the E6 protein acts to counteract this effect (Thomas et al., 2006).

HPV E6 proteins of both low and high risk HPV types have been shown to interact with p53 (Werness et al., 1990), a function common to a number of other viral oncoproteins including adenovirus E1B and the SV40 TAg (Thomas et al., 2006). Association of E6 proteins of high risk types results in degradation of p53, although low risk E6 proteins do not share this function (Scheffner et al., 1990). High risk E6 proteins first associate with the ubiquitin ligase E6 associated protein (E6-AP), which has been has been shown to positively regulate the stability of E6 (Tomaic et al., 2009), before forming a trimeric complex with p53 (Mantovani & Banks, 2001). This interaction results in ubiquitination of p53, marking the protein for degradation via the proteosome (Scheffner et al., 1993). The inability of low risk E7 and E6 proteins to degrade pRB and p53 respectively suggests that this function is not required for the life cycle of these low risk HPV types. Low risk HPV viruses replicate

within lower layers of the epithelium in which cellular proliferation has not yet ceased, whilst high risk types replicated within cells which have committed to differentiation and have exited from the cell cycle (Egawa et al., 2000; Peh et al., 2004). It is thought that initiation of DNA replication within the lower layers of the epithelium where the cellular replicative machinery is still present may negate the need to drive cells into S phase and therefore p53 degradation may not be required (Thomas et al., 2006).

Aside from degradation of p53, E6 has also been shown to down-regulate the activity of this protein through an alternative mechanism, with both high and low risk E6 proteins able to inhibit the transcriptional activity of p53 through inhibition of the DNA binding ability of the protein (Crook et al., 1994; Lechner & Laimins, 1994). Transcriptional activation of p53 is enhanced by association with the CBP/p300 transcriptional co-activator which recruits transcriptional machinery and can modify chromatin structure at promoter sequences. High risk E6 proteins have been shown to associate with, and inhibit CBP/p300 resulting in repression of p53 mediated transcription (Zimmermann et al., 1999; Patel et al., 1999)

1.9.3.2 p53-independent inhibition of apoptosis

E6 can also inhibit apoptosis in a p53-independent manner and E6 proteins of both high and low risk types can associate with and direct degradation of Bak, a pro-apoptotic protein belonging to the Bcl-2 family of proteins (Thomas & Banks, 1998; Thomas & Banks, 1999). Bak is an important regulator of apoptosis, located within the membrane of the mitochondria whereby in response to cellular stress signalling the protein promotes apoptosis by forming pores within the membrane enabling release of cytochrome c, a key modulator of the apoptosis pathway (Howie et al., 2009). A caspase signalling pathway is triggered by release of cytochrome c, ultimately resulting in cellular death. By targeted degradation of Bak, E6

acts to block induction of apoptosis through the mitochondrial intrinsic pathway (Thomas & Banks, 1998; Thomas & Banks, 1999).

Papillomaviruses have also evolved mechanisms to evade apoptosis mediated by the host immune system (Howie et al., 2009). E6 has been shown to counteract apoptotic signalling through the extrinsic pathway, mediated by cytokine receptors, by interacting with regulators of these signalling pathways including members of tumour necrosis factor (TNF), Fas and TNF-related apoptosis-inducing ligand (TRAIL) signalling pathways (Howie et al., 2009). Binding of the TNF ligand to its cognate receptor: TNFR1, triggers a signalling cascade resulting in recruitment of a number of proteins including the adaptor proteins: TNFR1-associated death domain (TRADD), followed by Fas-associated death domain (FADD) then the initiator protein procaspase 8 which leads to a caspase signalling cascade ultimately resulting in cell death. Association of E6 with TNFR1 has been shown to inhibit association of the receptor with the TRADD adaptor protein, preventing formation of a signalling complex which leads to apoptosis (Filippova et al., 2002). E6 also blocks transmission of apoptotic signalling through the Fas and TRAIL pathways through association and degradation of the FADD adaptor protein (Filippova et al., 2004) and caspase 8, which are shared adapter proteins of Fas and TRAIL mediated apoptosis (Garnett et al., 2006).

1.9.3.3 PDZ association

HPV E6 proteins of high risk virus types target a number of cellular proteins involved in cell polarity, including several proteins which contain a PDZ domain of around 80-90 amino acids (Thomas et al., 2008). The PDZ binding domain of high risk HPV E6 proteins is located at the extreme C-terminus of the protein and is composed of a tetra-peptide consensus binding sequence: X[T/S]X[V/L] (Thomas et al., 2008). Association of high risk E6 proteins has been

shown to result in degradation of a number of PDZ proteins including the human homologue of *Drosophila* discs-large (hDlg) tumour suppressor protein (Kiyono et al., 1997), MAGI-1,-2,-3 proteins (Glaunsinger et al., 2000; Thomas et al., 2002), human homologue of the *Drosophila* Scribble (hScrib) (Nakagawa & Huibregtse, 2000) and tumour suppressor protein MUPP1 (Lee et al., 2000). PDZ proteins are often located at sites of cell to cell contact and play a number of roles in epithelial adhesion, polarity, differentiation and proliferation. Degradation of these PDZ proteins by E6 disrupts these cellular pathways and is thought to contribute towards malignancy (Thomas et al., 2008).

Dlg is a component of cellular adhesion pathways and localises at sites of cell-cell contact. A reduction in Dlg correlates with progression from low grade lesions to cervical carcinogenesis, indicating that loss of this protein may contribute towards malignancy (Watson et al., 2002). The PDZ binding motif of E6 contributes towards reduced localisation of Dlg to sites of cellular adhesion and disruption of intercellular junction formation, and this domain has been shown to contribute towards E6-dependent morphological transformation of keratinocytes (Watson et al., 2003). Together these studies suggest that E6 may contribute towards carcinogenesis through disruption of cell to cell association. Phosphorylation of E6 proteins of HPV types 18 and 16 within a C-terminal PDZ binding domain by protein kinase A (PKA) has been shown to negatively regulate the ability of this viral protein to mediate degradation of the Dlg protein (Kuhne et al., 2000). Mutation of the PKA phosphorylation site resulted in enhanced morphological transformation indicating that E6 PKA phosphorylation may negatively regulate the transformation activity of E6 (Watson et al., 2003).

1.9.3.4   Activation of telomerase

In addition to preventing p53 activity, other roles for E6 in the immortalisation of human cells have been characterised. E6 has the ability to activate telomerase, a ribonucleoprotein responsible for the addition of telomere repeats to chromosome ends (Klingelhutz et al., 1996). Telomeres at the chromosome termini become progressively shorter with each successive cell division and in order for a cell to progress to malignancy activation of the protective cellular senescence pathway which is triggered by critically shortened chromosome ends must be avoided. By activating the expression of the catalytic component of telomerase, hTERT, E6 reconstitutes telomerase activity which is ordinarily absent from the normal cell and induction of cellular senescence can be avoided (Oh et al., 2001). E6 has also been shown to interact directly with hTERT and increase telomerase activity post-transcriptionally, representing a second mechanism by which E6 targets this enzymatic complex (Liu et al., 2009).

The combined functions of E6 allow the cell to eliminate DNA damage checkpoints, resulting in cellular proliferation in the presence of a damaged genome, an action that enables mutations to be transmitted to progeny cells with the possibility of development of a cancerous mutation.

1.10   The HPV E5 protein

The E5 protein is a small hydrophobic protein of around 80 amino acids which is localised to the membrane of the Golgi, the endoplasmic reticulum and the nuclear membrane (DiMaio & Mattoon, 2001). Whilst the E5 protein constitutes the major transforming activity of bovine papillomavirus (BPV), the HPV E5 protein possesses only a weak ability to transform cells and E5 proteins of these different papillomaviruses share little sequence homology.

Expression of E5 occurs during both early and late stages of the virus life cycle, however integration of the HPV genome into host DNA frequently results in loss of the E5 coding region, indicating that E5 is not required for carcinogenesis although it is thought that protein plays a role during a productive infection (Tsai & Chen, 2003).

Expression of HPV E5 results in increase cellular proliferation dependent upon expression of epidermal growth factor (EGF) and in the presence of this ligand, E5 has been shown to increase activation of the cognate receptor: EGFR (Crusius et al., 1998). Following binding of the ligand, EGFR is internalised by endocytosis whereby the receptor is either recycled to the cell surface, or more frequently undergoes degradation within lysosomes. It is proposed that E5 mediates enhanced EGFR activity by delaying internalisation of the receptor from the cell surface and through enhanced recycling of the receptor (Straight et al., 1993). E5 has also been shown to associate with the 16 kDa subunit of the vacuolar H+ ATPase and prevent this pump from acidifying endosomes which acts to enable disassociation of ligands from receptors (Conrad et al., 1993; Straight et al., 1995). E5 has also been shown to inhibit endocytic protein trafficking between cellular compartments, and it is proposed that by interfering with the process of endocytosis this may contribute towards the E5 function of delayed degradation of EGFR, resulting in increased EGFR signalling (Thomsen et al., 2000). The EGF signalling pathway results in activation of MAPK which have been shown to regulate a number of cellular activities including proliferation and differentiation. Activation of MAPK by E5 has been shown to occur in both an EGFR-dependent and independent manner, and it is proposed that induction of cellular signalling cascades which promote enhanced cellular growth enables the E5 protein to support a productive infection (Crusius et al., 2000).

Aside from the role of E5 in cellular proliferation, E5 proteins display a diverse range of other functions including the ability to contribute towards the differentiation-dependent late viral functions of viral genome amplification and late gene expression (Fehrmann et al., 2003). A further function for E5 has also been suggested in promotion of cellular survival, with E5 shown to prevent DNA damage induced apoptosis (Zhang et al., 2002). Furthermore, E5 proteins have been shown to down-regulate the expression of major histocompatibility complex (MHC) class I and class II expression at the cell surface which may contribute towards the ability of the virus to evade detection by T cells of the host immune system thus supporting a productive infection (Ashrafi et al., 2006).

1.11 The HPV L1 and L2 structural capsid proteins

1.11.1 Assembly of HPV virions

Assembly of viral genomes into infectious particles occurs within the upper layers of the epithelium, with expression of the L1 and L2 structural proteins, which constitute the viral capsid, occurring shortly after the onset of viral genome amplification within the granular layer of the epithelium (Pereira et al., 2009). Activation of the differentiation-dependent late promoter results in production of viral transcripts from which the late proteins L1 and L2 are transcribed. It is thought that the rare codon usage of L1 and L2 genes may also contribute towards restricting their expression to the upper layers of the epithelium where the abundance of rare tRNA increases (Zhou et al., 1999). Following expression, the capsid proteins relocate from the cytoplasm to the nucleus and assemble into icosahedral capsids in which the viral genomes are packaged (Conway & Meyers, 2009).

The HPV icosahedral capsid is composed of 72 capsomeres, each composed of five L1 major capsid proteins forming a pentamer, with a hole in the centre in which the minor capsid protein L2 is thought to associate (Holmgren et al., 2005). Whilst the exact number of L2 particles within each virion is unknown, estimates have been made with a ratio of L1:L2 particles at 30:1, with 12 L2 particles present within each virion, although estimates of L2 proteins have been made as high as 72 (Conway & Meyers, 2009). Although L1 can spontaneously self-assemble to form an icosahedral VLP , L2 plays an integral role in assembly of virions and is required for efficient encapsidation of viral DNA (Holmgren et al., 2005).

L2 accumulates at ND10 bodies (Day et al., 1998; Day et al., 2004), nuclear domains which are proposed to be sites of papillomavirus genome replication (Swindle et al., 1999). L2 has also been shown to associate with E2 and recruit this protein to ND10 bodies, and it is proposed that the E2 protein plays a role in encapsidation by recruiting viral genomes to PML domains through site-specific DNA interactions (Day et al., 1998; Heino et al., 2000). L2 association induces reorganisation of ND10 bodies, which is followed by recruitment of pre-formed L1 capsomeres and higher order assembly of L1 and L2 proteins into an icosahedral virion in which viral genomes are encapsulated (Florin et al., 2002a; Florin et al., 2002b). During assembly, the capsid undergoes a process of maturation in which it undergoes modifications which render it more stable and resistant to proteolytic degradation. As part of this process, disulphide bonds form between adjacent L1 proteins which act to stabilise the capsid (Buck et al., 2005). The exact mechanism of release of HPV virions is unknown although it is proposed that HPV infection may induce abnormalities to the cornified cell envelope (CCE), a highly insoluble structure which forms on the inside of the plasma membrane of cells on the outermost layer of the epithelium. Increasing the fragility of the

CCE may enable release of infectious particles from dead cornified squame cells, as opposed to through cellular lysis (Bryan & Brown, 2000). Restricted expression of the immunogenic capsid proteins to differentiated cells, and the retention of the infectious virion until the cell reaches the upper layers of the epithelium, acts to aid immune evasion of virally infected cells.

1.11.2   Viral entry

The L1 and L2 capsid proteins play key roles in establishment of a viral infection and are implicated in the process of cellular entry, uncoating and delivery of viral genomes to the nucleus. The L1 capsid protein can mediate internalisation of the virus although the presence of L2 enhances the infectivity of the virus (Pereira et al., 2009). The nature of the receptor that mediates internalisation of infectious papillomavirus virions is still under debate, although initial binding of the virion to the cell surface is shown to dependent upon heparin sulphate proteoglycan (HSPG) (Giroglou et al., 2001). L1 binding to HSPG is proposed to induce a conformational change which unmasks previously inaccessible sites of the virion to either facilitate entry via HSPG or a second receptor (Shafti-Keramat et al., 2003). Following internalisation, the viral capsid disassembles within the endosome compartment and N-terminal cleavage of L2 by furin, a proprotein convertase is necessary for translocation of the capsid protein into the cytoplasm (Richards et al., 2006). L2 mediates relocation of the viral genome from the endosomal compartment into the nucleus, where the L2-chaperoned viral genome accumulates at ND10 bodies (Day et al., 2004).

1.12   The HPV E1^E4 protein

The E4 ORF is located within the E2 ORF and although both proteins are encoded by the same DNA strand, translation occurs in separate reading frames and therefore the proteins

share no sequence homology. The E4 protein is expressed as a fusion protein with N-terminal residues of the E1 gene fused to the E4 gene, derived by RNA splicing of the first five codons of E1 to the ORF of E4 resulting in expression of an E1^E4 protein (Nasseri et al., 1987). Whilst the E4 ORF is lost upon integration of viral genomes and is therefore not expressed within malignant lesions, the E1^E4 protein is the most highly expressed protein in the HPV productive life cycle, with the HPV1 E1^E4 protein estimated to constitute around 30% of total wart proteins (Doorbar et al., 1986) and it is therefore proposed that this protein plays a major role in the virus life cycle.

The E4 ORF is located within the early region of the HPV genome and is present within early viral transcripts, although it is positioned as the third ORF of polycistronic mRNAs and therefore little of the protein is likely to be translated within undifferentiated cells (Longworth & Laimins, 2004). E1^E4 is therefore described as a late HPV protein due to expression predominantly being restricted to the suprabasal layers of the epithelia, concomitant with cellular differentiation, where usage of the late promoter results in positioning of E4 as the first ORF in a number of late transcripts (Figure 1.3). E1^E4 is first expressed within the upper layers of the epithelium, with expression coincident with the onset of vegetative viral genome amplification, suggesting a role for the protein in initiation of the productive phase of the HPV life cycle (Peh et al., 2002). The continued presence of E1^E4 in cells within the uppermost layers of the epithelium in which virion assembly occurs, also implies a function for this protein during the later stages of virus production and egress (Peh et al., 2002). The distribution of E1^E4 within the cell varies depending upon the HPV species from which it originates, but the protein is predominantly located within the cytoplasm although in several HPV types E1^E4 can also be observed within the nucleus (Doorbar et al., 1991; Roberts et al., 1993; Roberts et al., 1994; Pray & Laimins, 1995; Brown et al., 2004).

An implication of the overlapping nature of the E4 and E2 ORFs with HPV genomes is that the evolution of the E2 and E1^E4 proteins are interlinked. This could suggest that evolution of these proteins would be constrained and that E1^E4 proteins of different HPV types would be similar. However sequence alignment of E1^E4 proteins reveals that there is little sequence homology between HPV types, and furthermore there is great variation in the size of E1^E4 proteins. Whilst the HPV18 and HPV16 E1^E4 mucosal proteins for example, are both around 90 aa in length, the E1^E4 protein of the cutaneous HPV1 virus is much greater in size and contains 125 aa (Figure 1.4). The E4 ORF lies within the hinge region of the E2 ORF, which links the two major functional domains of the E2 protein and is an extremely divergent region of this protein (Zhu, 1999). This may explain why E1^E4 protein homology can vary so greatly between HPV types despite being contained with the E2 ORF.

Whilst there are considerable differences between E1^E4 proteins of different HPV types, particularly within the central region of the protein, certain domains show a greater degree of conservation. These include an N-terminal leucine-rich motif, which is highly conserved between E1^E4 proteins of alpha types and a C-terminal homology domain is also conserved between these HPV types (Figure 1.5). These domains have been shown to mediate a number of E1^E4 functions (Roberts et al., 1994; Roberts et al., 1997; Raj et al., 2004). Therefore whilst E1^E4 proteins may be highly divergent they share a number of functional domains which are conserved between different HPV types and E1^E4 proteins may therefore play similar roles during the HPV life cycle. Although the exact role of the most highly expressed protein in the productive life cycle of HPV is yet to be fully defined, evidence suggests that the E1^E4 protein may have multiple roles at different stages in the viral life cycle and a number of functions have been assigned to this protein with may of these functions shared between E1^E4 proteins of different HPV types.

```
HPV18 E1^E4          MADPEVPV-----------TTRYPLLSLLNS--YSTPPHRIPAPCPWAQRPTARRRLLHD 47
HPV16 E1^E4          MADPAA-------------ATKYPLLKLLGSTWPTTPPRPIPKPSPWA--PKKHRRLSSD 45
HPV1  E1^E4          MADNKAPQGLLGLLQYTPTTQPYPRVTPPSNRRPSTTPNSQDRGRPRRSDKDSRKHLYAD 60
                     ***                  **          * *        *          *  *

HPV18 E1^E4          -LDTVDSRRSSIV---------DLS----THFSVQ---LHLQAT--------TKDGNSVV 82
HPV16 E1^E4          -QDQSQTPETPAT---------PLSCCTETQWTVLQSSLHLTAH--------TKDGLTVI 87
HPV1  E1^E4          GLTDGEDPEVPEVEDKKKENQRPLG---HPDLSLLRETLEVYTQRLKRDILQDLDDFCRK 117
                                             *             *                 *

HPV18 E1^E4          VTLRL--- 87
HPV16 E1^E4          VTLHP--- 92
HPV1  E1^E4          LGIHPWSV 125
```

Figure 1.4 Sequence alignment of HPV1, HPV16 and HPV18 E1^E4 proteins. Alignment of E1^E4 protein sequences was performed using the multiple sequence alignment programme CLUSTAL W 2.0 (Larkin et al., 2007). The single letter amino acid code is used, with invariant positions indicated by asterisks.

1.12.1   Association and disruption of cellular structures

As the keratinocyte undergoes terminal differentiation, keratin filaments aggregate in a process known as keratinisation, providing a tough impermeable barrier to the cell. It is proposed that E1^E4 may function to jeopardise the keratinisation process, an action that could aid the later stages of the viral life cycle by facilitating the release of infectious virus particles (Doorbar et al., 1991). E1^E4 proteins associate with the keratin intermediate filament (IF) network within the epithelial cell and in the case of the HPV16 protein, this association has been shown to induce collapse of the cytokeratin matrix (Doorbar et al., 1991). Whilst both HPV1 and HPV16 E1^E4 co-localise with keratin filaments, only the high risk type 16 inducing keratin network collapse and it has been suggested that this ability may be specific to E1^E4 proteins of high risk HPV types (Roberts et al., 1993). The highly conserved leucine-rich motif (LLXLL) located at the N-terminus of the E1^E4 protein is required for cytokeratin interaction (Roberts et al., 1994), and a hydrophobic C-terminal region conserved between mucosal proteins has been shown to be essential for induction of cytoskeletal collapse (Roberts et al., 1997). This latter domain has also been shown to mediate E1^E4 oligomerisation suggesting that self-association of E1^E4 proteins may allow cross-linking of keratin filaments, leading to disruption of the cytokeratin network (Roberts et al., 1997). Disruption of the cytokeratin network alters the structural integrity of the keratinocyte and may facilitate infectious virion egress and this may indicate a role for E1^E4 in enhancing this later stage of the viral life cycle (Doorbar et al., 1991).

An additional means by which E1^E4 may act at this late stage in the viral life cycle may be via weakening of the cornified cell envelope (CCE) and an association between HPV11 E1^E4 and the CCE of keratinocytes has been demonstrated. E1^E4 may play a role in
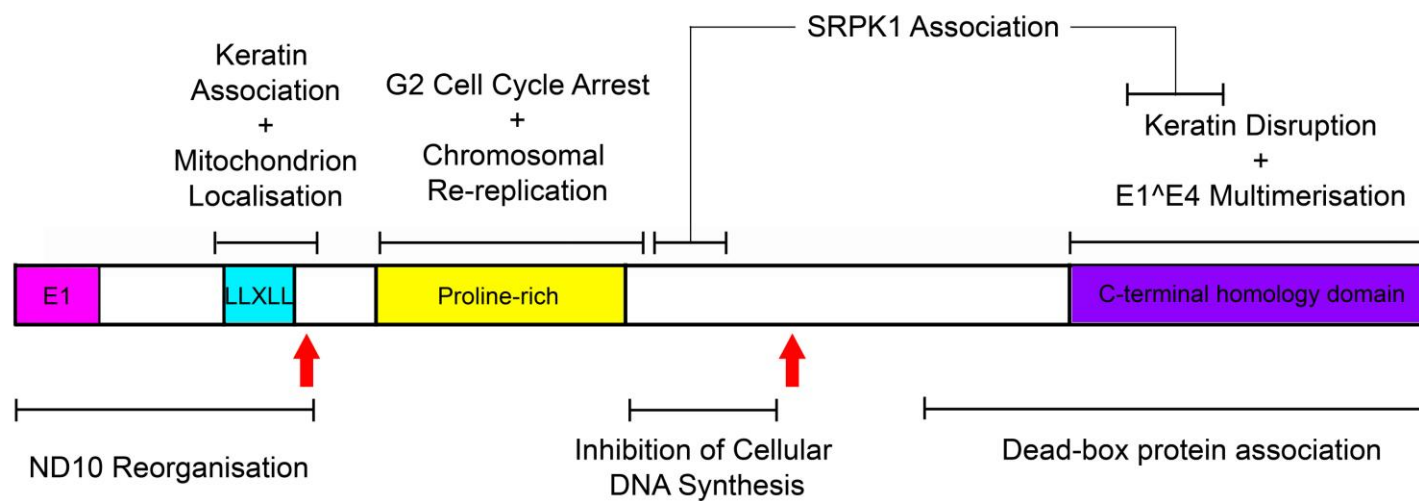
Figure 1.5: Functional domains of the E1^E4 protein. A schematic representation of the E1^E4 protein with known functional regions of the protein depicted. Positions of proteolytic cleavage of the HPV1 E1^E4 protein are indicated with red arrows. Adapted from Roberts 2006.

promoting the structural abnormalities and fragility of the CCE observed within HPV11-infected keratinocytes (Bryan & Brown, 2000).

A further mechanism by which E1^E4 may compromise the structural integrity of the keratinocyte is through reorganisation of cytoplasmic organelles and an association between HPV16 E1^E4 and mitochondria has been characterised, although a direct interaction has yet to be demonstrated (Raj et al., 2004). The observed redistribution of mitochondria from the microtubule network to clusters adjacent to the nucleus is dependent upon the leucine-rich cluster of E1^E4. Expression of E1^E4 also caused a marked reduction in mitochondrial membrane potential, with eventual induction of apoptosis which would render the cell more fragile and could aid the release of progeny virus (Raj et al., 2004).

1.12.2 Cell cycle disruption

E1^E4 has been implicated in modulation of the cell cycle, with expression of the E1^E4 proteins from HPV types: 1, 11, 16 and 18 in human keratinocytes resulting in induction of a G2/M arrest (Nakahara et al., 2002; Davy et al., 2002; Knight et al., 2004). The G2/M checkpoint ensures successful completion of DNA replication prior to entry into mitosis, preventing the proliferation of damaged cells and thereby maintaining genomic integrity (Houtgraaf et al., 2006). Progression of the cell cycle from G2 into mitosis is driven by the CDK1-cyclin B1 complex, which phosphorylates a whole host of proteins required during mitosis. The phosphorylation status of CDK1 determines the catalytic activity of the M phase CDK1-cyclin B1 complex. This kinase complex accumulates in the cytoplasm prior to entry into mitosis but is held in an inactive state by inhibitory phosphorylations of Thr14 and Tyr15 by Myt1 and Wee1 respectively. Dephosphorylation of these CDK1 residues by the phosphatase: Cdc25 activates the M-phase CDK, provoking passage from G2 into M phase.

DNA damage checkpoints converge upon proteins which regulate passage through these checkpoints rendering these proteins inactive until DNA replication is complete or until the damaged DNA is repaired (Houtgraaf et al., 2006).

Stemming from the requirement for the cellular replicative machinery, which is absent from the terminally differentiated cell, papillomavirus E6 and E7 oncoproteins together drive re-entry of the host cell into S phase in order to enable viral DNA replication (Longworth & Laimins, 2004). However within this environment, the virus must compete with the host for access to replication factors. It is proposed that in order to overcome this disadvantage, induction of a G2 arrest by the E1^E4 protein would force the cell into a 'pseudo S-phase' in which replicative enzymes become abundant but in which the cellular DNA fails to replicate, thus generating an environment in which viral DNA can replicate in the absence of competition (Knight et al., 2004). The ability of the E1^E4 protein to induce a G2 arrest is hypothesised to enable an environment conducive viral genome amplification.

Whilst E1^E4 proteins of both low and high risk HPV types, have been shown to induce a G2 arrest, the mechanisms by which this is conveyed differ, with E1^E4 proteins targeting different components of the G2/M transition to mediate this effect (Knight et al., 2004; Davy et al., 2005; Knight et al., 2006; Davy et al., 2006). Although expression of the full length 17 kDa E1^E4 protein of HPV1 alone does not perturb progression of the cell-cycle, a truncated 16 kDa E4 species, derived by N-terminal proteolytic cleavage of the full-length protein, is adept at inducing a G2 arrest (Knight et al., 2004). Following expression of the 16 kDa E1^E4 protein within cells, the levels of cyclin B were distinctly reduced and expression of endogenous cyclin B1 abolished the cell-cycle arrest suggesting that the G2 arrest induced by expression of the 16 kDa E4 protein is as result of the presence of low levels of active CDK1-

cyclin B complexes (Knight et al., 2006). Upon co-expression of both the full length 17 kDa and truncated 16 kDa HPV1 E4 proteins, a G2 arrest is also observed although this occurs by a distinct means from that of the 16 kDa protein alone. Within these cells the levels of the Wee1 kinase were maintained, resulting in induction of a G2 arrest by inhibitory phosphorylation of CDK1 (Knight et al., 2006). Co-expression of both 17 and 16 kDa E4 proteins prevents re-replication of the chromosomal DNA and it is hypothesised that these E4 proteins may act to block competing cellular DNA synthesis to enable efficient viral DNA replication (Knight et al., 2004).

The mechanism by which HPV16 E1^E4 instigates a G2 arrest does not appear to involve inhibition of CDK1-cyclin B activity but is instead dependent upon sequestration of this complex away from the nucleus in which the targets of the enzyme complex are located (Davy et al., 2005). HPV16 E1^E4 has also been shown to associate with, and cytoplasmically sequester, the CDK2-cyclin A complex, which has also been shown to play a role in the G2/M checkpoint (Davy et al., 2006). By preventing nuclear accumulation of these CDK-cyclin complexes, HPV16 E1^E4, prevents entry into mitosis, inducing a G2 arrest.

In addition to the ability of E1^E4 to induce a G2 arrest, the viral protein has also been shown to inhibit entry into S phase. Co-expression of both the full length HPV1 E1^E4 protein and an N-terminally truncated E4 species within keratinocytes prevented entry of these cells into S phase (Knight et al., 2004). Furthermore, the HPV1 E1^E4 protein has been shown to inhibit cellular DNA replication by blocking the recruitment of cellular licensing factors onto chromatin (Roberts et al., 2008). It is proposed that by inhibiting cellular DNA replication, E1^E4 generates an environment in which viral genomes can be replicated in the absence of competition from the host for the replicative machinery.

1.12.3   ND10 reorganisation

ND10 bodies constitute interchromosomal accumulations of proteins of which the promyelocytic leukemia protein (PML) is an integral component (Maul et al., 2000). Studies in HPV1 have demonstrated an association between E1^E4 and the PML protein, which results in redistribution of PML from ND10 bodies to the periphery of intranuclear E1^E4 inclusions during the productive phase of the life cycle (Roberts et al., 2003). Upon viral infection ND10 domains may segregate highly expressed viral proteins to these nuclear domains and it is hypothesised that the association of E1^E4 with PML is a strategy adopted by the virus to counteract this negative effect. It has also been proposed that ND10 domain disruption may allow the formation of HPV replication centres at which viral genome amplification can occur (Swindle et al., 1999), and viral genomes are directed to these domains by the minor capsid protein L2 (Day et al., 1998). E1^E4 mediated disruption of ND10 may generate an environment suitable for replication during the productive phase of the HPV life cycle thereby facilitating viral genome amplification prior to packaging into virions (Roberts et al., 2003).

1.12.4   Regulation of gene expression

Several lines of evidence indicate that E1^E4 may function during the later stages of the virus life cycle to regulate gene expression. HPV16 E1^E4 of has been shown to bind via its C-terminus to the E4-DEAD-box protein (E4-DBP), a member of the DEAD-box protein of RNA helicases (Doorbar et al., 2000). This family of RNA helicases display a number of functions linked to post-transcriptional regulation including the ability to regulate gene expression by modulating splicing, RNA turnover, ribosome biogenesis and mRNA export (Iost & Dreyfus, 1994; Fuller-Pace, 1994; Fuller-Pace, 2006). E4-DBP has been shown to associate with HPV16 RNA, including the late polycistronic transcript: E1^E4.L1. The

interaction between E4-DBP and E1^E4 has been shown to partially inhibit the RNA-independent ATPase activity of E4-DBP. Synthesis of HPV capsid proteins is regulated at the post-transcriptional levels of splicing, polyadenylation, mRNA stability and translation (Graham, 2008; Cumming et al., 2009) therefore the interaction with E4-DBP may represent a potential mechanism by which E1^E4 could regulate capsid protein synthesis (Doorbar et al., 2000).

A number of E1^E4 proteins of both low and high risk types have been shown to associate with the serine-arginine (SR) – specific protein kinase-1 (SRPK1) which plays a role in regulation of RNA splicing (Bell et al., 2007). This interaction results in the sequestration of SRPK1 to HPV1 E1^E4 inclusion bodies and results in the phosphorylation of this E1^E4 protein although phosphorylation of E1^E4 proteins of high risk types has not been observed. SRPK1 phosphorylates arginine-serine rich (RS) domains within SR substrate proteins many of which are involved in pre-mRNA splicing, mRNA transport and mRNA translation and it is hypothesised that the association of E1^E4 with SRPK1 may alter the regulation of SR proteins (Bell et al., 2007). Altered activity of SR proteins by E1^E4 may represent a mechanism by which processing of viral transcripts and the expression of viral genes could be regulated (Bell et al., 2007).

1.12.5 The role of E1^E4 during the papillomavirus life cycle

Whilst over-expression studies are useful systems for analysis of E1^E4 protein function, they do not recapitulate the virus replication cycle and therefore do not mimic the physiological environment. Since the advent of expression systems which support propagation of papillomaviruses, studies have been able to examine the role of the E1^E4 protein during the

virus replication cycle within differentiating keratinocytes (described further within Chapter 6).

1.12.5.1 Early stages of the virus life cycle

Whilst E1^E4 proteins are frequently undetectable, or present at low levels within undifferentiated cells, there is evidence to suggest that the E1^E4 protein contributes towards early stages of the viral life cycle. The first study of this kind examined the function of the E1^E4 protein utilising the cottontail rabbit papillomavirus (CRPV) animal model whereby following intraepithelial delivery of CRPV genomes, papilloma growth is induced within rabbits (Peh et al., 2004). The number of papillomas induced in rabbits by mutant CRPV genomes in which E1^E4 expression was abrogated, was considerably higher than that of wild type genomes, indicating that the E1^E4 protein may negatively regulate proliferation (Peh et al., 2004). Further evidence supporting a role for E1^E4 in cellular proliferation during early phases of the virus life cycle came from a study of HPV18 E1^E4 (Wilson et al., 2007). HFK cell lines containing HPV18 genomes consistently exhibited slower rates of growth than those containing mutant genomes in which E1^E4 expression was abrogated, indicating that the presence of HPV18 E1^E4 within these cells has an adverse effect on cellular proliferation.

A role for the E1^E4 protein in replication of viral genomes during early stages of the life cycle has also been proposed (Nakahara et al., 2005). The ability of HPV16 genomes to replicate extrachromosomally following transfection into immortalised keratinocyte NIKS cells was compromised by introduction of a stop codon resulting in a severe truncation of the E4 ORF, with expression of only the first 9 amino acids of the protein. This study indicates that the E1^E4 protein may play an important role during viral genome replication. An intact leucine-rich motif within the HPV16 E1^E4 protein was shown to be required for effective

plasmid DNA replication (Nakahara et al., 2005). However the full length E1^E4 protein has been shown to be dispensable for this viral function in other HPV types including 11, 18 and 31 indicating that type-specific differences are present between E1^E4 proteins (Wilson et al., 2005; Fang et al., 2006a; Wilson et al., 2007). It is possible however that these severely C-terminally truncated E1^E4 proteins may still be active within these cells.

1.12.5.2 Differentiation-dependent stages of the virus life cycle

Expression of E1^E4 is coincident with the onset of viral genome amplification (Peh et al., 2002) and a role in viral genome amplification has been assigned to the E1^E4 protein based on the findings of several studies, which analysed E1^E4 in the context of complete viral genomes. *In-situ* hybridisation analysis of lesions formed in rabbits with a CRPV DNA probe demonstrated that viral genome amplification was abrogated within lesions induced by E4 mutant genomes, and the L1 capsid protein was also undetectable, demonstrating that expression of E1^E4 is crucial for the later stages of the CRPV life cycle (Peh et al., 2004).

The E1^E4 protein of HPV16 has also been shown to contribute towards viral DNA amplification, demonstrated following transfection of NIKS cells with mutant HPV16 genomes in which the E4 ORF was truncated (Nakahara et al., 2005). Transfected NIKS cells were grown in organotypic raft cultures to induce differentiation and were analysed by *in-situ* hybridisation with HPV16 DNA probes. Whilst the presence of amplified viral genomes as detected within differentiated cells of wild type HPV16 raft cultures, the ability of cells to undergo viral genome amplification was significantly reduced by mutant genomes in which the E4 ORF was severely truncated or in which the leucine-rich motif of E4 was mutated, most likely due to the inability of these mutants to replicate during the early stages of the life cycle. Truncations of the E4 ORF preventing expression of the C-terminal third of the E1^E4

protein, resulted in a reduction of viral genome amplification, however more severe C-terminal truncations had a less adverse effect and increased amplification of the viral genome. It is hypothesised that this intriguing, and apparently contradictory observation, is the result of the smaller E1^E4 truncated proteins encoding gain of function mutations. In addition to the dependence upon E1^E4 expression for efficient amplification of viral genomes, the HPV16 E1^E4 protein is also required for enhancing host cell DNA synthesis within suprabasal cells, demonstrated by the reduced ability of HPV16 E4 mutant genome containing raft cultures to support suprabasal DNA synthesis. However, only those mutations which were unable to support viral DNA replication during the early stages of the virus life cycle, were reduced to levels comparable within HPV-negative raft cultures. It is hypothesised that by contributing towards the viral function of enabling terminally differentiated cells to continue undergoing suprabasal DNA replication, the E1^E4 protein enables efficient amplification of the viral genome (Nakahara et al., 2005).

The E1^E4 proteins of HPV18 and HPV31 also play a significant role in differentiation-dependent viral genome amplification, demonstrated by an impaired ability of HFK cells containing mutant genomes in which the E4 ORF is severely truncated, to undergo viral genome amplification upon differentiation (Wilson et al., 2005; Wilson et al., 2007). Following differentiation of these HFKs by suspension in semi solid medium, the ability of cells to undergo viral genome amplification was assessed by Southern blot analysis, demonstrating that in cells unable to express the full-length E1^E4 protein the amplification of viral genomes is significantly reduced. A reduction in viral genome amplification was also coupled with a reduction in late viral transcripts. However whilst suprabasal synthesis was reduced within HPV31 E4 mutant organotypic raft cultures (Wilson et al., 2005), consistent with observations from HPV16 (Nakahara et al., 2005), this function was not shared by

HPV18 (Wilson et al., 2007), indicating that type specific functional differences may exist between E1^E4 proteins.

Whilst a number of E1^E4 proteins have been shown to contribute towards viral genome amplification, following expression of E4 mutant HPV11 genomes within a human keratinocyte cell line immortalized by the catalytic subunit of human telomerase (N-Tert), the ability of cells to undergo viral genome amplification was retained (Fang et al., 2006a). Introduction of a stop codon into the E4 ORF resulting in expression of a severely truncated E1^E4 protein did not impede the expression of RNA transcripts or viral genome amplification, indicating that the E1^E4 protein is not required for these viral functions during the HPV11 life cycle in N-Tert cells (Fang et al., 2006a).

1.12.6 Post-translational modifications

Although the coding capacity of the ~8 kb papillomavirus genome is limited, the virus has evolved other ways of expanding its range of functions to enable complex biological activities to be performed. Encoding of the E4 gene within the E2 ORF presents an ingenious method by which the small coding capacity of the virus has been positively exploited to increase viral function without a subsequent increase in genome size. Similarly, post-translational modifications of viral proteins present another mechanism by which the virus can increase its repertoire of activities.

In natural HPV1 infections, sequential proteolytic cleavage at the N-terminus of the full-length 17-kDa E1^E4 protein generates 16, 11 and 10 kDa E4 proteins (Breitburd et al., 1987). Cleavage occurs in a progressive manner as the keratinocyte migrates up through the epidermis of the wart, such that the 17 kDa protein predominates in the parabasal layer and the truncated E4 species become successively more abundant within the upper layers.

Analysis of the function of these truncated proteins has shown that the 16 kDa protein displays a distinct gain of function (Roberts et al., 1994; Knight et al., 2004). Furthermore proteolysis of the full length E1^E4 species results in removal of functional motifs, including residues required for association with ND10 bodies and the leucine-rich motif which has been shown to mediate association with the keratin intermediate filament and with mitochondria (Roberts et al., 1994; Roberts et al., 2003; Raj et al., 2004). Thus the truncated E4 species looses the ability to perform the E1^E4 functions mediated by these motifs. There is also the possibility that the different E4 species cooperate with one another, thereby extending the functions of the protein (Knight et al., 2004). These studies identify a unique role for the truncated E4 species and since these E4 species are produced in natural infections this suggests that E1^E4 proteolysis is likely to be important for the activity of the protein during the HPV life cycle.

E1^E4 proteins are also targets for phosphorylation by several cellular kinases both *in vivo* and *in vitro* including PKA, SRPK1 and MAPK (Grand et al., 1989; Bryan et al., 2000; Bell et al., 2007). Whilst phosphorylation of E1^E4 by these kinase may be a mechanism of altering E1^E4 function, by changing the structure of the protein or localisation within the cell, E1^E4 may also act to alter the function of cellular kinases through interaction with these proteins.

1.13   **Hypothesis and aims**

The E1^E4 protein is proposed to be a major regulator of the HPV life cycle and thus understanding E1^E4-host interactions has the potential to enable identification of novel therapeutic targets for HPV intervention. The multiple and diverse activities associated with E1^E4 suggest that it is likely to be a multi-functional protein. The ability of E1^E4 to perform a multitude of functions at various stages during the HPV infectious cycle is likely to require regulation of the protein to coordinate these activities and it is hypothesised that post-translational modification of E1^E4 is one mechanism that enables this protein to perform diverse functions.   Knowledge of E1^E4 post-translational modifications is therefore necessary to understand the function of this protein during the HPV life cycle. Thus in this thesis I investigate post-translational modification of the E1^E4 protein of the oncogenic virus HPV18, and use a strategy that mimics the natural scenario of the HPV18 life cycle to test the significance of some of these modifications to the role of the E1^E4 protein in viral genome amplification.

The specific objectives of this thesis were to:

1   Investigate whether the HPV18 E1^E4 protein is a target for post-translational modifications of phosphorylation (*Chapter 3*) and proteolysis (*Chapter 5*).

2   Identify the sites of modification within the primary sequence of E1^E4 and to determine whether these modifications impinge upon the behaviour of the protein using *in vitro* assays and an E1^E4 over expression system (*Chapters 4 and 5*).

3   Mutate by site-directed mutagenesis, sites of modification within the E4 open reading frame within intact HPV18 genomes, and introduce into primary foreskin keratinocytes   (*Chapter 6*).

4   Use these HPV18 genome containing cell lines to assess the effect of these mutations on E1^E4 function during the HPV18 virus life cycle (*Chapter 6*).

CHAPTER 2

MATERIALS AND METHODS

2.1 Bacterial cell culture

2.1.1 Bacterial hosts, growth and storage

The DH5α strain of *Eschericia coli* (*E. coli*) was used as a bacterial host for growth of pcDNA 3.0 and pGEMII plasmid vectors for preparation of plasmid DNA. The *E. coli* strain BL21 (Stratagene) were used for growth of the pGEX-3X plasmid vector for expression of recombinant GST-fusion proteins. Bacteria were streaked and stored on agar plates for a short period of time, or for longer periods, stored as a glycerol stock.

Agar plates were made by dilution of 1.5% [wt/vol] agar in Luria-Bertani (LB) medium (1% [wt/vol] bacto-tryptone, 0.5% [wt/vol] bacto-yeast extract, 1% [wt/vol] NaCl). Following autoclaving, the agar was boiled then cooled to approximately 50°C, before the appropriate antibiotic was added to a final concentration of 50 μg/ml and 25 ml of agar dispensed into each Petri dish. Bacteria were streaked out from glycerol stocks by streaking of bacteria across the surface of an agar plate and plates were then incubated overnight at 37°C to allow growth of bacterial colonies.

Glycerol stocks were made following growth of bacteria in LB media supplemented with 50 μg/ml antibiotic. Bacteria were grown in LB media overnight at 37°C with shaking before a 400 μl aliquot of bacteria was taken and added to 600 μl of a sterile 80% [vol/vol] glycerol solution. Glycerol stocks were stored in 1 ml aliquots in cryovials at -80°C.

2.1.2 Preparation of chemically-competent E. coli

E. coli bacteria were streaked out from a glycerol stock onto a LB plate without selection antibiotic. A single bacterial colony was used to inoculate a 5 ml starter culture of LB media, incubated in a 15 ml polypropylene tube at 37°C with shaking overnight. The following day

the starter culture was used to inoculate 200 ml of LB medium, grown in a 2 L conical flask at 37˚C until the OD of the culture reached 0.4-0.6. Following centrifugation at 3,834 g for 15 min at 4˚C, the bacterial pellet was re-suspended in 40 ml of ice-cold transformation buffer TFB1: (30 mM potassium acetate, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15% [vol/vol] glycerol). The culture then underwent centrifugation at 3,834 g for 15 min at 4˚C prior to re-suspension of the pellet in 8 ml pre-cooled and sterile buffer: TFB2 (10 mM Na-MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl and 15% [vol/vol] glycerol). 100 µl aliquots were then added to pre-cooled cryovials before snap-freezing in liquid nitrogen and storage at -80˚C.

2.1.3   Transformation of competent *E. coli* with plasmid DNA

Chemically competent *E. coli* were first thawed on ice, prior to addition of 30-50 µl of cells into a pre-cooled Falcon tube. 1-2 µg of plasmid DNA was typically added directly into the thawed cells whilst for transformation of a ligation reaction 10 µl was used. Following incubation on ice for 30 min, *E. coli* were heat shocked at 42˚C for 45 sec, followed by a 2 min incubation on ice. 500 µl of LB media was subsequently added, followed by incubation at 37˚C with shaking for 1 h. 250 µl was then plated onto ampicillin selection agar plates and incubated overnight.

2.1.4   Bacterial expression of glutathione-S-transferase fusion proteins

For expression of GST-fusion proteins, BL21 cells containing pGEX-3X recombinant plasmids were grown for 16 h at 37°C with shaking in 50 ml LB media with 1% [wt/vol] glucose and 50 µg/ml ampicillin. This culture was then diluted into 200 ml of LB media and incubated at 37°C for 2 h, and expression induced for 3 h by addition of 50 µM isopropyl-ß-D-thiogalactopyranoside (IPTG, Sigma-Aldrich) and growth at 30°C. Bacterial cells were

pelleted at 3,834 g for 5 min at 4°C and re-suspended in 8 ml of ice-cold bacterial lysis (BL)

buffer (1X phosphate-buffered saline (PBS), 1% Triton X-100 (vol/vol), supplemented with

Complete™ protease inhibitor cocktail (Roche)). After brief sonication (2 periods of 30 s), at

40 watts using a Ultrasonic Processor (Model W380, Heat Systems, Ultrasonics), insoluble

material was removed by centrifugation at 13,186 g for 10 min at 4˚C and soluble GST

proteins were immobilised on glutathione *S*-agarose beads (Sigma-Aldrich) by rotation at 4°C

for 1 h. The beads were washed 5 times with BL buffer and GST proteins eluted into BL

buffer containing 50 mM reduced glutathione (Sigma). Soluble proteins were dialysed using

mini Geba Flex tube dialysis tubes (Gene-Bio Application Ltd) following manufacturer's

instructions, at 4°C in dialysis buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 2 mM DTT

and Complete™ protease inhibitor cocktail) for 16 h, prior to storage at -80˚C. GST fusion

proteins were quantitated by Coomassie staining following separation of proteins by SDS-

polyacrylamide gel electrophoresis (SDS-PAGE) with a known quantity of bovine serum

albumin (BSA) run as a protein standard (as described in section 2.6).

2.1.5   Isolation and purification of plasmid DNA from bacteria

2.1.5.1   Large scale preparation of plasmid DNA (Maxi-preparation)

A single colony of bacteria transformed with the desired plasmid was picked from an agar

plate and grown for 6-8 h at 37˚C with shaking in 5 ml LB media with appropriate selection

antibiotic. This culture was subsequently diluted into 200 ml LB media and grown for 16 h at

37˚C before centrifugation at 3,834 g for 15 min at 4˚C to pellet the bacteria. DNA was then

extracted from the bacteria using the HiSpeed® Plasmid Maxi Kit (Qiagen), following

manufacturer's instructions. DNA was eluted into 1 X Tris-EDTA (TE) buffer (Promega, 10

mM Tris-HCl, 1 mM EDTA pH 8.0) and stored at -20˚C. The concentration of DNA was determined with a Nanodrop™ spectrophotometer (Thermo Scientific).

2.1.5.2   Small scale preparation of plasmid DNA (Mini-preparation)

A single bacterial colony was grown in 5-6 ml LB media with selection antibiotic at 37˚C for 16 h. 3 ml of bacteria were subsequently pelleted by centrifugation at 3427 g at 4˚C for 5 min followed by DNA extraction using the QIAprep® Spin Miniprep Kit (Qiagen), following manufacturer's instructions. DNA was eluted into 30 µl of 1 X TE buffer and stored at -20˚C.

2.2   Molecular cloning

2.2.1   Plasmid DNA vectors

The pGEMII-HPV18 plasmid (a gift from Frank Stubenrauch, University of Tuebingen and used in a previous study (Wilson et al., 2007)) contains the total HPV18 genome (accession number: NC 001357) cloned into the pGEMII vector at the *Eco*RI restriction site at residue 2440.

The pGEX-3X-HPV18 E1^E4 plasmid was generated by cloning of HPV18 E1^E4 into the pGEX3X expression vector (Amersham) by Michelle McNally (Bell et al., 2007).

The pcDNA 3.1 codon-optimised E1^E4 plasmid was generated by Dr Ian Bell, by converting a number of codons within HPV18 E1^E4 cDNA (coE1^E4) to those more commonly used in genes of the human genome (Appendix III).

2.2.2   Polymerase chain reaction (PCR)

PCR reactions were typically performed with 1 µg of plasmid DNA using the Expand High Fidelity PCR system (Roche), following manufacturer's instructions. PCR reactions were

carried out in a total volume of 100 µl in expand high fidelity buffer supplemented with 15 mM MgCl$_2$, 0.2 mM dNTPs, 5 ng/µl forward and reverse primers, 1% [vol/vol] DMSO with 3.5 U of Taq DNA polymerase. PCR cycles were performed on a thermal cycler (2720, Applied Biosciences) with a 3 min 94˚C hot start followed by 25 cycles (1 min at 94°C, 1 min at 50°C and 2 min at 72°C) before cooling at 4˚C.

2.2.3 Agarose gel electrophoresis

For resolution of plasmid DNA, 1% agarose gels were routinely used which comprised of: 1% [wt/vol] agarose in 1 x Tris/Boric Acid/EDTA (TBE) running buffer (90 mM Tris, 90 mM Boric Acid, 2 mM EDTA) supplemented with 0.1% [vol/vol] ethidium bromide (EtBr). Solutions were dissolved by boiling, and following cooling to approximately 70˚C, agarose gels were cast in a Mini Sub Cell GT electrophoresis tank (Bio-Rad). DNA samples were mixed with 6 x loading buffer (30% [vol/vol] glycerol, 0.3% [wt/vol] bromophenol blue in 10 x TBE) in a total volume of 10-20 µl, and loaded into wells of the agarose gel. 1 kb plus DNA ladder (Invitrogen) was run on each gel for size determination. Gels were routinely run at 25 mA for 1 h in TBE running buffer and DNA bands were visualised with a Gene Flash UV light box (Syngene Bio Imaging). For resolution of genomic DNA isolated from HFK prior to Southern blot analysis 0.8% [wt/vol] agarose gels were prepared as described above and cast in Fisherbrand horizontal gel electrophoresis tanks (Fisher Scientific), with gels run at 50 V overnight.

2.2.4 DNA purification

For isolation of products from PCR amplification reactions or from restriction enzyme digestions, DNA was purified using the High Pure PCR Product Purification Kit (Roche

Diagnostics) following manufacturer's instructions. DNA was eluted in 50 μl 1 X TE buffer and stored at -20˚C.

For extraction of DNA from agarose gels, the DNA product was run on an agarose gel and the corresponding band was cut out with a sterile scalpel. The DNA was then extracted using the QIAquick Gel Extraction Kit (Qiagen) and eluted into 30 μl 1 X TE buffer and stored at -20˚C.

2.2.5 Restriction enzyme digestion

Restriction enzyme digests were typically performed using 2-10 μg of plasmid DNA incubated with restriction enzymes as follows in the appropriate buffer provided with the enzyme in a total volume of 20-50 μl. Restriction enzyme used in this study include *Eco*RI (20 units), *Bgl*II (10 units), *Bam*HI (20 units) (New England Biolabs) and *Dpn*I (10 units) (Roche). Following addition of the restriction enzyme, reactions were vortexed and centrifuged briefly, before incubation at 37˚C for 1-2 h or overnight. To ensure complete digestion, a 200 ng aliquot was analysed on an agarose gel. For digestion with multiple restriction enzymes these were carried out simultaneously by selecting an appropriate buffer in which both restriction enzymes are active.

2.2.6 Alkaline phosphatase treatment of plasmid DNA

Prior to ligations, vector DNA was treated with alkaline phosphatase to de-phosphorylate the 5' termini of the DNA and prevent self-ligation. Following restriction digest of vector DNA and PCR purification, alkaline phosphatase treatment was performed by incubation of digested plasmid DNA with 7 units of calf-intestinal alkaline phosphatase (New England

Biolabs) in the de-phosphorylation buffer provided with the enzyme. Reactions were incubated for 30 min at 37˚C prior to addition of 1/10 volume of 200 mM EGTA to terminate the ligation reaction, with incubation at 65°C for 10 min.

2.2.7 DNA ligation reactions

Following restriction digests, enzymes were heat-inactivated prior to a ligation reaction by incubation of reactions at 65˚C for 10 min, then cooling on ice for 2 min. Ligations were performed by incubation of vector and insert DNA at ratios typically of 1:3, 1:5 or 1:10 with 4-8 units T4 DNA ligase (New England Biolabs) in the buffer provided and incubated overnight at 16˚C. Ligated DNA was then transformed into bacteria and sequenced following mini-preparation of plasmid DNA (as described in section 2.1.5.2).

2.2.8 Cloning of DNA into plasmid vectors

2.2.8.1 Generation of a pGEX-3X HPV18 E7 expression vector

The HPV18 E7 coding sequence was amplified by PCR from the pGEMII-HPV18 vector with the following primers: 5' primer 5' GGG ATC CCC CAT GGA CCT AAG GCA ACA TTG C 3' and 3' primer 5' TGA ATT CCC TTA CTG CTG GGA TGC ACA CCA CGG 3'. The PCR product was gel extracted, digested with *Bam*HI and *Eco*RI and subsequently ligated into a pGEX-3X vector prepared with the same restriction endonucleases. Ligated DNA was then transformed into bacteria and sequenced with the pGEX promoter primer.

2.2.8.2 Cloning of E1^E4 cDNA into the pAP16 vector

E1^E4 cDNA underwent PCR amplification from the pMC3-HPV18 E1^E4 vector using the following primers: 5' primer 5'-GCGCAGATCTCCACCATGGCTGATCCAGAAG-3' and

3' primer 5'-CGCGAGATCTTTATAGGCGTAGTG-3' and the PCR product was subsequently gel extracted. The SV40 early replacement vector pAP16 (Roberts et al., 1993) was linearised with *Bgl*II, the DNA PCR-purified and subsequently treated with alkaline phosphatase. The purified E1^E4 PCR product was subsequently digested with *Bgl*II and ligated into the prepared pAP16 vector. Following transfection into bacteria, the vector was sequenced using the SV40 promoter primer as described previously, to ensure correct orientation of the E1^E4 sequence within the vector.

2.2.9   PCR sequencing

1 µg of plasmid DNA prepared by the maxi-preparation method (section 2.1.5.1) or alternatively 10 µl of DNA prepared by the mini-preparation method (section 2.1.5.2) were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) according to manufacturer's instructions. 25 cycles of 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min were performed on a thermal cycler. Oligonucleotide primers utilised for sequencing of plasmid DNA are listed in Table 1, and were generated by Altabiosciences, University of Birmingham. The DNA was subsequently precipitated by addition of 3 volumes of ethanol and sodium acetate to a final concentration of 0.3 M and incubated at room temperature for 20 min. Following centrifugation at 16,100g for 30 min at 4˚C, the supernatant was removed and the DNA pellet air dried for 20-30 min and re-suspended in 10µl Hi-dye (Applied Biosystems). The reaction was heated at 95˚C for 5 min before loading into a 96 well plate on 3100 ABI Prism™ DNA capillary sequencer (Applied Biosystems). The sequencer capillary array was filled with 3100 Performance Optimised Polymer 6™ Performance Optimized Polymer (Applied Biosystems). Sequencing data was collected using

the 3100 data collection software version 1.0 and was analysed using ABI Sequencer version 3.6.1.

Table 1: Sequencing primers

Primer	Direction	Primer Sequence
T7 Promoter	Forward (5')	5' TAA TAC GAC TCA CTA TAG GG 3'
BGH	Reverse (3')	5' CTG TGA ATG CTG TGG CAG GGG 3'
SV40 Promoter	Forward (5')	5' CCG CCC ATT CTC CGC CCC ATG GC 3'
SV40 Reverse	Reverse (3')	5' GCC ATG GGG CGG AGA ATG GGC GG 3'
pGEX	Forward (5')	5 'GGG CTG GCA AGC CAC GTT TGG TG 3'

2.2.10 Site-directed mutagenesis

Site-directed mutagenesis was performed using the QuikChange Kit (Stratagene) following manufacturer's instructions. The pCDNA-HPV18 codon-optimised E1^E4 plasmid vector or the pGEX3X-HPV18 E1^E4 expression vector were used as template DNA using the primers listed in Tables 1 and 2 respectively, with the number of PCR cycles used between 18 and 25. All of the oligodeoxyribonucleotide site-directed primers utilised in this study were generated by Altabiosciences, University of Birmingham. PCR products were digested with *Dpn*I and transformed into XL-1 blue *E. coli*. Colonies were grown in LB media and mini-preparation performed (as described in section 2.1.5.2). Extracted DNA was sequenced to confirm correct introduction of mutants.

2.2.10.1   Generation of mutations within the codon-optimised E1^E4 sequence

Deletions within the E1^E4 coding sequence were introduced by site-directed mutagenesis using the primer pairs in Table 2, with the pcDNA3.1 codon-optimised E1^E4 vector as template DNA. Bi-directional sequencing with a T7 promoter and a BGH reverse primer were used to verify the mutations within the E1^E4 cDNA.

Table 2:   Mutagenic primers for the introduction of deletion and substitution mutations into the pcDNA 3.1 codon-optimised E1^E4 expression vector

Mutant	Direction	Primer Sequence
ΔYP	Forward (5')	5' CCC GTG ACA ACC AGG CTG CTC AGC CTG CTC 3'
	Reverse (3')	5' GAG CAG GCT GAG CAG CCT GGT TGT CAC GGG 3'
ΔLLXLL	Forward (5')	5' CCG TGA CAA CCA GGT ACC CTA ACA GCT ACA GCA CCC CAC C 3'
	Reverse (3')	5' GGT GGG GTG CTG TAG CTG TTA GGG TAC CTG GTT GTC ACG G 3'
ΔLNS	Forward (5')	5' GGT ACC CTC TGC TCA GCC TGT ACA GCA CCC CAC CTC AC 3'
	Reverse (3')	5' GTG AGG TGG GGT GCT GTA CAG GCT GAG CAG AGG GTA CC 3'
ΔNSY	Forward (5')	5' CTC TGC TCA GCC TGC TCA GCA CCC CAC CTC ACA G 3'
	Reverse (3')	5' CTG TGA GGT GGG GTG CTG AGC AGG CTG AGC AGA G 3'
ΔSYS	Forward (5')	5' CTG CTC AGC CTG CTC AAC ACC CCA CCT CAC AGG 3'
	Reverse (3')	5' GAT CCT GTG AGG TGG GGT GTT GAG CAG GAG CAG 3'

ΔYST	Forward (5')	5' CTC AGC CTG CTC AAC AGC CCA CCT CAC AGG ATC CCC 3'
	Reverse (3')	5' GGG GAT CCT GTG AGG TGG GCT GTT GAG CAG GCT GAG 3'
ΔSTP	Forward (5')	5' GCC TGC TCA ACA GCT ACC CTC ACA GGA TCC CCG C 3'
	Reverse (3')	5' GCG GGG ATC CTG TGA GGG TAG CTG TTG AGC AGG C 3'
ΔTPP	Forward (5')	5' TGC TCA ACA GCT ACA GCC ACA GGA TCC CCG CCC C 3'
	Reverse (3')	5' GGG GCG GGG ATC CTG TGG CTG TAG CTG TTG AGC A 3'
L^{14}L^{15}>PP	Forward (5')	5' GAC AAC CAG GTA CCC TCC ACC CAG CCT GCT CAA CAG C 3'
	Reverse (3')	3' GCT GTT GAG CAG GCT GGG TGG AGG GTA CCT GGT TGT C 3'
L^{14}L^{15}>RR	Forward (5')	5' GAC AAC CAG GTA CCC TCG ACG CAG CCT GCT CAA CAG C 3'
	Reverse (3')	3' GCT GTT GAG CAG GCT GCG TCG AGG GTA CCT GGT TGT C 3'
L^{17}L^{18}>SS	Forward 5'	5' CCA GGT ACC CTC TGC TCA GCT CGT CAA ACA GCT ACA GCA CCC CAC C 3'
	Reverse 3'	5' GGT GGG GTG CTG TAG CTG TTT GAC GAG CTG AGC AGA GGG TAC CTG G 3'
T^{23}>A	Forward 5'	5' CAA CAG CTA CAG CGC CCC ACC TCA CAG 3'
	Reverse 3'	5' CTG TGA GGT GGG GCG CTG TAG CTG TTG 3'
T^{23}>D	Forward 5'	5' GCT CAA CAG CTA CAG CGA CCC ACC TCA CAG GAT C 3'
	Reverse 3'	5' GAT CCT GTG AGG TGG GTC GCT GTA GCT GTT GAG C 3'
S^{58}>A	Forward 5'	5' GCA GGA GGA GCG CCA TCG TGG ACC 3'
	Reverse 3'	5' GGT CCA CGA TGG CGC TCC TCC TGC 3'
S^{58}>D	Forward 5'	5' GCA GGA GGA GCG ACA TCG TGG ACC 3'
	Reverse 3'	5' GGT CCA CGA TGT CGC TCC TCC TGC 3'

2.2.10.2   Generation of mutations within the pGEX-3X E1^E4 sequence

Substitution mutations within the E1^E4 coding sequence were introduced by site-directed mutagenesis using the primer pairs in Table 3, with the pGEX-3X E1^E4 vector as template DNA. Sequencing with a pGEX promoter primer were used to verify the mutations within the E1^E4 sequence.

Table 3:   Mutagenic primers for the introduction of mutations into the

              pGEX-3X E1^E4 expression vector

Mutant	Direction	Primer Sequence
$S^{57}S^{58}$>NN	Forward (5')	5' GGA CTC GCG GAG AAA CAA CAT TGT GGA CCT G 3'
	Reverse (3')	5'CAG GTC CAC AAT GTT GTT TCT CCG CGA GTC C 3'
S^{57}>N	Forward (5')	5' GGA CTC GCG GAG AAA CAG CAT TGT GGA CCT G 3'
	Reverse (3')	5' CAG GTC CAC AAT GCT GTT TCT CCG CGA GTC C 3'
S^{58}>N	Forward (5')	5' GGA CTC GCG GAG AAG CAA CAT TGT GGA CCT G 3'
	Reverse (3')	5' CAG GTC CAC AAT GTT GCT TCT CCG CGA GTC C 3'

2.2.10.3   Generation of an N-terminally HA tagged E1^E4 expression vector

An haemagglutinin (HA) epitope tag was introduced at the N-terminus of the cDNA sequence of E1^E4 within the HPV18 coE1^E4 expression construct to generate an N-HA E1^E4 plasmid. The HA tag was introduced by site-directed mutagenesis using the following primer pairs: 5' primer 5' CCC ATA CGA TGT TCC AGA TTA CGC TGC CGA CCC CGA GGT GCC C 3' and 3' primer 5' GGG CAC CTC GGG GTC GGC AGC GTA ATC TGG AAC

ATC GTA TGG G 3'. The vector was sequenced with the T7 forward promoter and the BGH

polyadenylation sequence reverse primer as described below, to verify introduction of the HA

tag.

2.3   Maintenance of cell lines

2.3.1   Cell lines and growth medium

Table 4:   Cell lines used in this study

Cell Line	Cell Type
911	Adenovirus 5 E1A-transformed epithelial cell line
293T	Human embryonic kidney cell line containing the SV40 Large-T antigen
SAOS-2	Human cell line derived from an osteosarcoma
COS-1	African green monkey kidney epithelial cell line
HeLa	Derived from a cervical adenocarcinoma and contain chromosomally integrated HPV18 genomes
SVJD	Simian virus 40 [SV40]-immortalised human epidermal keratinocytes
SCC12F	Human squamous cell carcinoma

Cell lines: 911, 293T, SAOS-2, COS-1 and HeLa were all cultured in Dulbecco's modified eagles medium (DMEM) HEPES modification (Sigma) supplemented with 2mM L-glutamine and 10% [vol/vol] fetal bovine serum (FBS). SVJD keratinocytes were grown in Jokliks media supplemented with 10% [vol/vol] FBS and 2mM L-glutamine. SCC-12F keratinocytes were grown in DMEM HEPES modification / Ham's F12 (3:1) supplemented with 5% [vol/vol] FBS, 2mM L-glutamine and 0.4 µg /ml hydrocortisone. All cells were maintained in an incubator at 37˚C with 5% CO_2 and were grown routinely in 100 mm dishes (Iwaki) or for large scale transfections in T150 flasks (Iwaki).

2.3.2 Passaging of cells

Cells were harvested by first removing media, before washing the cells twice with PBS and addition of 1 ml 0.25% Trypsin with 1 mM EDTA (Gibco-Invitrogen), followed by incubation at 37˚C. Cells were checked frequently under a microscope to monitor cellular adhesion, and dishes were tapped gently to aid release of cells. Cells were subsequently re-suspended in 10 ml appropriate media before transferral to a 50 ml conical tube. The dish was then washed with 5 ml media and the cells centrifuged at 538 g at room temperature for 5 min. The supernatant was then removed and the cells re-suspended in 10 ml media. Viable cells were then counted using a cell counter following mixing of an equal volume of 0.4% Trypan Blue to a sample of the cell suspension and incubation at room temperature for 5 mins.

2.3.3 Freezing and thawing of cells

Cells were grown to a confluency of between 70-80% prior to freezing and were harvested before re-suspension of cells at 2×10^6 / ml in appropriate media supplemented with 10% [vol/vol] DMSO. Cells were subsequently transferred in 1 ml aliquots into Nunc cryovials

(Nalgene-Nunc, Thermo Fischer Scientific) and stored at -80˚C in a slow-cool freezing chamber with isopropanol (Nalgene® Mr Frosty, Nalgene-Nunc) for 16 h prior to long term storage in liquid nitrogen.

To defrost cells, cryovials containing cells were removed from liquid nitrogen storage and thawed at 37˚C in a water bath. Cells were transferred with a transfer pipette into 10 ml pre-warmed media in a drop-wise manner before centrifugation at 538 g for 5 min at room temperature. The cells were then re-suspended in appropriate media and grown in 100 mm dish culture dishes.

2.3.4   Transfection of cells with plasmid DNA

Transfections were performed by incubation of Lipofectamine™ 2000 (Invitrogen) transfection reagent (2µl/µg DNA to be transfected) with 200µl of Optimem® reduced serum media (Invitrogen) for 5 min at room temperature. In a separate tube 5-10 µg of plasmid DNA / 100 mm dish or 15 µg / T150 flask of cells was typically incubated with 200 µl of Optimem®. Following 5 min incubation at room temperature the Lipofectamine / Optimem® solution was dispensed directly into the DNA / Optimem® solution and incubated at room temperature for 30 min. Cells were grown to 80% confluency prior to transfection and following removal of media and 2 x washes with 1 X PBS, cells were incubated with the 400 µl transfection mixture at 37˚C for 4-6 h. Dishes were rocked every 1 – 2 h to ensure even coverage of the transfection mixture and were then grown in appropriate media at 37˚C.

2.3.5   Treatment of cells with protease inhibitors

All protease inhibitors used in the study were obtained from Calbiochem. 10 X stock solutions of protease inhibitors were made up in DMSO and were added to the medium 2 hrs

prior to transfection. Following transfection, cells were grown in media containing the appropriate inhibitor until harvested. The following inhibitors were used at the indicated final concentrations E64 (500 μM), AEBSF (100 μM), Calpeptin (1 μM, 0.5 μM), ALLM (1 μM, 100 nM) and ALLN (1 μM, 100 nM).

2.3.6 Infection of cells with recombinant virus

Cells were grown to 70% confluency prior to infection. Media was first removed and cells washed with 10 ml PBS before virus was added with a multiplicity of infection of 100 in a total volume of 200 μl with titres routinely of ~ 5-10 X 10^{11} pfu/ml. Following 2 h incubation at 37˚C, with regular rocking of dishes to ensure even coverage of the virus, appropriate media was replaced and the cells were then grown at 37˚C until harvested, 36-48 h post-infection.

2.4 Generation of a recombinant SV40 E1^E4 virus

2.4.1 Preparation of pAP16-HPV18 E1^E4 for electroporation

Generation of the pAP16-HPV18 E1^E4 vector was previously described in section 2.2.8.2. 25 μg of this plasmid DNA was digested with *Eco*RI (as described in section 2.2.5) and heat inactivated by incubation at 65˚C for 10 min. The digest was subsequently divided into two ligation reaction with total volumes of 500 μl. Following a ligation reaction (as described in 2.2.7), a chloroform extraction was performed by addition of 500 μl chloroform prior to centrifugation at 16,100 g for 5 min. The upper aqueous layer was then precipitated with 2.5 volumes of ethanol and sodium acetate to a final concentration of 300 mM and incubated at -20˚C for 6 h. Following centrifugation at 16,100 g for 20 min at 4˚C, the supernatant was

removed and the DNA pellet washed with 70% [vol/vol] ethanol before a further 10 min centrifugation at 16,100 g. The ethanol was then removed, and the pellet air dried briefly, before re-suspension in 25 μl sterile water.

2.4.2 Electroporation into COS-1 cells and harvesting of SV40 recombinant virus

COS-1 cells at a confluency of 80% were harvested with trypsin, washed in PBS and re-suspended at 2.5 x 10^6/ml in DMEM supplemented with FBS and L-glutamine. An 800μl aliquot of cells was transferred into a cuvette before addition of 15μg of pAP16-HPV18 E4 vector or 15μg pcDNA as a control. Following 10 min incubation on ice, cells were electroporated at 960mF/0.22kV using a Gene Pulser X Cell™ electroporator (Bio-Rad). 1 ml of media was subsequently added drop-wise into the cuvette before transferral of the cells drop-wise with a transfer pipette into 10 ml media. The cells were then divided into two 10 cm tissue culture tissues an incubated at 37˚C.

Following electroporation with pAP16-HPV18 E1^E4, COS-1 cells were grown until a cytopathic effect (CPE) was observed, which would routinely occur after approximately 7 days. Cells were then harvested by scraping of cells into the medium with a cell scraper followed by centrifugation at 538 g for 2 min. The supernatant was then removed and the cell pellet re-suspended with 2 ml medium which was subsequently divided into 2 eppendorf tubes and stored at -80˚C for 20 min. The vial was then thawed at 37˚C, followed by two repeated cycles of freeze/thawing before centrifugation at 136.8 g for 10 min. The supernatant was then aliquoted into vials and frozen at -80˚C as the P_0 passage of the virus. The P_0 virus was then used to infect 20 dishes of 80% COS-1 cells. Following removal of media and washing with PBS, 200 μl of P_0 virus was added to the cells. The dishes were then

incubated at 37°C with regular rocking to ensure even coverage of virus. Following 2 h incubation, media was replaced and the cells were grown at 37°C until CPE was observed. The P_1 virus was harvested in the same manner as the P_0 virus, aliquoted into freezing vials and stored at -80°C until required.

2.5   Immunofluorescence and microscopy

2.5.1   Transfection and infection of cells grown on glass slides

Cells were harvested and re-suspended in appropriate medium as described previously, at a concentration of $5 \times 10^4 – 1 \times 10^5$ cells / 100 μl. Cells were seeded onto spots on glass multi-spot microscope slides (Hendley Essex) in 100 μl aliquots and grown overnight in a humidified incubator at 37°C with 5% CO_2.

Transfections were performed essentially as described in section: 2.3.4, with 250 ng DNA in 100 μl transfection mixture added to each spot following removal of media and washing of each spot with 100 μl Optimem®. Cells were incubated at 37°C for 4-6 h prior to flooding of the dish with 15 ml of appropriate media.

For viral infection, following replacement of media of recombinant virus was added to each spot in a total volume of 10-20 μl. Following incubation at 37°C for 2 h dishes were flooded with 15 ml of appropriate media.

2.5.2   Fixation of cells

Cells were fixed 36-48 h following transfection or infection and were washed briefly in saline, pre-cooled to 4°C, prior to fixation in acetone at -20°C for 10 min. Alternatively, for paraformaldehyde fixation, cells were fixed with 4% [wt/vol] paraformaldehyde in PBS for 5

min, prior to two 5 min washes in PBS before permeabilisation in acetone at -20˚C for 10 min. Slides were then air dried briefly.

2.5.3 Immunodetection

100 µl of blocking buffer (PBS with 20% [vol/vol] heat-inactivated goat serum (HINGS), 0.5% [wt/vol] BSA, 0.5% [wt/vol] sodium azide) was added to each spot and the slides incubated in a humidifying chamber at 37˚C for 1 h. Following removal of blocking buffer, primary antibodies were diluted in blocking buffer and added in 100 µl to each spot. HPV18 E1^E4 was detected either using a mouse monoclonal antibody (MAb), 1D11 at a dilution of 1:5 (Roberts et al., 2003), or using a rabbit polyclonal E1^E4 antibody, R424 raised against a GST-HPV18 E1^E4 fusion protein (Wilson et al., 2007) at a dilution of 1:1000. Cyclin proteins were recognized using mouse MAbs to cyclin A (AT10.3), obtained from Cancer Research UK Research Services, at a working dilution of 1:250 and cyclin B1 (AB3GNS1), obtained from Neomarkers, used at 1:100. Keratin 18 was detected with a mouse MAb anti-cytokeratin K18 antibody (CK5) obtained from Sigma and used at a dilution of 1:1000. Flag-tagged HPV1 E1 was detected with a rabbit anti-Flag antibody (1:1000, Sigma-Aldrich).

Following 2 h incubation in a humidifying chamber at 37˚C, slides were washed twice in PBS for 15 min at room temperature. Slides were dried briefly before addition of secondary antibodies. Anti-rabbit IgG Alexa® 488 or anti-mouse IgG Alexa® 596 conjugates (Molecular probes) secondary antibodies were diluted 1:500 in blocking buffer and 100 µl added to each spot. Following 1 h incubation at 37˚C, slides were washed twice in PBS with addition of 5 µl 4', 6'-diamidino-2-phenylindole (DAPI, Sigma-Aldrich) / 100 ml PBS for the

final wash. After drying, coverslips were mounted onto slides using 1,4-diazabicyclo[2.2.2]octane (DABCO) mounting medium.

2.5.4 Microscopy

Epifluorescence microscopy was performed on a Nikon Eclipse E600 microscope and images were captured using a Nikon DXM1200F digital camera. Confocal microscopy was performed on a LSM 510 META confocal microscope (Zeiss) using multi-track imaging. Images were assembled in Adobe Photoshop® CS2.

2.6 Protein Biochemistry

2.6.1 Cell harvesting and protein extraction

Cells were harvested from dishes with the use of a cell scraper or by trypsin treatment (described in section 2.3.2) and were pelleted by centrifugation at 538 g for 5 min before washing in PBS and re-pelleting prior to protein extraction. Cell lysates were prepared for SDS-polyacrylamide gel (SDS-PAGE) electrophoresis by solubilisation in urea lysis buffer (8M Urea, 25 mM Tris-HCl pH 8.0, 0.15M β–Mercaptoethanol) supplemented with 1X Complete™ protease inhibitors. Following incubation on ice for 20 min, lysate underwent sonication for 20 sec at 40 watts using a Ultrasonic Processor (Model W380, Heat Systems, Ultrasonics) followed by removal of insoluble material by high speed centrifugation at 16,100 g for 20 min at 4˚C. Cell lysates were prepared for immunoprecipitation and co-precipitation experiments by lysis in an NP40 lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% [vol/vol] NP40, supplemented with 1 X Complete™ protease inhibitors) and lysate was cleared by high speed centrifugation.

2.6.2   Protein determination by the Bradford method

Protein determination was performed by the Bradford method using a Bio Photometer (Eppendorf). Protein standards were prepared by dilution of 0, 4, 8, 12, 16 and 20 µg of bovine serum albumin (BSA) in 1 ml Protein Assay reagent (Bio-Rad). A known volume of lysate was diluted in reagent and the concentration determined.

2.6.3   Lambda phosphatase (λ-PPase) treatment

Cellular lysates were incubated with 400units (λ-PPase) (New England Biosciences) in provided 1 x λ-PPase reaction buffer, supplemented with 2 mM $MnCl_2$, at 30°C for 30 min, before the reaction was terminated by addition of Laemelli sample loading buffer (Bio-Rad, 62.5 mM Tris-HCl, pH 6.8, 25% [vol/vol] glycerol, 2% [wt/vol] SDS, 0.01% [wt/vol] Bromophenol Blue) supplemented with 5% [vol/vol] β-Mercaptoethanol.

2.6.4   *In vitro* proteolysis assay

For *in vitro* proteolysis experiments, cell pellets underwent three cycles of freeze/thawing whereby cell pellets were frozen at -80˚C prior to thawing at 4˚C, in repeated cycles. Cell lysate was subsequently incubated at 25 or 37°C and following incubation for various times, the reaction was stopped by addition of Laemelli loading buffer supplemented with 5% [vol/vol] β-Mercaptoethanol.

2.6.5   Immunoprecipitation assays

For immunoprecipitation assays using HA-tagged E1^E4 proteins, cells were lysed in NP40 lysis buffer (as described in section 2.6.1). Cleared lysate was subsequently incubated with 0.5 µg / 100 mm dish or 1 µg / T150 flask of cells, of mouse MAb HA antibody (16B12,

Covance) at 4˚C with rotation. Following 2 h incubation, 50 μl of a 50% slurry of protein G-sepharose (Cancer Research UK) in NP40 lysis buffer was added and a further incubation of 90 min performed. Beads were subsequently washed 5 times in NP40 lysis buffer, before re-suspension in Laemelli loading buffer supplemented with 5% [vol/vol] β-Mercaptoethanol for Western Blot analysis.

2.6.6 Co-precipitation assays

Co-precipitation assays for the detection of CDK-associated activity were performed using cellular lysates prepared in NP40 lysis buffer. Lysates (500 μg) were mixed with 500 ng of GST fusion protein immobilized to glutathione-S-agarose beads (Sigma) at 4°C with rotation for 2 h. Beads were washed three times in the NP40 lysis buffer and twice in kinase reaction buffer containing: 20 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT and Complete™ protease inhibitors, prior to analysis by *in vitro* kinase assay.

2.6.7 *In vitro* kinase assays

Following co-precipitations with GST fusion proteins, *in vitro* phosphorylation assays, were performed in kinase reaction buffer containing: 50 mM ATP, 10 μCi [γ-^{32}P]-ATP (3000 Ci/mMol, Perkin Elmer) with 5 μg Histone H1 (Sigma-Aldrich) as a substrate, incubated at 30°C for 15-30 min and stopped by addition of Laemelli sample loading buffer supplemented with 5% [vol/vol] β-Mercaptoethanol. To examine for CDK-specific activity, *in vitro* phosphorylation assays were performed in the presence of 50 μM of the CDK inhibitor Roscovitine (Sigma-Aldrich), in DMSO; DMSO alone was added to control reactions.

In vitro E1^E4 phosphorylation assays were performed using the following recombinant active kinases in the buffers provided: p42 MAPK (5 ng, New England Biolabs), PKA (500 ng, New England Biolabs), CDK2-cyclin A (0.3 ng, New England Biolabs) and CDK1-cyclin B (10 ng, New England Biolabs). *In vitro* kinase assays with CDK2-cyclin E (60 ng, Cell Signaling Technology) were carried out in 40 mM HEPES, pH7.0, 20 mM $MgCl_2$, 10 mM EGTA and for CDK1-cyclin A (50 ng, Cell Signaling Technology) in 5 mm MOPS, pH7.2, 2.5 mM β-glycerophosphate, 1 mM EGTA, 0.4 mM EDTA, 5 mM $MgCl_2$, 50 nM DTT. Kinase assays were performed with recombinant active kinase in the appropriate buffer containing 50 mM ATP, 10 μCi $[γ-^{32}P]$-ATP (3000Ci/mMol) with 5 μg of soluble GST fusion proteins as substrates. In some experiments 5 μg of Histone H1 was used as a positive control. *In vitro* kinase reactions were incubated at 30°C for 15 min and stopped by addition of Laemelli sample loading buffer supplemented with 5% [vol/vol] β-Mercaptoethanol. Samples were heated at 100°C for 10 min and resolved on a 15% polyacrylamide SDS-PAGE gel.

2.6.8 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

12.5% polyacrylamide separating gels (350 mM Tris-HCL, 12.5% [wt/vol] polyacrylamide (acrylamide : bis-acrylamide 37.5 : 1), 1% [wt/vol) SDS, polymerised with 0.4% [wt/vol] ammonium persulphate (APS) and 0.08% [vol/vol] TEMED) were routinely used to separate proteins prior to Western blot analysis, whilst separation of proteins following *in vitro* kinase assays was carried out using 15% polyacylamide gels whereby the acrylamide concentration [wt/vol] of the gel was varied accordingly. Polyacrylamide 10-17.5% gradient separating gels were used to separate proteins for mass spectrometry, to allow enhanced separation across a range of molecular masses. Following polymerisation, separate solutions of 10 and 17.5% polyacrylamide gels, were placed into the inner and outer chambers of a gradient former

(Bethesda Research Laboratories) respectively, with a small magnetic stir bar placed within the inner chamber and the gradient maker centred on top of a magnetic stirrer to facilitate mixing. An attached Perista® pump (ATTO Bioscience and Technology) was subsequently turned on and the valve separating the chambers opened, enabling the 10% polyacrylamide gel solution to flow gradually into the 17.5% solution thus allowing formation of a gradient from the bottom to the top of the gel as it is poured.

To ensure a level surface of all polyacrylamide gels, a layer of isopropanol was added to the surface immediately following pouring. Once the separating gels had set, isopropanol was removed and a stacking gel (125 mM Tris pH 6.8, 4.5% [wt/vol] acrylamide, 1% [wt/vol] SDS, polymerised with 0.5% [wt/vol] APS and 0.125% [vol/vol] TEMED) was poured on top of the resolving gel and the gel combs inserted. Once set, the wells were filled with running buffer (25 mM Tris, 192 mM Glycine, 0.1% [wt/vol] SDS, pH 8.3).

Protein samples were prepared for SDS-PAGE by addition of Laemelli sample loading buffer supplemented with 5% [vol/vol] β-Mercaptoethanol, and boiled at 100°C for 10 min. Samples were loaded into wells of the gel with a Hamilton syringe (Hamilton Company), or for kinase assays with gel-loading pipette tips. Electrophoresis was performed in running buffer using vertical gel slab units. The Mini Protean® 3 Cell (Bio-Rad) was run at 25 mA for 1-2 h for smaller gels whilst for larger gels, the Sturdier system (Amersham Biosciences) was used and run overnight at 55 V or until the correct separation had been achieved.

2.6.9 Two-dimensional gel electrophoresis

First dimension tube gels (9 M Urea, 4% [wt/vol] acrylamide, 5% [vol/vol] pH 3.9-9.5 carrier ampholytes (Biolyte®, Bio-Rad), 2% [vol/vol] Nonidet P-40, polymerised with: 0.02%

[wt/vol] APS, 0.1 [vol/vol] TEMED (N,N,N',N' tetramethylethylene-diamine) were cast in 2 mm capillary tubes to a height of 120 mm. Cells were solubilised in 9 M Urea, 0.2% [vol/vol] ampholyte, 10 mM β-Mercaptoethanol, supplemented with 1 X Complete™ protease inhibitors (Roche), sonicated and insoluble material removed (as described in section 2.6.1). 10 μg of sample protein in a final volume of 20 μl was added to each tube gel, followed by 20 μl overlay buffer (5 M Urea, 1% [vol/vol] β-Mercaptoethanol, 0.5% [vol/vol] ampholyte). The tube gels were then inserted into a Gel Electrophoresis Cell (Bio-Rad) with catholyte (10 mM phosphoric acid) and anolyte (20 mM NaOH) added to the upper and lower electrophoresis chambers respectively. Iso-electrofocussing (IEF) was then performed at 400 V for 4 h before the tube gels were extracted from the capillary tubes and incubated at 60°C in Laemmli sample buffer (Bio-Rad) for 20 min. Proteins were then resolved in the second dimension by SDS-PAGE by placing the tube gel across the top of a 15% polyacrylamide gel and electrophoresis was carried out at 65 V overnight. Protein standards (Precision Plus dual standard colour markers, Bio-Rad) were run on each gel for determination of molecular mass.

2.6.10 Staining and de-staining of SDS-PAGE gels

2.6.10.1 Silver staining

For analysis of C-terminally tagged HA E4 proteins prior to mass spectrometry, silver staining was carried out with Silver Stain Plus Kit (Bio-Rad), according to the manufacturer's instructions.

2.6.10.2 Coomassie staining

For analysis of GST recombinant proteins and Histone H1, the proteins were fixed and visualized by addition of fixation buffer (40% [vol/vol] ethanol, 10% [vol/vol] acetic acid)

with 1% [wt/vol] Brilliant-Blue R (Sigma-Aldrich), with gentle agitation at room temperature for 1 h, before de-staining with multiple changes of fixation buffer until clear protein bands can be visualised. Following staining, gels were dried and autoradiography performed.

2.6.11 Mass spectrometry

For analysis of E4 products following HA immunoprecipitation of proteolytically cleaved products silver stained bands were extracted from the gel with a sterile scalpel and analysed by mass spectrometry on a LCQ DECA XP PLUS (Thermo Electron Corporation) using liquid chromatography-tandem mass spectrometry (LC MS/MS) by the School of Biosciences, University of Birmingham.

For analysis of the site of E4 phosphorylation, the kinase reaction was performed in the absence of $[\gamma\text{-}^{32}P]$-ATP and the Coomassie-stained GST-HPV18 E4 band was extracted from the gel with a sterile scalpel, digested with trypsin and analysed by the mass spectrometry method of Neutral Loss - Electron Capture Dissociation (NL-ECD MS/MS) by the School of Biosciences, University of Birmingham, as described elsewhere (Sweet et al., 2006). Briefly, tryptic peptides separated by liquid chromatography were introduced via a nanospray source into a Fourier transform ion cyclotron resonance mass spectrometer where the mass of each peptide was calculated and ions then subjected to collision-induced disassociation. The mass spectrometry data was scanned for site-determining ions whose presence indicates phosphorylation of a particular residue (Beausoleil et al., 2006).

2.6.12 Western blot analysis

2.6.12.1 Electrophoretic transfer of proteins

Following separation by SDS-PAGE, proteins were transferred to BioTrace® NT nitrocellulose blotting membrane (Pall Life Sciences) using a Trans-Blot® Cell transfer tank (Bio-Rad), in transfer buffer (25 mM tris, 192 mM glycine, 20% [vol/vol] methanol, pH 8.3), at 350 mA for 3½ h, or overnight at 15 V. To ensure transferral was successful, membranes were stained briefly in Ponceau stain (1% [wt/vol] Ponceau Red, 3% [wt/vol] trichloroacetic acid) to visualise proteins, before de-staining with repeated washes in Tris-buffered saline (50 mM tris, 100 mM NaCl, pH 7.4) containing 0.1% (vol/vol) Tween-20 (TBS-T).

2.6.12.2 Immunodetection of proteins

Nitrocellulose membranes were incubated in blocking buffer consisting of: 5% [wt/vol] dried skimmed milk in TBS-T at room temperature for 30-60 min. Primary antibodies were diluted in 5% milk in TBS-T and incubated with membranes for between 2 h and overnight. HPV18 E1^E4 was detected using the anti-GST-HPV18 E1^E4 rabbit polyclonal antibody, R424 used at a dilution of 1:1000, or the mouse MAb, ID11 used at a dilution of 1:5. Involucrin was detected with a mouse MAb (at 1:200 dilution; #SY5, Sigma-Aldrich) and Keratin 1 with a mouse MAb (at a 1:500 dilution; Sigma-Aldrich). To control for equal loading, β-actin levels were determined by using an anti-β-actin MAb (1:40,000 dilution; Sigma-Aldrich), or GAPDH was detected with a mouse anti-GAPDH MAb (1:200 dilution; SantaCruz Biotechnology). Following incubation with primary antibodies, membranes were subjected to two 15 min washes in TBS-T prior to incubation with secondary antibodies, diluted in 5% milk TBS-T for 1 hour at room temperature. Primary antibodies were detected with a horseradish peroxidase-conjugated anti-mouse (Sigma-Aldrich) used at a dilution of 1:2000 or

-rabbit immunoglobulins (DAKO) at a dilution of 1:3000 and were visualised by chemiluminescence detection in accordance with manufacturer's instructions (Amersham Pharmacia) following two 15 min washes of membranes in TBS-T at room temperature. For detection of the R424 antibody, the nitrocellulose membrane was extensively washed after incubation with secondary antibody, for up to 6 h in TBS containing 1% (vol/vol) Tween-20 (TBS-T).

2.6.12.3   Generation of an HPV18 E1^E4 T23 phospho-specific antibody

An HPV18 E1^E4 threonine 23 phospho-specific antibody was raised against the peptide: H_2N - LNS YST(PO_3H_2) PPH RIP A - $CONH_2$. An HPV18 E1^E4 non-phospho antibody was generated against the peptide: H_2N - LNS YST PPH RIP A - $CONH_2$. Antibodies were raised in rabbit by Eurogentec and used at a working dilution of 1:100 for Western blot analysis.

2.6.13   Densitometry

Protein levels from Western blot films were examined using a Scanning Densitometer GS-800 (Bio-Rad) and images were analysed used the Quantity One 4.6.7 software (Bio-Rad).

2.7   Generation and maintenance of HFK cell lines containing HPV18

genomes

Generation of HFK cell lines was undertaken according to the methods described by Wilson

and Laimins (Wilson & Laimins, 2005) unless otherwise stated.

2.7.1   Maintenance of HFKs prior to transfection

Normal primary human foreskin keratinocytes (HFK) were isolated from neonatal foreskin by

Dr Sally Roberts following the protocol described by Meyers and Laimins (Meyers &

Laimins, 1994) or were purchased from Clonetics. HFKs were cultivated in serum-free

keratinocyte growth medium (KGM, Invitrogen-Gibco), replaced every 2 days and cells were

grown to a confluency no greater than 80%. All cells were maintained in an incubator at 37˚C

with 5% CO_2. Cells were passaged by removal of media followed by washing twice in PBS

before addition of 1 ml 0.05% trypsin/ EDTA and incubation at 37˚C. After 5-10 min, trypsin

was inactivated by addition of 1 ml trypsin-neutralising solution (TNS, Invitrogen 0.25 mg/ml

soybean trypsin inhibitor in Dulbecco's PBS without calcium or magnesium, pH 7.2) before

transfer of cells to a 15 ml conical tube. The dish was rinsed twice with 2 ml KGM and the

cells pelleted by centrifugation at 538 g for 5 min at room temperature. The supernatant was

subsequently removed and the pellet re-suspended in 10 ml KGM and transferred to a 100

mm tissue culture dish or, prior to transfection into a 50 mm dish.

2.7.2   Freezing of HFKs

Cells were frozen as described in section 2.3.3, with re-suspension of HFK at 2×10^6/ml in:

KGM supplemented with 10% [vol/vol] FBS and 10% [vol/vol] DMSO or following

transfection with HPV18 genomes in E-medium supplemented with 10% [vol/vol] FBS and 20% [vol/vol] glycerol.

2.7.3 Preparation of E-medium

Following transfection of HFKs with HPV18 genomes, cells were grown in E-medium. To make 2 litres of E-medium the following components were mixed: 1200 ml DMEM HEPES modification, 640 ml Ham's F-12, 20 ml of 100 X cocktail (described below), 10 ml of 100 X Penicillin/Streptomycin (PAA Laboratories), 100 ml FBS (5% [vol/vol]), 2 ml 1000 X Cholera toxin (ICN Biomedical) and 2 ml 1000 X Hydrocortisone (Sigma).

100 X Cocktail: To make 200 ml of 100 X cocktail the following components were mixed: 20 ml 0.18 M Adenine (Sigma, 0.486 g in 15 ml H_2O with addition of 10 M HCL until dissolved, then addition of 5 ml H_2O) , 20 ml 5 mg/ml Insulin (Sigma, 0.1 g added to 20 ml 0.1 M HCL), 20 ml 5 mg/ml transferin (Sigma, 0.1 g added to 20 ml PBS), 20 ml 2×10^{-8} M 3,3',5-Triiodo-L-thyronine (T_3) in PBS (Sigma, 13.6 g added to 100 ml 0.02 M NaOH to make 2×10^{-4} M T_3 further diluted in PBS to make 2×10^{-8} M). The 100 X cocktail mix was filter sterilised and frozen prior to use.

Following mixing, E-medium was then filter sterilised and kept at 4°C in the dark. Immediately prior to use media was supplemented with 2 mM L-glutamine and 5 ng/ml epidermal growth factor (EGF, BD Biosciences). E-medium was routinely made by a laboratory technician (Ms Emma Yates).

2.7.4 Maintenance and irradiation of J2-3T3 mouse fibroblasts

Following transfection of HFK with HPV18 genomes, cells were cultivated with a feeder layer of irradiated J2-3T3 mouse fibroblasts. J2-3T3 cells were routinely grown in Dulbecco's Modified Eagle Medium (DMEM) HEPES modification (Sigma-Aldrich) supplemented with 10% [vol/vol] new born bovine serum, in 100 mm tissue culture dishes (Iwaki) and were grown to a confluency of no greater than 80% before passaging. J2-3T3 cells were grown up to passages 20-25 before replacement with earlier passage cells. J2-3T3 cells were passaged as described previously in section 2.3.2 and were frozen in DMEM supplemented with 20% [vol/vol] bovine serum, and 10% [vol/vol] DMSO (as described in section 2.3.3).

In preparation for irradiation, J2-3T3 cells were harvested and re-suspended in E-medium at 2 x 10^6 cells/ml in a 50 ml conical tube. Cells were exposed to 50 Gy of cobolt-60 γ-rays from a Pantec irradiator. Irradiated J2-3T3s were plated at 2 x 10^6 cells per 100 mm dish in E-medium at least 2 h prior to addition of HFKs. If not required within 24 h post-irradiation, J2-3T3 cells were stored in E-medium at 4˚C for up to 3 days.

2.7.5 Generation of HFK HPV18 cell lines

2.7.5.1 Construction of HPV18 mutant genomes

To generate mutations within the HPV18 genome the primer pairs listed in Table 5 were utilised in site-directed mutagenesis using the QuikChange Kit (Stratagene) following manufacturer's instructions. The pGEMII-HPV18 plasmid was used as the template DNA and the number of PCR cycles used were between 18 and 25.

Table 5:   **Mutants generated in pGEMII-HPV18 genome expression vector**

Mutant	Direction	Primer Sequence
L^{14}L^{15}>PP	Forward (5')	5' GAC GAC ACG GTA TCC GCC ACC CAG CTT GTT AAA CAG C 3'
	Reverse (3')	5' GCT GTT TAA CAA GCT GGG TGG CGG ATA CCG TGT CGT C 3'
L^{14}L^{15}>RR	Forward (5')	5' GAC GAC ACG GTA TCC GCG ACG CAG CTT GTT AAA CAG C 3'
	Reverse (3')	5' GCT GTT TAA CAA GCT GCG TCG CGG ATA CCG TGT CGT C 3'
^{14}LLSLL^{18}> RRSSS	Forward (5')	5' GAC GAC ACG GTA TCC GCG ACG CAG CTC GTC AAA CAG CTA CAG CAC AC 3'
	Reverse (3')	5' GTG TGC TGT AGC TGT TTG ACG AGC TGC GTC GCG GAT ACC GTG TCG TC 3'

The complete HPV18 mutant genomes were sequencing using the primers listed in Table 6, which cover the entire HPV18 genome sequence, to ensure that the only mutations generated by site-directional mutagenesis were intentional.

Table 6: **Sequencing Primers for the total HPV18 genome**

Primer	Direction	Primer Sequence
E6	Forward (5')	5' CC GAA AAC GGT CGG GAC CG 3'
E7	Forward (5')	5' CAC AAC ATA GCT GGG CAC TAT AG 3'
E1(1)	Forward (5')	5' CCA GAA GGT ACA GAC GGG GAG 3'
E1 (2)	Forward (5')	5' GA CGG GGG CAC AGA GGG C 3'
E1 (3)	Forward (5')	5' GCC CTG TTG CGT TAC AAA TG 3'
E1 (4)	Forward (5')	5' G CAA ACA TTA TAG GCG AGC CC 3'
E1 (3)	Reverse (3')	5' G CAG ACA CCG AAG GAA ACC C 3'
2F	Forward (5')	5' CCA GCA AAG GAT AAT AGA TGG CC 3'
3F	Forward (5')	5' GGC CCT ACA AGG CCT TGC AC 3'
E2F	Forward (5')	5' GGA GAT TGT ATT ATG TAA AGC 3'
4F	Forward (5')	5' GGA ATA CTG ACT GTA ACA TAC C 3'
5F	Forward (5')	5' CCC CTG CCA CAG CAT TCA CAG
L2 (1)	Forward (5')	5' CAC GTC CCC CAG TGG TTA TTG 3'
L2 (2)	Forward (5')	5' CCT GAG TTT CTT ACA CGT CCA TCC 3'
L1 (1)	Forward (5')	5' GCC TGT ATA CAC GGG TCC TG 3'
L1 (2)	Forward (5')	5' CA CAA CGT TTA GTG TGG GCC 3'
L1 (3)	Forward (5')	5' GCA GAT CCT TAT GGG GAT TCC ATG 3'
L1 (4)	Forward (5')	5' G CAG TAT AGC AGA CAT GTT GAG G 3'
End (1)	Forward (5')	5' GCG TGT GCG TGT ACG TGC CAG 3'
End (2)	Forward (5')	5' CAA TTG GCG CGC CTC TTT GG 3'

85

2.7.5.2   Preparation of HPV18 genomes for transfection of HFK

To extract the HPV18 genome from the bacterial sequences of the pGEMII vector, 10 µg of pGEMII-HPV18 wild type or mutant genomes were digested with *Eco*RI in a total volume of 50 µl (as described in section 2.2.5). Following confirmation of complete digestion, verified by running a 200 ng aliquot on an agarose gel, the reaction was heat inactivated at 65°C for 20 min. HPV18 genomes were then re-circulised in a large, 900 µl total volume ligation reaction to encourage self-ligation of HPV genomes, with T4 DNA ligase (400 U/ml, New England Biolabs) in the buffer provided and incubated overnight at 16°C. The DNA was then precipitated by addition of 2 volumes of isopropyl alcohol and 1/5 volume 5 M NaCl followed by vortexing before incubation at -20°C overnight. Following centrifugation at 16,100 g for 30 min at 4°C, the supernatant was removed and the pellet washed with 70% ethanol (pre-cooled to -20°C), prior to centrifugation at 16,100 g for a further 15 min. The supernatant was subsequently removed and the DNA pellet re-suspended in 12 µl of 1 X TE buffer. The DNA concentration was then estimated by running a 200 ng aliquot on an agarose gel.

2.7.5.3   Transfection of HFKs with HPV18 genomes

 HPV18 genomes were co-transfected into HFK with pcDNA 3.1 (Invitrogen), a plasmid which carries the neomycin resistance gene. Control transfections were also carried out in parallel, with the pcDNA 3.1 vector alone, with a green fluorescent protein (GFP Clontech) reporter expression vector and with no plasmid. For each transfection, 94 µl of KGM was dispensed into a polypropylene Falcon tube, followed by addition of 6 µl FuGene 6 Transfection Reagent (Roche). In a separate Falcon tube, 100 µl of KGM, 1 µg of HPV genomic plasmid DNA or control plasmid together with 1 µg of Neomycin resistance plasmid

were mixed, prior to addition of the KGM/FuGene 6 mixture. The Falcon tube was then tapped gently to mix, and then incubated at room temperature for 30 min. HFK at 50-60% confluency were selected, and the media replaced with SFM media (Invitrogen-Gibco), prior to addition of the transfection mixture and gentle rocking of the dish to mix.

The transfection efficiency of the GFP transfection was established the following day using a Zeiss Axiovert 100 microscope, with GFP-expressing cells typically representing 5% of the population. Transfected HFK were then trypsinised and re-plated onto 100 mm tissue culture dishes with 2×10^6 irradiated J2-3T3 fibroblasts in the presence of E-media. Over the next 8 days the HFK underwent a selection period in which on alternate days the cells were treated with G418 or with the addition of irradiated J2-3T3 cells. The selection period began with replacement of E-media with E-media supplemented with 100 µg/ml G418 (PAA Laboratories). On days following selection 1×10^6 irradiated J2-3T3 cells were added. The concentration of G418 was increased to 200 µg/ml for the latter two of the four selection days and following the last day of G418 treatment the media was replaced with E-media without selection and 2×10^6 irradiated J2-3T3 cells. In the event that cells reached 80% confluency during the selection process, the dish was split equally onto two dishes of irradiated J2-3T3 fibroblasts in the presence of E-media and selection was resumed the following day.

Once colonies reach approximately 2 cm in diameter, routinely observed after 1-2 weeks, the two plates were pooled and plated onto a single plate with irradiated J2-3T3 cells. Upon reaching 80% confluency the HFKs were passaged into five 100 mm tissue culture dishes and once 70-80% confluency is reached, cells from four of the five dishes were harvested and stored in liquid nitrogen (as described previously 2.7.2).

2.7.5.4   Maintenance of HFKs following transfection of HPV18 genomes

HFKs were routinely grown on a feeder layer of 2 x 10^6 irradiated J2-3T3 fibroblasts, in E-media supplemented with EGF prior to use. The media was replaced every two days and the cells grown to a confluency of no greater then 80% prior to passaging. Irradiated J2-3T3 cells were plated out in E-media at least 2 h, or on the previous day, prior to addition of HFK. HFK were passaged by first removing the J2-3T3 feeder layer, using 0.5 mM EDTA in PBS. Once the feeder cells became detached, a transfer pipette was then used to aspirate the EDTA and the dish was then washed twice in PBS with gentle aspiration. HFKs were then removed with 2 ml trypsin/EDTA (0.25% [wt/vol] trypsin/ 1 mm EDTA) and incubation at 37˚C for 5-10 min with regular checks. The cells were harvested in 10 ml media and transfer to a 15 ml conical tube. A further 3 ml of media was then used to wash the dish and the HFKs spun by centrifugation at 538 g for 5 min at room temperature. The supernatant was removed and the cells washed in 10 ml PBS, prior to a further 5 min centrifugation and re-suspension in E-media. HFK keratinocytes were routinely plated at 2 x 10^5 cells/dish on irradiated J2-3T3 cells. The plates were rocked several times to ensure thorough mixing and the cells incubated at 37˚C. The media was replaced every 2 days and cells typically took between 5-7 days to reach a confluency of 80%.

2.8   Analysis of HFK cell lines harbouring HPV18 genomes

2.8.1   HFK cellular proliferation assays

HFKs were plated at 2 x 10^4 cells / well in a 6 well plate (Iwaki) containing 2 x 10^5 irradiated J2-3T3 cells / well. Cells were harvested at various times, with each time point performed in duplicate. Viable HFKs were counted following removal of feeder fibroblasts.

2.8.2　　　　Differentiation of HFKs in semi-solid media

2.8.2.1　　　　Preparation of semi-solid medium

Semi-solid medium (1.5% [wt/vol] methylcellulose in E-media supplemented with 5% [vol/vol] FBS), was prepared by first dispensing 6 g of methylcellulose (Sigma) into a Pyrex bottle with a stir bar and autoclaving. E-medium (200 ml) without supplementation with EGF (E-media –EGF), was subsequently added, and the solution gently swirled before placing in a water bath at 60°C. At regular intervals of approximately 5 min, the solution was gently swirled and after 30 min the bottle was cooled on ice for 2 min. A further 180 ml of E-media –EGF was subsequently added before wrapping the bottle in foil and stirring vigorously at 4°C overnight. The following day, 20 ml FBS was added and a further 2 h of stirring at 4°C was performed prior to storage of the semi-solid media at 4°C for up to 2 weeks.

2.8.2.2　　　　Suspension of HFK in semi-solid medium

HFKs at a confluency of approximately 80%, were harvested following removal of J2-3T3 fibroblasts as described previously. Cells were pooled and re-suspended in E-media –EGF to form a single cell suspension. The cell suspension was then added to a Petri dish containing 25 ml semi-solid media, pre-warmed to 37°C, by dispensing cells in a volume of 1 ml to each dish in a drop-wise manner. To ensure an even suspension of cells, the semi-solid media was mixed carefully but extensively before incubation at 37°C.

2.8.2.3　　　　Harvesting of HFK from semi-solid medium

HFK were harvested at 24 or 48 h following suspension in methylcellulose. Using a cell scraper, the semi-solid medium was transferred equally into 4 x 50 ml tubes and the dish washed three times in ice-cold PBS to remove remaining cells. The 50 ml tubes were filled with PBS and inverted 2-3 times prior to centrifugation at 537 g for 10 min at 4°C. Following

centrifugation the supernatant was removed to leave 10 ml remaining in each tube and the cells re-suspended in this volume and pooled into 2 tubes. PBS was added up to 50 ml to each tube and cells were pelleted. Finally the cells were pooled into 1 tube and after centrifugation were re-suspended in 10 ml PBS, transferred to a 15 ml conical tube, pelleted, then divided into aliquots for analysis by Northern, Southern and Western blotting.

2.8.3 Southern blot analysis

2.8.3.1 Isolation of genomic DNA

HFK were harvested following removal of J2 3T3 fibroblasts (as described in section 2.7.4) or following harvest from methylcellulose as described previously and pelleted cells were washed once in PBS and re-suspended in 3 ml lysis buffer (400 mM NaCl, 10 mM Tris-HCl, 10 mM EDTA, pH 7.4). RNase A was added to a final concentration of 50 μg/ml, cell lysates vortexed, and then incubated for 15 min at room temperature. Proteinase K (Roche) was subsequently added to a final concentration of 50 μg/ml and SDS to 0.2% [wt/vol], prior to vortexing and incubation at 37°C for 6 h or overnight. Cellular DNA was subsequently sheared by passing lysate through an 18-gauge needle, repeated 10 times. Phenol-chloroform extraction was performed by addition of 6 ml phenol/chloroform/isoamylethanol (25:24:1) (Sigma-Aldrich) prior to centrifugation at room temperature for 5 min at 4838 g. The upper aqueous phase was retained and the extraction repeated twice, before extraction of the aqueous phase with choloroform/isoamylethanol (24:1) (Sigma-Aldrich). DNA was then precipitated by addition of two volumes ethanol and 1/10th volume 3 M sodium acetate and stored at -80° C for 1 h. Following centrifugation at 4°C for 30 min at 4838 g, the supernatant was removed and the DNA pellet washed twice, first in 100% ethanol, then in 70% ethanol before re-suspension in 100 ml 1 x TE buffer and incubation at 37°C for 30 min or 60°C for

10 min. The DNA yield was determined with a Nanodrop™ spectrophotometer (Thermo Scientific) and samples were stored at 4˚C until required.

2.8.3.2 Preparation of genomic DNA for Southern blot analysis

For analysis of viral episomes, genomic DNA extracted from HFKs was subjected to restriction digestion with *Bgl*II, which has no restriction sites within the HPV18 genome. For analysis of linear genomes, a DNA digest was performed with *Eco*RI which has one restriction site within the HPV18 genome. *Dpn*I digests of all DNA samples were carried out to remove any residual DNA input. The 5 μg of digested genomic DNA samples were analysed by electrophoresis in 0.8% [wt/vol] agaorse gel (as described in section 2.2.3). To enable quantification of genome copy numbers, copy number standards were generated by digestion of pGEMII-HPV18 with *Eco*RI to release the HPV18 genome from the vector and the equivalent of 5 and 50 genome copies/per cell were run as standards on the agarose gel.

2.8.3.3 Capillary transfer of DNA from agarose gel to nylon membrane

Prior to blotting, the agarose gel was washed twice in 250 mM HCl for 20 min at room temperature with shaking before washing in 0.4M NaOH for 30 min. To set up the transfer of DNA from the agarose gel to a nylon membrane, a tray was first filled with 1.5 l of 400 mM NaOH and a glass plate rested on top of the tray. A 24 x 33 cm single sheet of Whatman™ 3MM paper was soaked in 0.4M NaOH and laid across the glass plate with both ends submerged in the NaOH, thus forming a wick. A pipette was then rolled across the Whatman paper, ensuring that bubbles were removed. Three more layers of 24 x 33 cm Whatman paper were then placed on top, in the same manner as before and the agorose gel placed on top of the Whatman paper wick with the loading wells facing downwards. Bubbles were subsequently removed with a pipette prior to placing a 20 x 22.5 cm sheet of Gene Screen

Plus nylon membrane (Perkin Elmer) pre-soaked in NaOH on top, followed by four layers of 21 x 23.5 cm Whatman paper, placed one at a time, with care taken to remove bubbles. Two stacks of absorbent paper towels, approximately 10 cm in height, were placed on top, covering the Whatman paper, with care taken to ensure there were no gaps between the stacks. A second, smaller glass plate was placed on top of the towels, and a weight centred on top. SaranWrap was then placed between the top and bottom layers of the filter paper to ensure they were not in contact and the transfer blot was left overnight. Following marking of the position of the loading wells on the nylon membrane with pencil, the transfer stack was disassembled and the DNA auto-crosslinked to the membrane using a UV crosslinker (Stratalinker, Stratagene) on auto-crosslink mode. The membrane was then frozen at -20°C following soaking in 2x SSC and wrapping in SaranWrap or alternatively was used immediately for hybridisation.

2.8.3.4   Preparation of radiolabelled DNA probe

In preparation for generation of a DNA probe the pGEMII-HPV18 vector was linearised by *EcoRI* digestion (as described previously 2.2.5). 50 ng of digested vector was subsequently diluted into 45 μl of 1 X TAE buffer and the DNA denatured by heating at 95°C for 5 min followed by incubation on ice for 2 min. The denatured DNA was then used as a template for radiolabelled probe generation using the Ready To Go DNA labelling beads − dCTP kit (Amersham) following manufacturer's instructions. Following re-suspension of the labelling beads with the denatured linear DNA and addition of 50 μCi of [^{32}P] dCTP (Perkin Elmer) a 30 min incubation at 37°C was performed. To purify the labelled probe DNA from the reaction the Illustra Probe Quant G-50 microcolumn (Amersham), a radiolabelled probe purification kit, was used following manufacturer's instructions. The microcolumn first

underwent centrifugation at 801 g at room temperature for 1 min prior to transfer of the labelling reaction to the tube and centrifugation at 801 g for a further 2 min to elute the probe.

2.8.3.5 Hybridisation of radiolabelled probe to immobilised DNA

To prepare the hybridisation buffer: a 2 X hybridisation solution (5 X SSC (1.5 M sodium citrate, 750 mM NaCl, pH 7.0), 10 X Denharts (0.2%[w/vol] Ficoll 400, 0.2%[w/vol] Polyvinylpyrolilone, 0.2% [wt/vol] BSA fraction V (Sigma-Aldrich)) and 20% [wt/vol] Dextran sulphate), which was frozen at -20°C until required, was diluted immediately prior to use 1:1 with formamide (Sigma-Aldrich) and addition of SDS to a final concentration of 0.1% [wt/vol]. 200 μl of (10 mg/ml) salmon sperm DNA (Invitrogen) was denatured by heating at 95°C for 5 min, followed by cooling on ice for 2 min and was subsequently diluted into 10 ml hybridisation buffer. The nylon membrane on which the DNA was immobilised was then rolled and carefully placed into a glass hybridisation canister using forceps with the blotted DNA side facing inwards. The hybridisation buffer containing salmon sperm DNA was then added and the canister placed into a hybridisation oven with rotation at 42°C for 1 h. Following this pre-hybridisation step, hybridisation buffer containing the radiolabelled DNA probe was prepared by addition of 200 μl of (10 mg/ml) salmon sperm DNA to the probe, followed by boiling for 5 min, cooling and addition to the 10 ml hybridisation buffer as described previously. Following careful pouring to remove the pre-hybridisation buffer from the canister, the hybridisation buffer containing the probe was added and the canister incubated at 42°C with rotation overnight.

2.8.3.6 Stringency washes

Following removal from the hybridisation canister, the membrane was rinsed briefly in 2 x SSC (600 mM sodium citrate, 300 mM NaCl, pH 7.0), 0.1% [wt/vol] SDS, using a sponge to

wipe over the surface over the membrane, before 2 x 15 min washes in 500 ml 2 x SSC, 0.1% [wt/vol] SDS were performed at room temperature with agitation. Further stringency washes of: 2 x 15 min in 500 ml 0.5 x SSC , 0.1% [wt/vol] SDS, followed by 2 x 15 min in 500 ml 0.1 x SSC , 0.1% [wt/vol] SDS were carried out at room temperature. A final stringency wash was carried out in 500 ml 0.1 x SSC, 1% [wt/vol] SDS at 55˚C for 15 min before the membrane was wrapped in SaranWrap and autoradiography performed. Exposure of the membrane to a Phosphor Screen (Molecular Dynamics, GE Healthcare) followed by analysis with STORM imaging on a Storm860 (GE Healthcare) enabled quantification of band intensities using the ImageQuant 5.0 software (GE Healthcare).

2.8.4   Northern blot analysis

2.8.4.1   Extraction of RNA from HFKs

To prevent degradation of RNA through contamination with RNase during the extraction process, surfaces and pipettes were wiped with RNase AWAY (Invitrogen) and gloves were changed frequently. Following harvest of HFKs the cellular pellet was immediately re-suspended in 1 ml RNA STAT-60™ (AMS Biotechnology) and incubated at room temperature for 5 min or alternatively frozen at -80˚C until required. 200 μl of choloroform was subsequently added followed by inverting the mixture vigorously for 15 sec and incubating at room temperature for 2 min. The cell extract was subsequently centrifuged at 16,100 g for 15 min at 4˚C prior to transferral of the upper aqueous phase into an eppendorf tube and addition of 500 μl isopropanol. Following incubation at room temperature for 10 min and centrifugation at 16,100g for 10 min at 4˚C the supernatant was removed. The RNA pellet was subsequently washed in 75% [vol/vol] ethanol, vortexed and centrifuged at 16,100g for 5 min at 4˚C. The pellet was then air-dried briefly before re-suspension in 100 μl of

nuclease-free water and the RNA sample stored at -80°C until required. The RNA yield was determined with a Nanodrop™ spectrophotometer (Thermo Scientific).

2.8.4.2   RNA gel electrophoresis

Prior to use all gel equipment was rinsed with 3% [vol/vol] hydrogen peroxide to prevent contamination of RNase, and then rinsed thoroughly with sterile water. RNA formaldehyde agarose gels (1% [wt/vol] agarose in MOPS buffer (20 mM MOPS, 5 mM NaOAc, 1 mM EDTA, pH 7.0) 6% [wt/vol] formaldehyde and 1 µg/ml EtBr) were prepared by boiling of agarose in MOPS buffer prior to cooling to 60-65°C before addition of formaldehyde and EtBr. 10 µg of RNA was loaded into gels in 20 µl RNA load dye (Ambion) and gels were run in MOPS buffer at 50 V overnight.

2.8.4.3   Capillary transfer of RNA from formaldehyde agarose gel to nylon membrane

To transfer of RNA from the formaldehyde agarose gel to a nylon membrane, a tray was first filled with 1.5 l of 10X SSC (3 M sodium citrate, 1.5 M NaCl, pH 7.0), and a glass plate rested on top of the tray. A bridge was made with three sheets of 24 x 33 cm single sheet of Whatman™ 3MM paper as for Southern blotting (section 2.8) and the gel was then placed on top with the wells face down. Following wetting of single sheet of Zeta-Probe membrane (Bio-Rad), this was laid on top of the gel followed by two sheets of Whatman paper and a stack of paper towels as described previously.

2.8.4.4   Hybridisation of radiolabelled probe to immobilised RNA and stringency washes

Following transfer, the position of the wells was marked and the blots were washed in 2 X SSC before UV cross-linking (as described in section 2.8.3.3), and placing in a hybridisation

canister with RNA hybridisation buffer (1 mM EDTA, 0.5 M Na_2HPO_4 and 7% [wt/vol] SDS) for 10 min at 65˚C. The HPV18 probe was generated in the same manner as for Southern blotting, and following addition of equal volumes of 0.2 M NaOH, heating at 95˚C for 5 min and cooling on ice for 2 min, the probe was incubated with nylon membranes in RNA hybridisation at 65˚C overnight. The following day the membrane was washed for 5 min in 2 X SSC, 10% [wt/vol] SDS at room temperature, followed by a 15 min wash at 55˚C in 0.2 X SSC, 1% [wt/vol] SDS. Autoradiography and phosphor imaging was then performed as described for Southern blot analysis.

2.8.5 Growth and harvesting of organotypic raft cultures

Organotypic raft cultures were prepared on my behalf by Dr Sally Roberts as described in (Wilson & Laimins, 2005) and were harvested by peeling of HFKs from the collagen layer followed by homogenisation in 1 ml raft lysis buffer (50 mM Tris pH7.4, 150 mM NaCl, 1 mM EDTA, 1% [vol/vol] NP40 supplemented with protease inhibitors), performed on ice in a glass Dounce homogeniser. Following a 30 min incubation, lysate was transferred with a transfer pipette into a pre-cooled eppendorf tube and underwent centrifugation at 16,100 g for 15 min at 4˚C. The supernatant was subsequently removed and the pellet was solubilised in 250 μl of Laemelli buffer supplemented with 5% [vol/vol] β-Mercaptoethanol, before centrifugation at 16,100 g for 10 min at 4˚C. The supernatant was then heated at 100˚C for 5 min prior to SDS-PAGE and Western blot analysis (as described in sections 2.6.8 and 2.6.12 respectively).

2.9   List of suppliers and addresses

Altabiosciences

University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Amersham Pharmacia Biotech UK Ltd

Amersham Place, Little Chalfont, Bucks HP7 9NA,UK

AMS Biotechnology

63B Milton Park, Abingdon OX14 4RX, UK

Applied Biosystems

Lingley House, 120 Birchwood Boulevard, Birchwood, Warrington, WA3 7QH, UK

ATTO Bioscience and Technology

1-5-32 Yushima, Bunkyo-ku, Tokyo 113-0034, JAPAN

BD Biosciences

Two Oak Park, Bedford, MA 01730, USA

Bethesda Research Laboratories

8717 Grovemont Circle, Gaithersburg, Maryland 20877, USA.

Bio-Rad Laboratories

Bio-Rad House, Maxted Road, Hemel Hempstead, Hertfordshire, HP2 7DX, UK

Cancer Research UK

P O Box 123, Lincoln's Inn Fields, London, WC2A 3PX, UK

Cell Signaling Technology Inc

3 Trask Lane, Danvers, MA 01923, USA

Clonetics Biowhittaker

8830 Biggs Ford Rd,Walkersville, MD 21793-8415, USA

Covance

210 Carnegie Center, Princeton State, 08540, USA

DAKO UK Ltd

Cambridge House, St Thomas Place, Ely, CB7 4EX, Cambridgeshire, UK

Fischer Scientific UK Ltd

Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG, UK

Eppendorf UK Ltd

Endurance House, Vision Park, Chivers Way, Histon, Cambridge, CB24 9ZR, UK

Eurogentec Ltd

Forest Business Centre, Fawley Rd, Fawley, Southampton, Hampshire SO45 1FJ, UK

Fisher Scientific UK Ltd

Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG, UK

GE Healthcare

Pollards Wood, Nightingales Lane, Chalfont St Giles, Bucks, HP8 4SP, UK

Gene-Bio Application Ltd

Kfar Hanagide, 76 875, Israel

Gibco-Invitrogen

3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK

Hamilton Company

Via Crusch 8, CH-7402 Bonaduz, GR, Switzerland

Heat Systems, Ultrasonics

1938 New Highway, Farmingdale, NY 11735, USA

Hendley Essex

12 Oakwood Hill Industrial Estate, Oakwood Hill, Loughton, IG10, UK

ICN Biomedical

15 Morgan Irvine CA 92618-2005, USA

Invitrogen

1600 Faraday Avenue, Carlsbad, CA 92008, USA

Molecular Probes

29851 Willow Creek Road, Eugene, OR 97402, USA

Neomarkers

46360 Fremont Blvd., Fremont, CA 94538, USA

New England Biolabs

New England Biolabs, 240 County Road, Ipswich, MA 01938-2723, USA

Nikon Inc

1300 Walt Whitman Road, Melville, NY 11747-3064, USA

PAA Laboratories Ltd

Termate Close, Houndstone Business Park, Yeovil Somerset, BA22 8YG, UK

Pall Life Sciences

Europa House, Havant Street, Portsmouth, PO1 3TD, Hampshire, UK

Perkin Elmer Ltd

Post Office Lane, Beaconsfield, Buckinghamshire, HP9 1QA, UK

Promega

Delta House, Chilworth Research Centre, Southampton, SO16 7NS, UK

Qiagen

QIAGEN House, Fleming Way, Crawley, West Sussex, RH10 9NQ, UK

Roche Diagnostics

Bell Lane, Lewes, East Sussex, BN7 1LG, UK

SantaCruz Biotechnology Inc

Bergheimer Str. 89-2, 69115 Heidelberg, Germany

School of Biosciences

University of Birmingham, Edgbaston, Birmingham, UK

Sigma-Aldrich Company Ltd

Fancy Road, Poole, Dorset, BH12 4QH, UK

Stratagene

11011 N. Torrey Pines Road, La Jolla, CA 92037, USA

Syngene Bio Imaging

Beacon House, Nuffield Road, Cambridge, CB4 1TF, UK

Thermo Electron Corporation

85 1st Ave, Waltham, MA 02451-1105, United States

Thermo Scientific

3411 Silverside Rd, Bancroft Building, Wilmington, DE 19810, USA

Zeiss

15 - 20 Woodfield Road, Welwyn Garden City, Hertfordshire, AL7 1JQ, UK

CHAPTER 3

HPV18 E1^E4 IS PHOSPHORYLATED *IN VITRO*

BY CYCLIN-DEPENDENT KINASES

AND PROTEIN KINASE A

Introduction

3.1   Protein phosphorylation

Protein phosphorylation is a major post-translational modification (PTM), which substantially influences protein structure and function and is critical to the regulation of a wide and diverse range of cellular biological activities; from metabolism and cell motility, to cell-cycle progression and cellular differentiation. Aberrant protein phosphorylation has been implicated in the pathogenesis of many diseases with abnormalities in the phosphorylation status of proteins contributing towards disease progression. Irregularities in this mechanism have been shown to play a role in carcinogenesis and several oncogenes can be classified as protein kinases, the mediators of phosphorylation. Phosphorylation is a highly prevalent regulatory mechanism and it is estimated that approximately 30% of cellular proteins are phosphorylated (Cohen, 2002a), highlighting the important and ubiquitous nature of this modulatory biological process.

Phosphorylation is a reversible covalent modification mediated by protein kinases, a class of enzymes which catalyse the transfer of the terminal gamma phosphate group of adenosine triphosphate (ATP), to the polar hydroxyl group of serine, threonine or tyrosine residues within the substrate protein. This modification is reversed by a second family of enzymes: protein phosphatases which catalyse the removal of the phosphate group. Protein kinases constitute the largest family of homologous proteins and it is estimated that nearly 2% of genes within the human genome code for these proteins (Manning et al., 2002). The opposing action of protein kinases and phosphatase determines the overall phosphorylation status of a protein and the presence of multiple phosphorylation sites within many proteins can result in a

multitude of permutations in phosphorylation states and enables a complex level of protein regulation.

Protein phosphatases are fewer in number than protein kinases, with 40 genes encoded within the human genome (Cohen, 2002b), and they tend to have broader specificity for their target substrates. Protein kinases however can be either highly specific or target a broad range of protein substrates and can be classified by either their mode of regulation or their specificity (Hanks & Hunter, 1995). The majority of kinases primarily target serine and threonine residues and are classified as protein-serine/threonine kinases (STK). More rarely, tyrosine residues are phosphorylated, which is catalysed by the protein-tyrosine kinase (PTK) group of enzymes which specifically targets these residues although some kinases retain dual-specificity and are able to phosphorylate threonine and tyrosine residues. Whilst some kinases display high specificity for their substrates others target a wide range of proteins. The tertiary structure of proteins play a major role in kinase recognition, although substrate specificity of a kinase can also often be conferred by the presence of consensus sequences within target proteins, which denote these proteins as substrates for phosphorylation. This can allow a range of different proteins with vastly varying structures to undergo phosphorylation by a single kinase thus enabling the function of a subset of proteins to be selectively and simultaneously modulated upon activation of the kinase. Many kinases play major roles in the regulation of signal transduction pathways whereby in response to a stimulus, kinases sequentially phosphorylate and activate downstream kinases and substrate proteins in the pathway resulting in a signalling cascade which ultimately elicits a cellular response such as gene expression or cellular proliferation.

Phosphorylation is a highly versatile and flexible modification, allowing the cell to quickly respond to external stimuli or intra-cellular control mechanisms, by rapidly but specifically modulating the activity of a particular protein or subset of proteins. Phosphorylation acts like a "molecular switch", whereby the addition of a phosphate group results in the rapid switch of a protein from one state to another which can regulate protein function in a variety of ways. The covalent attachment of a phosphate group to a protein results in the addition of two negative charges, often inducing a major conformational change which substantially alters the structure of the protein. A structural modification may perhaps expose or mask a protein interaction domain, thus enabling or preventing protein-protein interactions. Alternatively phosphorylation of a domain may form part of the recognition site for a protein, allowing or preventing a protein from being recognised by other proteins. The addition of a phosphate group may also stabilise a protein or target the protein for degradation or could alter the sub-cellular localisation, aiding movement of a protein between biological compartments, thereby allowing access to previously inaccessible proteins. The biological activity of a protein can also increase or decrease as a result of phosphorylation by converting a protein into an active or inactive conformation and the activity of protein kinases can themselves be regulated in this manner.

3.2   Phosphorylation of HPV E1^E4 proteins

Phosphorylation of the E1^E4 protein was first identified by Breitburd *et al. 1987,* who demonstrated that incubation of wart extracts with [^{32}P] orthophosphate resulted in the incorporation of radioactivity into both the 17 kDa HPV1 E1^E4 protein and the 16 kDa protein derived from N-terminal proteolysis of the full-length E1^E4 protein. Further characterisation of HPV1 E1^E4 phosphorylation demonstrated that multiple-site

phosphorylation of the protein occurs at both serine and threonine residues *in vivo* but not at tyrosine residues (Grand et al., 1989). Furthermore it was shown that the E4 16 kDa polypeptide underwent little phosphorylation during [^{32}P] labelling of wart extracts although this protein, together with the full-length HPV1 E1^E4 protein, was shown to be phosphorylated by cAMP-dependent protein kinase A (PKA) during an *in vitro* kinase assay (Grand et al., 1989). Poor incorporation of phosphate into an HPV1 E4 protein lacking the five N-terminal E1 residues was observed in comparison with the full-length E1^E4 protein, following radioactive labelling of HPV1 E4 transfected cells (Rogel-Gaillard et al., 1993).

In vitro kinase assays have confirmed that HPV11 E1^E4 is also targeted by PKA and also undergoes phosphorylation by mitogen-activated protein kinase (MAPK) (Bryan et al., 2000). Based on PKA and MAPK consensus phosphorylation sites of RRXS/T and PXXS/TP respectively, it was envisaged that T36 and S44 of HPV11 E1^E4 would be the phospho-acceptor residues for PKA with T53 a putative MAPK target site and these hypothesises were verified by mutational analysis (Bryan et al., 2000). To demonstrate that phosphorylation of HPV11 E1^E4 also occurs *in vivo*, HPV11-infected human xenograft tissue from athymic mice was radioactively labelled and consistent with the *in vitro* studies, the immunoprecipitated E1^E4 protein was shown to be phosphorylated on both serine and threonine residues (Bryan et al., 2000). Whilst mutation of both PKA consensus sites did not alter HPV11 E1^E4 localisation in E1^E4-transfected primary human keratinocytes, threonine 53 was shown to be necessary for localisation of E1^E4 to the cytoplasm suggesting that MAPK phosphorylation may play a role in localisation of the E1^E4 protein.

The HPV1, 16 and 18 E1^E4 proteins have been shown to associate with the serine-arginine (SR) – specific protein kinase-1 (SRPK1) (Bell et al., 2007), which phosphorylates arginine-

serine rich (RS) domains within a subgroup of substrate proteins termed SR proteins (Graveley, 2000). During *in vitro* kinase assays, whilst HPV16 and 18 E1^E4 proteins were not phosphorylated by SRPK1, an HPV1 E4 species, most likely derived from proteolysis of the full-length E1^E4 protein, undergoes phosphorylation by SRPK1, dependent upon the ability of the E4 protein to associate with the kinase. This interaction results in the sequestration of SRPK1 to HPV1 E1^E4 inclusion bodies. It has been hypothesised that the association of E1^E4 with SRPK1 may alter the regulation of SR proteins, many of which are involved in pre-mRNA splicing and that this could modify processing of viral transcripts and thus modulate the expression of viral genes.

Phosphorylation of E1^E4 proteins is not restricted only to the low risk HPV types, as the E1^E4 protein of high risk HPV16 can also undergo modification by phosphorylation (Davy et al., 2006). *In vitro* kinase assays have demonstrated that HPV16 E1^E4 is a substrate for CDK2-cyclin A and the target phospho-acceptor site has been mapped by both mutational analysis and mass spectrometry, to serine 32 (Davy et al., 2006).

Although phosphorylation of several E1^E4 proteins has been observed, the consequences of E1^E4 phosphorylation on the function of the protein is not yet known and I intend to examine this in the context of the E1^E4 protein of HPV18. Phosphorylation of HPV18 E1^E4 has yet to be characterised and analysis of HPV18 E1^E4 phosphorylation and subsequent identification of phosphorylation sites within the protein may enable an improved understanding of the biological function of E1^E4.

Results

3.3   Identification of potential phosphorylation sites within HPV18 E1^E4

The presence of kinase consensus sites within the HPV18 E1^E4 sequence would give an early indication that the protein may be a target for phosphorylation. A search for potential phosphorylation sites of HPV18 E1^E4 was carried out using ScanProsite (Hulo et al., 2006), developed by the Swiss Institute for Bioinformatics (www.expasy.ch) and NetPhosK 1.0 server (Blom et al., 1999) which scanned the primary amino acid sequences of HPV18 E1^E4 for the presence of kinase consensus sites. Predicted phosphorylation sites were identified within consensus motifs of HPV18 E1^E4 for a number of protein kinases including: Protein kinase C (PKC), PKA, Casein kinase II (CKII), CDK, MAPK, and Ribosomal S6 kinase (RSK) (Figure 3.1).

Six potential CDK phosphorylation sites were predicted within HPV18 E1^E4 which were subsequently examined together with the known consensus site for CDKs: S/TPXR/K, in which the underlined residue is targeted for phosphorylation. Out of the six predicted residues, only threonine 23 of HPV18 E1^E4 (^{23}TPPHR^{27}) is located within an optimal consensus site for CDK kinases, due to the adjacent proline residue at the +1 position relative to the phospho-acceptor, which is an absolute requirement for CDK phosphorylation (Brown et al., 1999). In addition to this there is also a requirement for a positively charged residue at the +3 position which is usually an arginine or lysine residue. Whilst an arginine residue is present at +4, the presence of a histidine at +3 is likely to meet this requirement as this residue can be either positively charged or neutral, making threonine 23 a likely candidate for phosphorylation by CDK. This residue is also located within a consensus site for MAPK which can be either PXXS/TP or PXS/TP as optimal motifs, or in some cases the minimal

A

Site	NetPhosK	Scan Prosite
T-9	PKC	PKC
S-20	CDK	
S-22	CDK	
T-23	MAPK	
T-23	CDK	
T-40	PKC	PKC
T-51	PKC	
T-51	CDK	
S-54		PKC
S-57	PKA	
S-58		CKII
S-58	RSK	
S-58	PKA	PKA
S-63	CDK	
T-64	PKC	
T-75	PKC	PKC
T-75		CK2
T-76	PKC	
S-81	CDK	
T-85	PKC	PKC

B

HPV18 E1^E4

```
        9            20 22,23                      40
MADPEVPVTTRYPPPSSSNSYSTPPHRIPAPCPWAPQRPTARRR


    51  54  57,58   63,64        75,76   81    85
LLHDLDTVDSRRSSIVDLSTHFSVQLHLQATTKDGNSVVVTLRL
```

Figure 3.1: Putative phosphorylation sites within HPV18 E1^E4. The primary amino acid sequence of HPV18 E1^E4 was scanned by the ScanProsite and NetPhosK algorithms which predict phosphorylation sites dependent upon the presence of kinase consensus sites within a protein. Phosphorylation sites were predicted for: Protein kinase C (PKC), cAMP-dependent protein kinase A (PKA) and Casein kinase II (CKII), Cyclin-dependent protein kinase (CDK), Mitogen-activated protein kinase (MAPK), and Ribosomal s6 kinase (RSK). **(A)** Table of predicted phosphorylation sites. **(B)** Putative phosphoacceptor sites are indicated in red typeface with numbers above the sequences indicating the position within the HPV18 E1^E4 sequence.

requirement for MAPK phosphorylation is only an S/TP di-peptide site (Bardwell, 2006). HPV18 E1^E4 contains a single ^{23}TP^{24} MAPK site with a predicted phospho-acceptor residue at threonine 23, making it possible that this residue is targeted by multiple kinases.

Whilst eight putative phosphorylation sites for PKC were predicted, further analysis revealed that only a subset of these residues were contained within a consensuses site for PKC (S/TXR/K, K/RXXS/T and K/RXS/T) threonine residues at ^{9}TTR^{11}, ^{40}TAR^{43}, ^{75}TTK^{77} and ^{85}TLR^{87}. Predicted sites for CKII phosphorylation were also identified within HPV18 E1^E4 at ^{58}SIVD^{61} and ^{75}TTKD^{78} which are good consensus sequences for the kinase (S/TXXD/E). Serine residue 58 was also predicted as a site for RSK phosphorylation, however the characterised consensus motif for this kinase is RXRXXS/T, and the predicted target residue of HPV18 E1^E4 is located within the sequence ^{53}DSRRSS^{58}, which only weakly corresponds to the consensus motif indicating that HPV18 E1^E4 may be an unlikely substrate for the kinase. However this region of the E1^E4 protein (^{55}RRSS^{58}) is a good consensus site for PKA which targets RXS/T and RR/KXS/T sites, with both serine 57 and 58 predicted to be target residues for PKA phosphorylation. In addition HPV18 E1^E4 also contains a consensus site for PKA within the sequence ^{38}RPT^{40}.

3.4   HPV18 E1^E4 is a substrate for epithelial cell kinases in crude cell extract

The presence of a number of kinase consensus sites within the HPV18 E1^E4 protein sequence suggests that the protein is a substrate for phosphorylation by cellular kinases. To establish whether HPV18 E1^E4 can undergo phosphorylation, the protein was expressed as a GST-fusion protein (GST-HPV18 E1^E4) and incubated with total cellular lysate derived from HeLa cells in an *in vitro* kinase reaction (Figure 3.2 A). A major phosphorylated species corresponding to the GST-HPV18 E1^E4 protein was detected which was absent from control
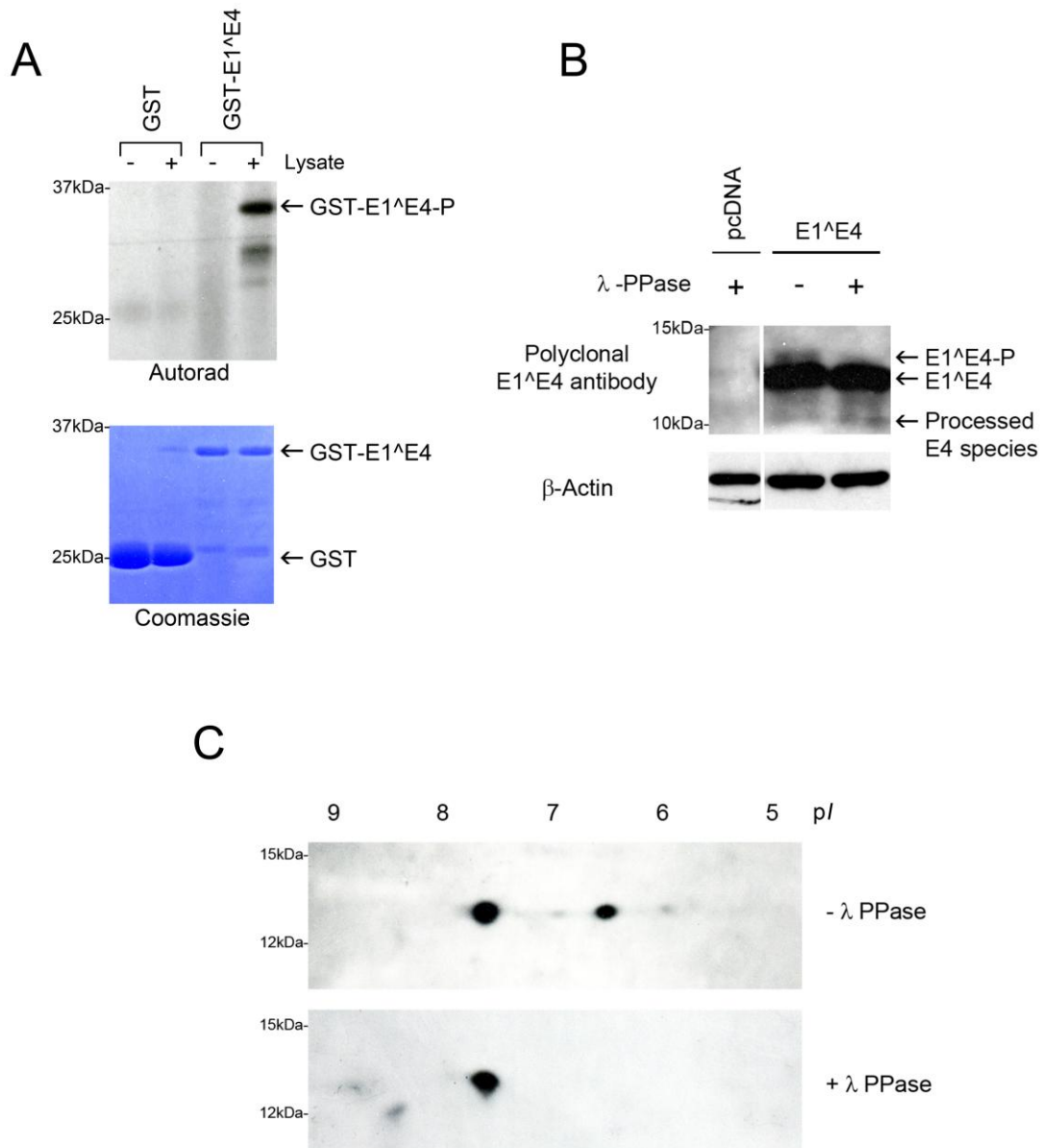
Figure 3.2: HPV18 E1^E4 is phosphorylated both *in vitro* and *in vivo*. (A) Phosphorylation of HPV18 E1^E4 by crude cell extract. Bacterially-expressed GST or GST-HPV18 E1^E4 proteins were incubated with total cell lysate derived from HeLa cells during an *in vitro* kinase assay. Whilst GST is not a substrate for the cellular kinases present, GST-HPV18 E1^E4 was phosphorylated. (B) Western blot analysis of HPV18 E1^E4 expressed from a codon-optimised (CO) cDNA in 293T cells. In addition to a major E1^E4 protein at 11.7 kDa, multiple E4 species were detected with a polyclonal E1^E4 antibody. Treatment of cell lysate with the serine- threonine-specific lambda protein phosphatase (λPPase) resulted in a loss of a slower migrating E4 species. (C) Extracts from 293T cells transiently expressing CO HPV18 E1^E4 cDNA were separated by two-dimensional gel electrophoresis and analysed with a polyclonal E1^E4 antibody. Upon treatment of lysate with λPPase only a major E4 species could be observed and the multiple E4 species observed in untreated lysate were absent. Isoelectric points (p*I*) are indicated above the gel.

reactions containing GST or in the absence of cellular lysate. HeLa cells were chosen to allow analysis of E1^E4 phosphorylation within HPV18 positive cells. This finding demonstrates that HPV18 E1^E4 is a substrate for cellular kinases and suggests that the protein may undergo phosphorylation within epithelial cells.

3.5 HPV18 E1^E4 is phosphorylated *in vivo*

Expression of E1^E4 proteins of several anogenital types can prove to be problematic as considerably low levels of the these proteins have been reported to be generated by transient transfection (Roberts et al., 1993; Raj et al., 2004). This common characteristic of anogenital E1^E4 proteins may be due to the presence of inhibitory elements which alter RNA stability or secondary structure or alternatively may be due to instability of the E1^E4 protein. It has also been hypothesised that this may reflect the rare codon usage of HPV DNA in comparison with that of its host cell (Roberts et al., 1993; Zhou et al., 1999). In an effort to try to overcome these issues, a codon-optimised HPV18 E1^E4 expression construct was generated by Dr. I Bell in which a number of codons were converted within HPV18 E1^E4 cDNA (coE1^E4) to those more commonly used in genes of the human genome (Appendix III). A greatly enhanced expression of the E1^E4 protein was observed within epithelial cells transfected with coE1^E4, in comparison with unmodified E1^E4 cDNA, with the greatest expression of coE1^E4 observed in 293T cells, a human embryonic kidney cell line containing the SV40 Large-T antigen, therefore these cells were used in all future experiments examining E1^E4 expression by Western blot analysis.

Lysates from 293T cells transfected with codon-optimised HPV18 E1^E4 were subjected to SDS-PAGE and major E1^E4 species of 11.7 kDa was detected by Western blotting with a polyclonal E1^E4 antibody. A faster migrating species of 10.5 kDa was also detected which

corresponds to an N-terminally truncated E4 protein which has been characterised in Chapter 5. In addition to this, a slower migrating species could be detected just above the major E1^E4 protein (Figure 3.2 B). Protein phosphorylation often causes slower migration of the protein in SDS-PAGE gels due to the altered mass to charge ratio as a result of the additional negative charges of the phosphate group. The higher molecular mass species of E1^E4 observed may therefore be indicative of E1^E4 phosphorylation. To establish whether this species corresponds to a phosphorylated E1^E4 product, lambda protein phosphatase (λ-PPase), a phosphatase which removes phosphate groups from serine, threonine or tyrosine residues was used to treat E1^E4 lysate. Incubation of the E1^E4 lysate with λ-PPase, resulted in a loss of detection of the slower migrating band indicating that this is a phosphorylated E1^E4 species.

Phosphorylated E1^E4 appears to be a minor subset of E4 species as the corresponding band is significantly lower in intensity relative to the major E1^E4 species. However this experiment was not readily repeatable as the slower migrating band was not always visible, and it was difficult to establish under which conditions this species could be best detected. Improved resolution of E4 proteins could address this issue.

Two-dimensional (2D) gel electrophoresis separates proteins according to both mass and charge, and can be used to detect protein phosphorylation. Phosphorylation results in the transfer of a negatively charged phosphate group to the hydroxyl group of the target residue, altering the overall charge, and therefore the isoelectric point (p*I*) of the protein. Two-dimensional gel electrophoresis resolves proteins in the first dimension by isoelectric focusing (IEF), separating proteins by their net charge, whereby differentially phosphorylated proteins with varying isoelectric points will be separated. Proteins are then resolved by SDS-PAGE in

which the proteins are separated by mass. In order to analyse HPV18 E1^E4 phosphorylation, 2D gel electrophoresis was carried out with lysate from 293T cells transfected with the codon-optimised E1^E4 construct. Four E4 species of differing isoelectric points and of varying intensities could be detected with a polyclonal E1^E4 antibody with the predominant protein spot having an isoelectric point of approximately p*I*7.8 (Figure 3.2 C). A less intense spot was also detected with a p*I* of approximately 6.5 and in addition to two faint spots with approximate p*I*s of 7 and 6 indicating that E4 exists as differentially charged species. All E4 spots except the major species were lost upon prior treatment of the lysate with λ-PPase, indicating that these species are differentially phosphorylated E4 proteins.

A number of kinase consensus sites have been identified within the HPV18 E1^E4 protein sequence (Figure 3.1) and the corresponding protein kinases may account for the phosphorylation of the E1^E4 protein observed in epithelial cells. To identify kinases that target HPV18 E1^E4 for phosphorylation the ability of several kinases to phosphorylate the protein was assessed by *in vitro* kinase assay using bacterially-expressed GST-HPV E1^E4 fusion proteins.

3.6   **HPV18 E1^E4 is phosphorylated by PKA at serine 58 *in vitro*.**

HPV18 E1^E4 contains a consensus site for PKA within the sequence ^{38}RP$\underline{T}^{40}$ with threonine 40 acting as the phospho-acceptor residue, and two overlapping consensus sites within the residues ^{55}RR$\underline{SS}^{58}$ in which both serine residues may be targeted for phosphorylation by PKA. Since E1^E4 proteins from HPV types 1 and 11 have been shown to be phosphorylated by PKA and a number of other E1^E4 proteins also contain consensus sites for the kinase (Figure 3.2), phosphorylation by PKA may be a conserved feature amongst E1^E4 proteins. To ascertain whether HPV18 E1^E4 is a substrate for PKA phosphorylation, an *in vitro* kinase

assay was carried out using a GST-HPV18 E1^E4 recombinant protein purified from bacteria (Figure 3.3 A). Incubation of GST-HPV18 E1^E4, but not GST alone with PKA resulted in incorporation of radioactive isotope into the protein, indicating that HPV18 E1^E4 is phosphorylated by PKA *in vitro*.

To identify the residue within HPV18 E1^E4 targeted by PKA, a GST E4 mutant protein was generated by site-directional mutagenesis of the pGEX-HPV18 E1^E4 expression construct in which both serine residues of the ^{55}RR$\underline{SS}^{58}$ double PKA motif were mutated to asparagine (serine 57+58 -> asparagine, S57N,S58N) (These codon changes were chosen as in the context of the total HPV18 genome they are the only mutations at these residues that can be made without altering the ORF of the E2 gene see Appendix I). The S57N,S58N mutant severely abrogated phosphorylation by PKA comparable to that of the wild type GST-HPV18 E1^E4, indicating that this consensus site is indeed the target site for PKA phosphorylation *in vitro* and the second consensus site: ^{38}RP$\underline{T}^{40}$ is therefore redundant (Figure 3.3 A). To further elucidate which residues are targeted for phosphorylation within the overlapping PKA consensus sites, mutations of each serine residue were generated separately (serine 57 -> asparagine, S57N and serine 58 -> asparagine, S58N). E1^E4 phosphorylation by PKA was retained at comparable levels to that of the wild type protein by the S57N mutant (Figure 3.3 B), however in contrast, mutation of S58N abrogated phosphorylation of the GST-HPV18 E1^E4 protein, demonstrating that serine 58 is the target residue for phosphorylation of HPV18 E1^E4 by PKA *in vitro*.
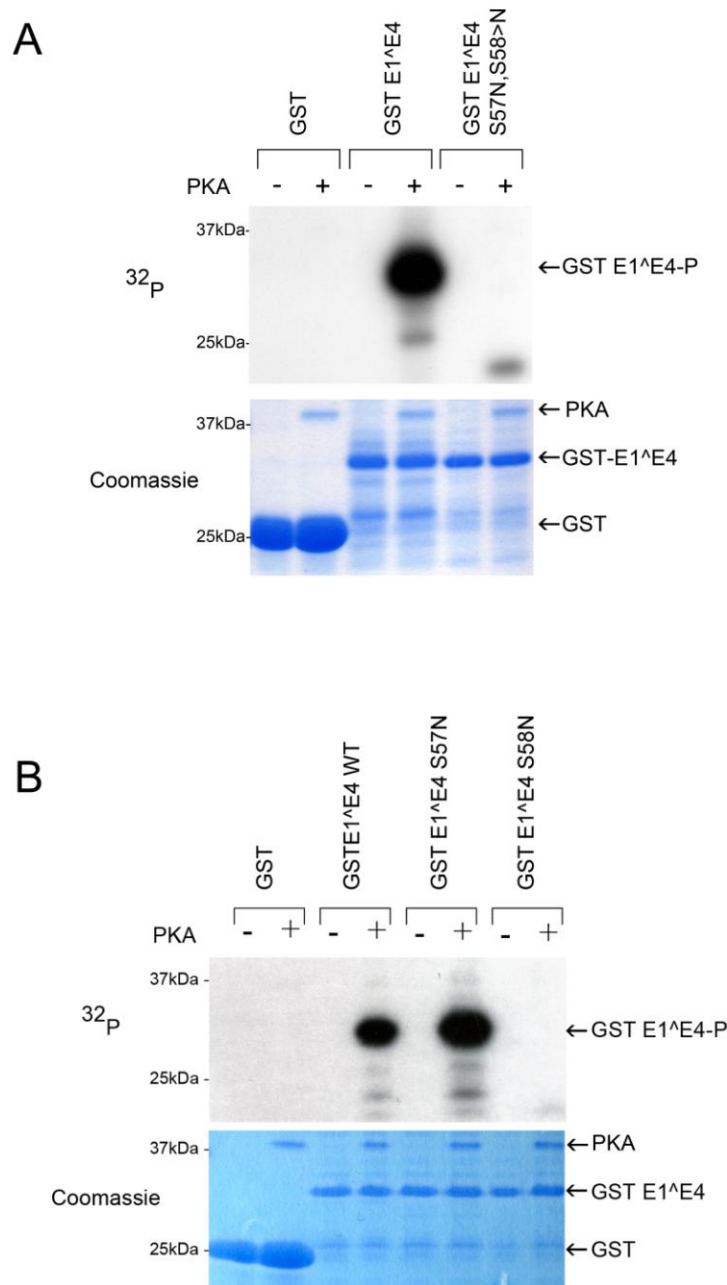
Figure 3.3: HPV18 E1^E4 is phosphorylated at serine 58 by PKA *in vitro*. (A) *In vitro* phosphorylation of GST and GST-HPV18 E1^E4 proteins by recombinant cAMP-dependent Protein Kinase A (PKA) catalytic subunit. GST-HPV18 E1^E4 is phosphorylated by PKA but GST and a GST-HPV18 E1^E4 in which serine 57 and 58 are substituted (S57N,S58N) are not effective substrates for this kinase. Equal levels of GST proteins are confirmed by the Coomassie stained gel and the data is representative of two experiments. (B) To identify the target site within HPV18 E1^E4 for PKA phosphorylation, an *in vitro* kinase assay with PKA was carried out using GST HPV18 E1^E4 proteins containing substitutions at S57 (S57N) and S58 (S58N). GST-HPV18 E1^E4 S57N is an effective substrate for PKA phosphorylation whilst the S58N mutant GST-fusion protein is not phosphorylated by this kinase. The Coomassie stained gel reveals equal loading of GST proteins and data is representative of three experiments.

3.7   **HPV18 E1^E4 is not a substrate for p42 MAPK *in vitro*.**

HPV18 E1^E4 contains a single consensus site at [23]TP[24] which meets the minimal requirements of an S/TP di-peptide site for MAPK phosphorylation. To determine whether GST-HPV18 E1^E4 is a substrate for this kinase, an *in vitro* kinase assay was carried out, and GST-HPV16 E1^E4 as a known substrate (Wang et al., 2009), was also included in the assays as a positive control (Figure 3.4). Following incubation of GST proteins with recombinant p42 MAPK, a major phosphorylated band corresponding to GST-HPV16 E1^E4 was detected. However, incorporation of radioactive isotope was not detected in either GST or GST-HPV18 E1^E4 proteins indicating that the GST-HPV18 E1^E4 fusion protein is not a substrate for p42 MAPK *in vitro*. Whilst Coomassie staining indicates that more GST-HPV16 E1^E4 protein was added to the experiment than with GST-HPV18 E1^E4, the high level of radioactive isotope incorporated into the GST-HPV16 E1^E4 protein, with no detectable phosphorylation of the GST-HPV18 E1^E4 protein, suggests that the latter protein may not be a target for MAPK *in vivo*. This would demonstrate that type-specific differences occur between E1^E4 proteins and could suggest that these proteins may act or be regulated in different ways. However it is possible that this result may reflect an inability of the kinase to phosphorylate the HPV18 E1^E4 protein due to the presence of the GST tag precluding access to the target site.

3.8   **HPV18 E1^E4 is phosphorylated *in vitro* at threonine 23 by CDK-cyclin complexes**

HPV18 E1^E4 has a single consensus site used by CDKs [23]TPPH[26] (S/TPXR/K) in which threonine 23 is the putative target residue. To assess whether HPV18 E1^E4 is a substrate for CDK, an *in vitro* kinase assay was carried out with GST-HPV18 E1^E4. In the presence of
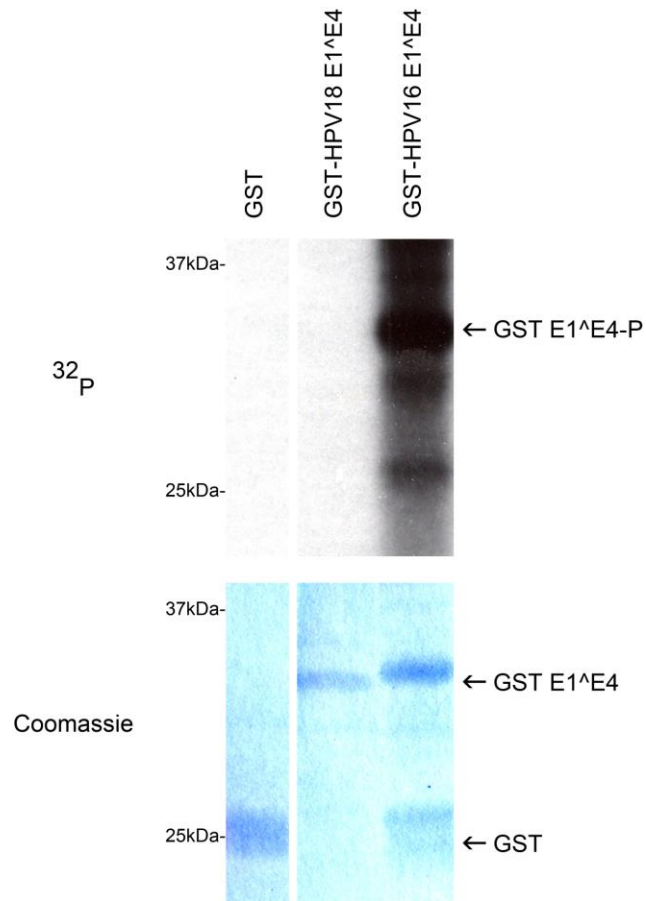
Figure 3.4: HPV18 E1^E4 is not a substrate for MAPK phosphorylation *in vitro. In vitro* phosphorylation of GST, GST-HPV18 E1^E4 and GST-HPV16 E1^E4 proteins by recombinant p42 MAPK, with a Coomassie stained gel demonstrating loading of the GST proteins. Whilst GST-HPV16 E1^E4 is effectively phosphorylated by MAPK, GST and GST-HPV18 E1^E4 proteins are not substrates for the kinase. This result is representative of three experiments.

CDK2-cyclin A, a phosphorylated band corresponding to the GST-HPV18 E1^E4 fusion protein was detected which was not present in the GST reaction, demonstrating that HPV18 E1^E4 is phosphorylated by CDK2-cyclin A *in vitro* (Figure 3.5 A).

To identify the target residue for CDK2-cyclin A, the assay was repeated in the absence of radioactive isotope and the Coomassie-stained band corresponding to GST-HPV18 E1^E4 was extracted from the gel and analysed by the mass spectrometry method of Neutral Loss - Electron Capture Disassociation (NL-ECD MS/MS) (Sweet et al., 2006) by School of Biosciences, University of Birmingham. This method allows determination of sites of phosphorylation and assigns a probability value of the likelihood of phosphorylation at that site. Briefly, and in overview, a tryptic digest of the GST-HPV18 E1^E4 sample was subjected to liquid chromatography and the peptides eluted via a nanospray source into a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer where the mass of each peptide is analysed. The ions were then subjected to collision-induced disassociation (CID) which results in a loss of phosphoric acid (H_3Po_4) from phosphothreonine and phosphoserine residues and a loss of a phosphate group (HPo_3) from phosphotyrosine residues, which correspond to the loss of 98 and 80Da respectively. These neutral losses from a phosphor peptide can be detected by MS/MS and if observed, the precursor is subjected to ECD where the ions are irradiated with electrons, fragmenting the peptide backbone but retaining any PTM. The mass spectrometry data is then scanned for the presence of site-determining ions which provide evidence of phosphorylation occurring at that site. A neutral loss from the E4 peptide: [12]YPLLSLLNSYS**T***PPHR[27] was identified, and site-determining ions were observed which were highly indicative of threonine 23 phosphorylation with an ambiguity score (Ascore) value of 28.5 obtained for this residue. The Ascore, is a probability-based score measuring the probability of correct phosphorylation localisation
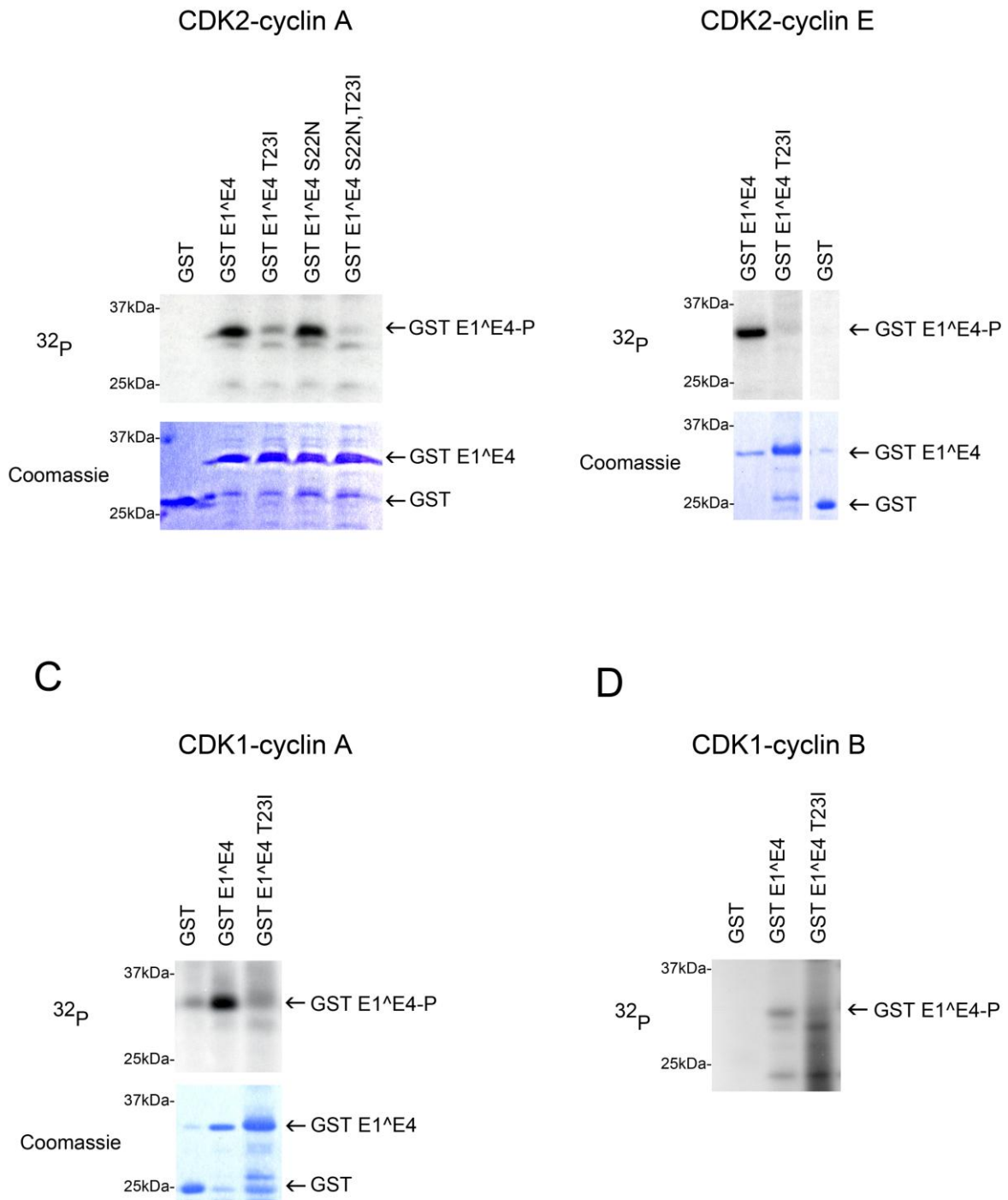
Figure 3.5: Phosphorylation of HPV18 E1^E4 at T23 by CDK-cyclin complexes. *In vitro* phosphorylation of GST and GST-HPV18 E1^E4 proteins by: **(A)** CDK2-cyclin A, **(B)** CDK2-cyclin E, **(C)** CDK1-cyclin A and **(D)** CDK1-cyclin B. Coomassie stained gels indicate loading of GST proteins and each experiment is representative of at least three independently performed kinase assays. Whilst GST-HPV18 E1^E4 is phosphorylated *in vitro* by both CDK1 and CDK2, phosphorylation of the GST-HPV18 E1^E4 T23I mutant protein by these kinases is severly reduced, indicating that T23 is a target residue within HPV18 E1^E4 for CDK phosphorylation.

based on the intensity and frequency of site-determining ions in MS/MS spectra compared with the probability of their detection by chance alone (Schwartz & Gygi, 2005; Beausoleil et al., 2006). Sites with Ascores >19 correspond to sites have been localised with P<0.01 which is defined as near certainty, therefore phosphorylation of HPV18 E1^E4 by CDK2-cyclin A is highly likely to occur at T23. Site-determining ions were also identified corresponding to serine 22 phosphorylation, making it possible that T23 may be the preferential site for CDK phosphorylation and that S22 may be a secondary and minor site for phosphorylation. However phosphorylation at the S22 site was deemed unlikely by NL-ECD MS/MS analysis.

To confirm the results obtained by mass spectrometry analysis, GST-HPV18 E1^E4 mutant proteins were generated using substitutions which retain the coding capacity of the E2 ORF in the context of the total HPV18 genome (Appendix I). Substitutions of threonine 23 (threonine 23 -> isoleucine, T23I), the adjacent serine residue (serine 22 -> asparagine, S22N) and, to allow consideration that both residues may be phosphorylated, a double substitution incorporating both mutations (S22N,T23I) were generated, and these mutant GST proteins were subsequently assayed for their ability to undergo CDK2-cyclin A phosphorylation *in vitro* (Figure 3.5 A). Phosphorylation of the S22N, T23I double mutant GST-fusion protein was severely abrogated compared with that of the wild type GST-HPV18 E1^E4 protein indicating that phosphorylation does indeed occur at one or both of these residues *in vitro*. Whilst the S22N protein was phosphorylated by CDK2-cyclin A to a similar extent as that of the wild type protein, phosphorylation of the T23I mutant protein was severely reduced, indicating that threonine 23 and not serine 22, is the phospho-acceptor residue *in vitro* (Figure 3.5A). This result was repeated in two subsequent experiments and the observation is consistent with the mass spectrometry data in the conclusion that threonine 23 is the target residue for HPV18 E1^E4 *in vitro* phosphorylation by CDK2-cyclin A.

To assess whether HPV18 E1^E4 is phosphorylated by other CDKs, *in vitro* kinase assays were carried out using purified forms of CDK2-cyclin E, CDK1-cyclin A and CDK1-cyclin B with GST-HPV18 E1^E4 proteins as substrates (Figure 3.5 B-D). The wild type HPV18 E1^E4 GST-fusion protein was phosphorylated by all CDK-cyclin complexes assayed, whilst phosphorylation of the T23I mutant GST-E1^E4 protein was severely reduced in all reactions, indicating that T23 is a shared phospho-acceptor residue for CDK-cyclin complexes *in vitro*. The extent of phosphorylation of GST-E1^E4 by CDK1-cyclin B was repeatedly poor (Figure 3.5 D), despite efficient phosphorylation of the CDK substrate: Histone H1 (data not shown).

3.9 Effect of HPV18 E1^E4 phosphorylation on protein stability

Phosphorylation can substantially alter the structure of a protein and may modulate the stability or turnover of the protein. Given that CDKs and PKA target HPV18 E1^E4 residues 23 and 58 respectively during *in vitro* phosphorylation assays, it is possible that phosphorylation of these residues may alter the stability of the E1^E4 protein within cells. To establish whether phosphorylation of HPV18 E1^E4 results in an alteration to the steady state levels of the protein, phospho-acceptor residues within the codon–optimised HPV18 E1^E4 expression construct were mutated to alanines, with the CDK target residue (threonine 23 -> alanine, T23A), the PKA target residue (serine 58 -> alanine, S58A) and both residues mutated simultaneously (T23A, S58A) to prevent phosphorylation at these sites (Figure 3.6 A). In an attempt to imitate phosphorylation, residues can be mutated to aspartic acid, which introduces a negative charge and can mimic constitutive phosphorylation at the mutated site. A CDK phosphorylation mimic (T23D), a PKA phosphorylation mimic (S58D) and a double mutation (T23D, S58D) were also generated within the E1^E4 cDNA sequence (Figure 3.6 A).

A

CDK Site	S/TPXR/K
E1^E4	^{23}TPPH^{26}
E1^E4 T23A	APPH
E1^E4 T23D	DAPH

PKA Site	R/KRXS/T
E1^E4	^{55}RRSS^{58}
E1^E4 S58A	RRSA
E1^E4 S58D	RRSD

B
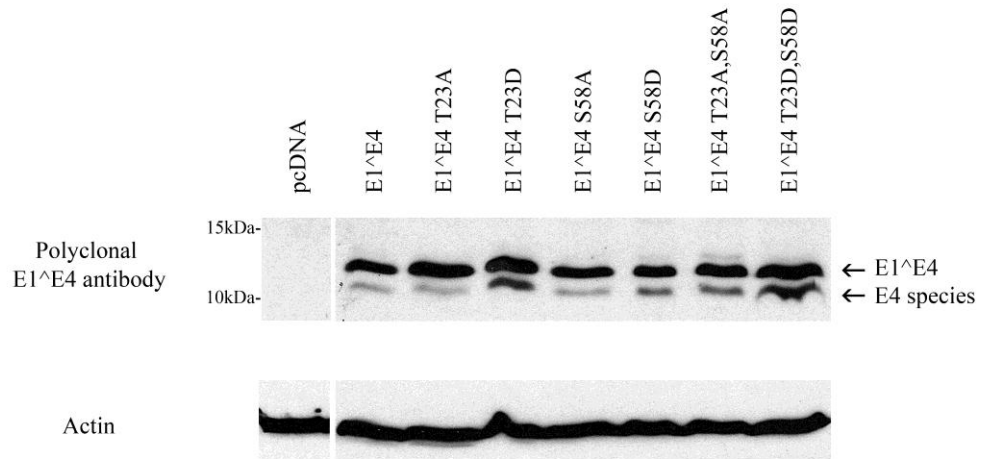
Figure 3.6: Steady-state levels of E1^E4 phosphoacceptor mutant proteins. **(A)** Substitution mutations of threonine 23 CDK and serine 58 PKA phosphoacceptor residues with alanine or aspartic acid. Mutations were introduced into the codon-optimised HPV18 E1^E4 expression vector. **(B)** Following transfection of 293T cells with the E1^E4 phospho mutants, Western blot analysis was performed with a polyclonal E1^E4 antibody and β-actin as a loading control. This result is representative of two independent experiments.

To examine the contribution of these residues to the steady state level of the E1^E4 protein the mutants proteins were expressed in 293T cells, and the cellular lysate was subjected to Western blot analysis with a polyclonal E1^E4 antibody and the protein levels compared (Figure 3.6 B). Whilst slight variations were observed between the E1^E4 mutant proteins, no significant differences were observed between experiments compared with the wild type E1^E4 protein indicating that an inability to undergo phosphorylation, or mimicking of phosphorylation, does not impact upon protein levels within these cells. In addition to the major E1^E4 species a smaller E4 species was also observed, representing a truncated polypeptide. The presence of this cleaved E4 protein in the phospho-acceptor mutant proteins indicates that phosphorylation at these sites is not required for E1^E4 proteolysis.

3.10   Effect of HPV18 E1^E4 phosphorylation on the biological action of cellular localisation

Phosphorylation can often modulate the localisation of a protein and it is possible that the cellular distribution of the HPV18 E1^E4 protein may be altered by CDK or PKA phosphorylation. Typical cellular distributions of the HPV18 E1^E4 wild type protein can be observed in Figure 3.7. The E1^E4 protein of HPV18 frequently displays a filamentous distribution, forming extensive cytoplasmic fibrous networks (Figure 3.7 A,a). These networks can progressively collapse, and a partially collapsed E1^E4 filamentous network is observed in Figure 3.7 A,b. Collapsed E1^E4 networks produce juxta-nuclear bundles (Figure 3.7 A, c). Co-staining of cells transfected with HPV18 E1^E4 with the polyclonal E1^E4 and monoclonal keratin 18 (K18) antibodies confirmed colocalisation of HPV18 E1^E4 filamentous and collapsed networks with the keratin intermediate filaments (IF) (both
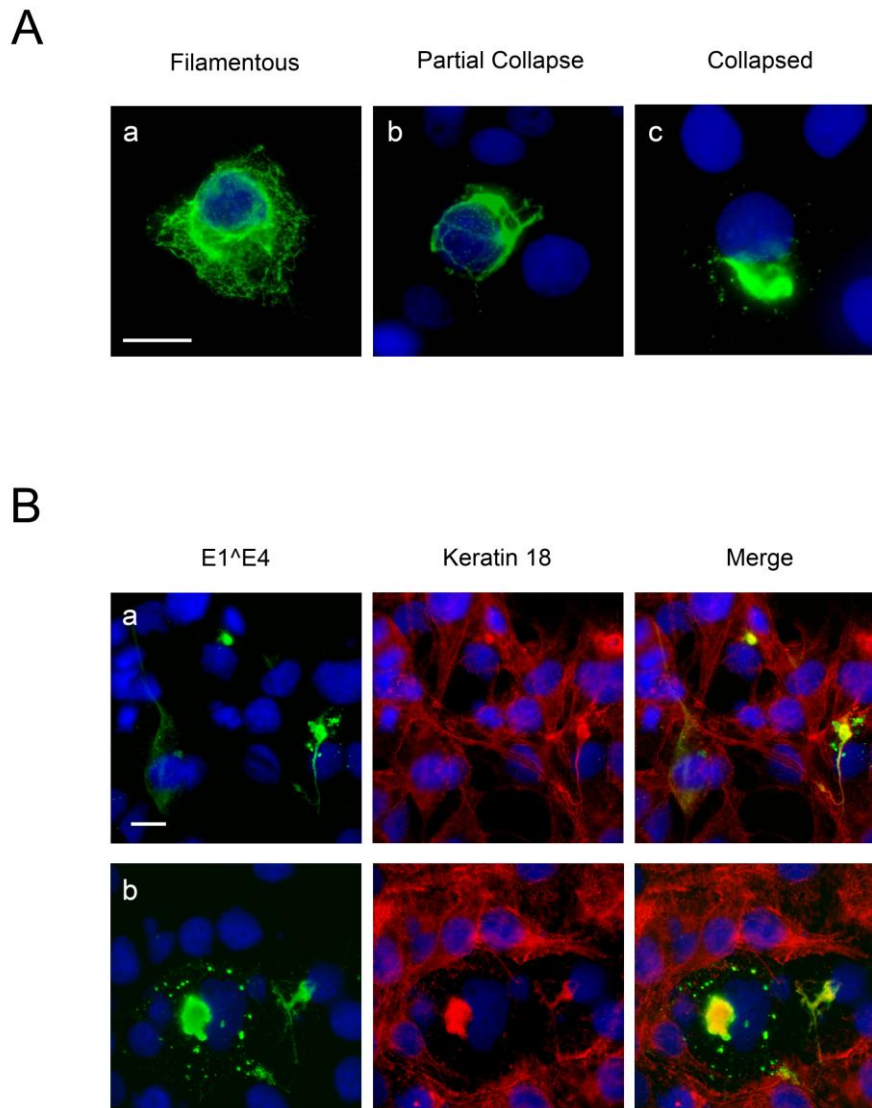
Figure 3.7: Typical distributions of the HPV18 E1^E4 protein within epithelial cells. Following a 36 hr transfection with coE1^E4, COS-1 cells were fixed in 4% paraformaldehyde prior to permeabilisation in acetone, before staining with a polyclonal E1^E4 anti-rabbit antibody (A), or dual-staining with a polyclonal E1^E4 anti-rabbit antibody and a keratin 18 anti-mouse antibody (B). Immune complexes were visualised with an anti-rabbit Alexa Fluor 488 antibody (green) and an anti-mouse Alexa Fluor 596 antibody (red) and counter-stained with Dapi for detection of nuclei (blue). **(A)** HPV18 E1^E4 typically dispays the following distributions, **(a)** cytoplasmic filamentous network, **(b)** partially collapsed filamentous distribution, or **(c)** collapsed juxta-nuclear bundle. **(B)** HPV18 E1^E4 co-localises with the keratin IF network. **(a)** Fibrous cytoplasmic E1^E4 networks and collapsed filaments co-localise with cytokeratin. **(b)**Within cells with collapsed E1^E4 networks the E1^E4 protein can also be commonly observed as punctate, bright cytoplasmic foci.   Bar, 10 μm. Observations are representative of at least three independently performed experiments.

phenotypes observed in Figure 3.7B,a), confirming findings from a previous study (Nakahara et al., 2002). Within cells with collapsed E1^E4 networks the HPV18 E1^E4 protein can also be commonly observed as punctate, bright cytoplasmic foci in addition to the collapsed juxta-nuclear bundle (Figure 3.7 B,b).

To establish whether an inability to undergo phosphorylation or mimicking of phosphorylation at the PKA and CDK *in vitro* phospho-acceptor sites alters the localisation of HPV18 E1^E4, mutant proteins were expressed in SVJD cells, simian virus 40 [SV40]-immortalised human epidermal keratinocytes, and the distribution of these protein was examined by immunofluorescence microscopy (Figure 3.8). SVJD cells were chosen for this experiment as these cells were available at the time of performing these experiments. Wild type E1^E4 was typically observed to co-localise with keratin within juxta-nuclear bundles. Expression of T23A, S58A and T23A,S58A E1^E4 mutant proteins in which these phospho-acceptor residues were substituted with an alanine, displayed a phenotype indistinguishable from that of the wild type E1^E4 protein (Figure 3.8 b). Introduction of an aspartic acid to either the CDK or PKA target sites, or both residues simultaneously, similarly did not alter the distribution of the E1^E4 protein (Figure 3.8 c).

3.11   Analysis of E1^E4 threonine residue 23 phosphorylation with a phospho-specific E1^E4 antibody

Phospho-specific antibodies can be valuable tools in the analysis of *in vivo* protein phosphorylation and can be used to discriminate between phosphorylated and un-phosphorylated proteins. Use of an antibody which specifically recognises phosphorylated forms of the E1^E4 protein could prove to be useful for examining the role of HPV18 E1^E4
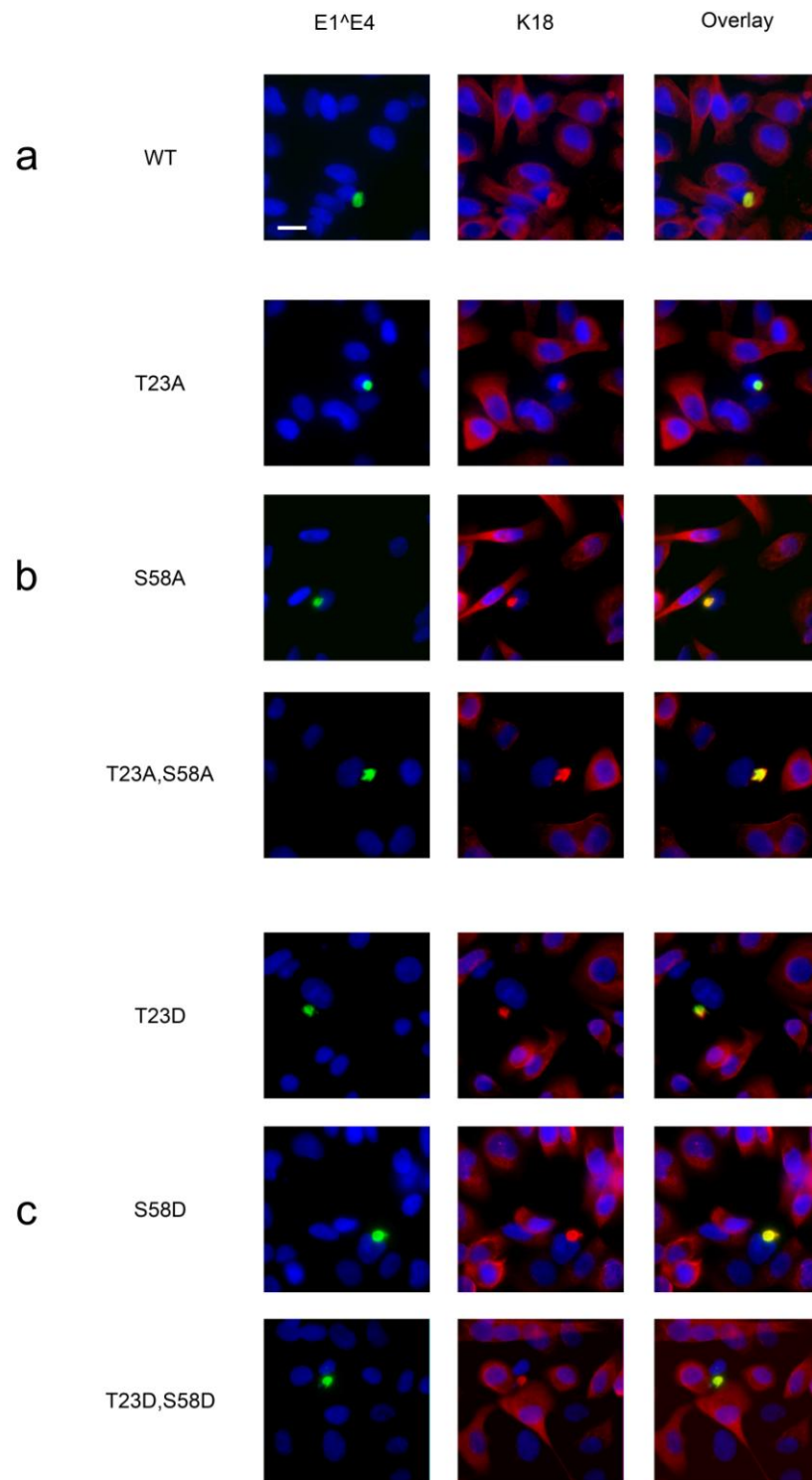
Figure 3.8: Localisation of E1^E4 phosphoacceptor mutant proteins. Immunofluorescence microscopy was performed following transfection of SVJD cells with the phosphoacceptor substitution mutations of threonine 23 CDK and serine 58 PKA phosphoacceptor residues with alanine or aspartic acid. Cells were dual stained for E1^E4 (green) and keratin 18 (red) and counterstained with Dapi for detection of nuclei (blue). This was repeated in two independently performed experiments with similar observations. Bar, 10 μm.

phosphorylation in cells. Therefore a phospho-specific antibody was generated in rabbits to detect phosphorylation of threonine 23 of the HPV18 E1^E4 protein (Eurogentec, Chapter 2: Materials and Methods, 2.6.12.3). An antibody was raised against the E1^E4 phospho peptide sequence: ^{18}LNSYST(PO$_3$H$_2$)PPHRIPA^{30} (E1^E4 T23-P) and in addition a control antibody was raised against the non-phosphorylated E1^E4 epitope: ^{18}LNSYSTPPHRIPA^{30} (E1^E4 non-P).

To confirm the specificity of the E1^E4 antibodies, GST-E1^E4 proteins were phosphorylated by CDK2-cyclin A *in vitro*, performed in the absence of radioactive isotope, and detection of the protein was examined by Western blot analysis with the E1^E4 T23-P and the E1^E4 non-P (Figure 3.9). Whilst the E1^E4 non-P antibody was specific for the GST-E1^E4 proteins but not GST alone, the antibody did not distinguish between the wild type and T23I mutant proteins, or between assays carried out in the absence or presence of CDK2-cyclin A, with equal recognition of these GST-fusion proteins. The E1^E4 T23-P antibody however had the highest affinity for the phosphorylated wild type GST-E1^E4 and whilst some detection of un-phosphorylated wild type and the T23I mutant protein was observed, this was at a greatly reduced level. A degree of non-specific reactivity is often observed with phospho-specific antibodies.

Next, to investigate HPV18 E1^E4 phosphorylation at threonine 23 within a physiologically relevant expression system, human foreskin keratinocytes (HFK) stably maintaining HPV18 genomes as episomes were induced to differentiate by suspension of cells in semi-solid media. This system is detailed within Chapter 6 and has been successfully used within our laboratory to enable analysis of differentiation-dependent stages of the HPV18 life cycle in which E1^E4 expression is observed (Wilson et al., 2007). An HPV18 mutant cell line had
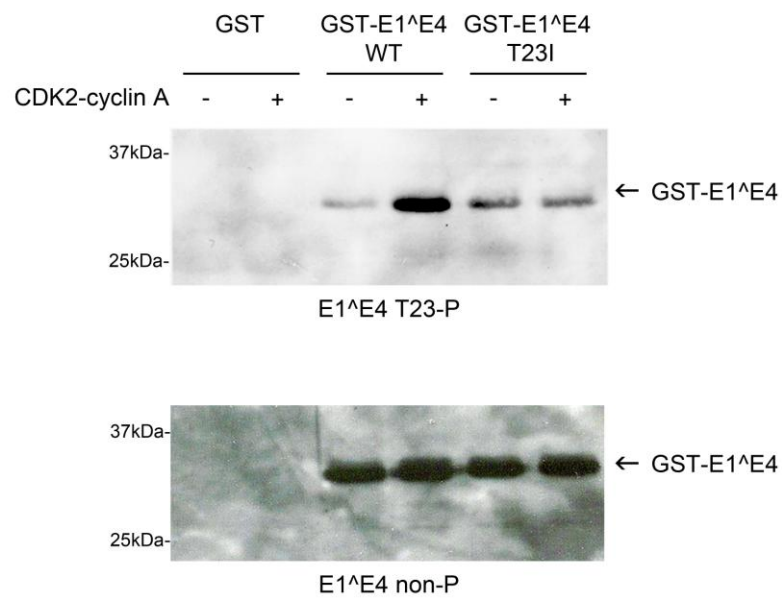
Figure 3.9 An E1^E4 threonine 23 phospho-specific antibody for detection of HPV18 E1^E4 phosphorylation. CDK2-cyclin A *in vitro* phosphorylation of indicated GST-HPV18 E1^E4 fusion proteins. Western blot analysis was subsequently performed with antibodies raised against an E1^E4 T23 phospho peptide (E1^E4 T23-P) or a non phospho E1^E4 peptide (E1^E4 non-P). Western blots are representative of three seperate experiments.

previously been generated by Dr G. L. Knight, in which the codon encoding for threonine residue 23 of the E1^E4 protein is mutated to an isoleucine, a substitution which alters the E4 ORF but introduces a silent mutation into the overlapping E2 coding sequence. Upon differentiation of keratinocytes harbouring HPV18 genomes, expression of E1^E4 proteins was examined by Western blot analysis. Following 48 h suspension in semi-solid media, E1^E4 proteins from both wild type and the T23I E1^E4 mutant cell lines were detectable with a monoclonal E1^E4 antibody but neither were detectable with the T23 phospho-specific E1^E4 antibody (data not shown). Whilst the T23 E1^E4 phospho-specific antibody had detected phosphorylated GST-E1^E4 protein, the inability to detect E1^E4 proteins within differentiating keratinocytes may be due to the considerably reduced level of E1^E4 protein expression from the HPV18 genome in comparison with bacterial expression. Alternatively the HPV18 E1^E4 protein may not be a target for CDK phosphorylation within differentiating keratinocytes or phosphorylation may occur at a later stage of the virus life cycle, than simulated by suspension of cells in semi-solid media. It is possible that the E1^E4 threonine 23 phospho-specific antibody may allow improved detection of E1^E4 phosphorylation within cells by immunofluorescence and future work could examine this possibility.

Discussion

It has been demonstrated for the first time that HPV18 E1^E4 can exist as a phosphoprotein within cells and that this protein is a substrate for *in vitro* phosphorylation by a number of cellular kinases. Mutational analysis has shown that serine residue 58 of HPV18 E1^E4 is the target of *in vitro* phosphorylation by PKA, with threonine residue 23 the phospho-acceptor residue for a number of CDKs including: CDK2-cyclin A, CDK2-cyclin E, CDK1-cyclin A and CDK1-cyclin B. However the HPV18 E1^E4 protein was not shown to be a substrate for

MAPK. Phosphorylation of E1^E4 proteins by cellular kinases is likely to regulate the function of the protein and may act to induce or impair the ability of the E1^E4 protein to form protein-protein interactions. Examination of HPV18 E1^E4 CDK and PKA phospho-acceptor mutants however revealed a comparable cellular distribution to that of the wild type protein and the stability of these proteins was unaltered in epithelial cells.

3.11   PKA phosphorylation of E1^E4 proteins

Increased levels of the secondary messenger cAMP in response to extracellular signalling can exert diverse effects on cellular proliferation, migration and differentiation. Many of these effects are mediated by PKA whose activity is dependent upon binding of cAMP to regulatory subunits of the kinase to release and activate the catalytic subunits. PKA has been implicated in induction of keratinocyte differentiation, with increased expression of terminal differentiation markers within keratinocytes observed in response to increased intracellular cAMP levels in a PKA-dependent manner (Mammone et al., 1998). It is possible that PKA phosphorylation may act to modulate the function of the HPV18 E1^E4 protein during differentiation-dependent stages of the virus life cycle within suprabasal keratinocytes.

In addition to HPV18, E1^E4 proteins of other HPV types including 1, 11 and 16 have also been shown to undergo *in vitro* phosphorylation by PKA (Grand et al., 1989; Bryan et al., 2000; Wang et al., 2009), suggesting that this modification is likely to be a conserved mechanism of modulating E1^E4 function. Comparison of the location of PKA putative phosphorylation sites reveals that these consensus sites are restricted to the central region of the E1^E4 proteins (Figure 3.10). PKA has also been shown to modify the function of other papillomavirus proteins including E1 and E6. PKA phosphorylation of E1 may reduce E1 DNA binding activity (Zanardi et al., 1997), whilst phosphorylation of E6 has been shown to
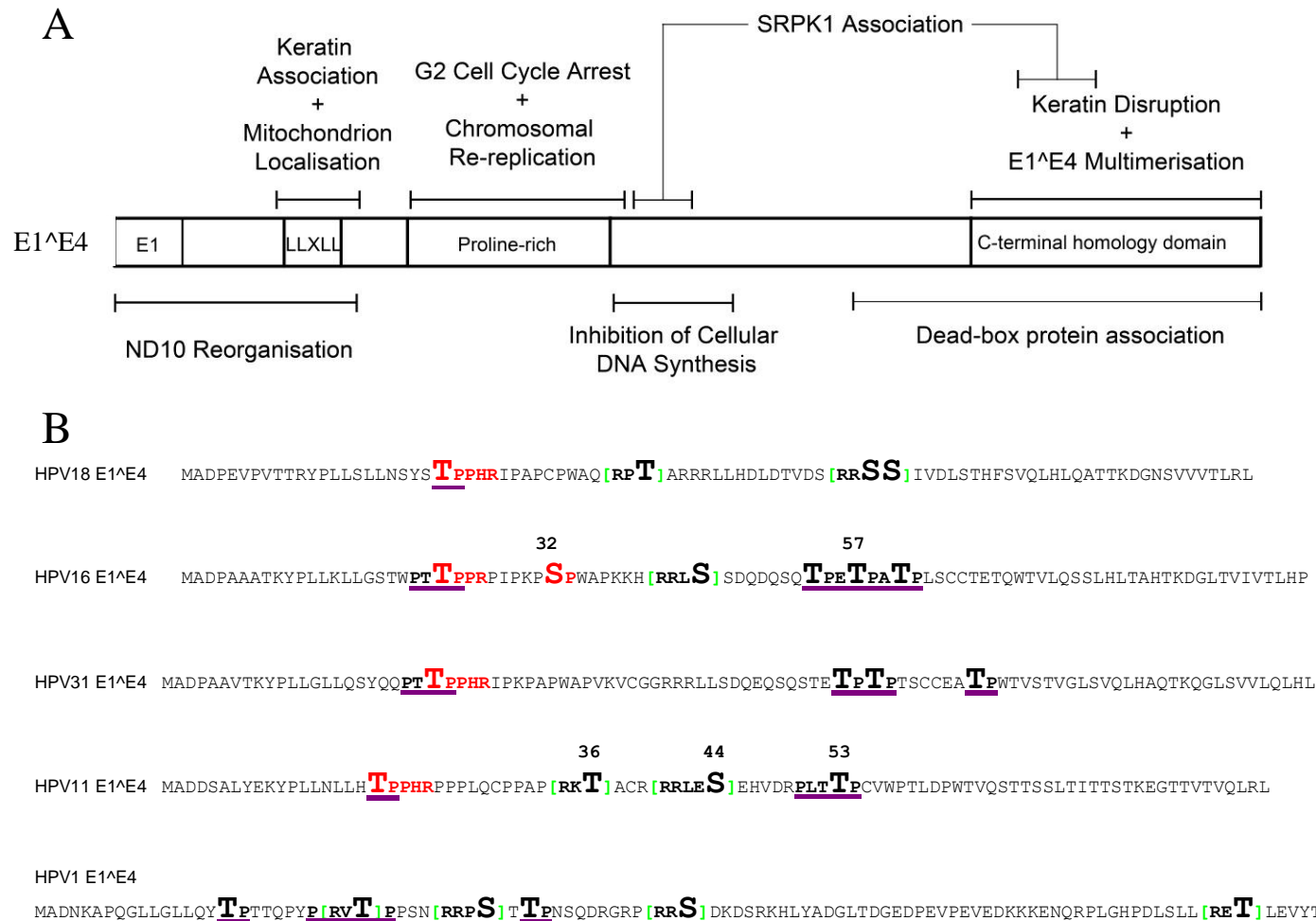
Figure 3.10: Kinase consensus motifs for CDK, MAPK and PKA within HPV E1^E4 proteins. **(A)** A schematic representation of the functional regions of E1^E4 proteins. (B) Amino acid sequences of E1^E4 proteins from both cutaneous and mucosal HPV types. Numbers above residues indicate known phosphorylation sites (Bryan et al. 2000;Davy et al. 2006;Wang et al. 2009). Kinase consensus motifs are indicated as follows: MAPK: PXXS/TP, PXS/TP, S/TP, PKA: [RXS/T] and [RR/KXS/T] and CDK: S/TPX(K/R/H). Putative phosphoacceptor sites are indicated in large typeface with numbers above the sequences indicating the position of known sites of phosphorylation.

regulate the interaction of E6 with PDZ proteins including Dlg (Kuhne et al., 2000). It appears that PKA phosphorylation is a common method employed by the virus to regulate viral protein activity and it is possible that phosphorylation by this kinase may regulate the biological actions of the E1^E4 protein. Use of a phospho-specific antibody for detection of E1^E4 PKA site phosphorylation could enable examination of the role of this phosphorylation event in organotypic raft cultures of HPV18 genome-containing keratinocytes.

3.12 Phosphorylation of E1^E4 proteins by CDK-cyclin complexes

This study has demonstrated that HPV18 E1^E4 undergoes *in vitro* phosphorylation at threonine residue 23 by a number of CDK-cyclin complexes including CDK2-cyclin A, CDK1-cyclin A and CDK2-cyclin E. Whilst *in vitro* phosphorylation of the E1^E4 protein by CDK1-cyclin B was observed, this appeared to be less efficient in comparison with other CDK-cyclin complexes suggesting that in may not be a physiologically relevant phosphorylation. Although mutation of the CDK phospho-acceptor site did not alter the steady state levels or localisation of the HPV18 E1^E4 protein during over-expression studies it is possible that this may vary within differentiating keratinocytes.

The HPV16 E1^E4 protein is also a substrate for *in vitro* phosphorylation by CDK2-cyclin A, with serine residue 32 as the target residue (Davy et al., 2006) and this site has also more recently been shown to undergo phosphorylation by the CDK1-cyclin B complex *in vitro* and within cells (Wang et al., 2009). Putative CDK phosphorylation sites are also observed in a number of HPV E1^E4 proteins (Figure 3.10) indicating that this may be a conserved mechanism for regulation of E1^E4 proteins.

The timing of CDK activity regulates progression through the cell cycle and the *in vitro* phosphorylation of E1^E4 by CDK2-cyclin A and CDK2-cyclin E complexes involved in S phase entry and progression, and by CDK1-cyclin A which regulates the G2/M transition, suggests that the function of the E1^E4 protein may be modulated in a cell cycle-dependent manner. Indeed, E1^E4 proteins have been shown to act at multiple stages of the cell cycle. The ability to induce a G2 arrest is a conserved function of a number of E1^E4 proteins (Nakahara et al., 2002; Davy et al., 2002; Knight et al., 2004) and HPV1 E1^E4 is also able to prevent entry into S phase (Knight et al., 2004). Expression of full-length E1^E4 proteins of HPV types 16, 18 and 31 has been shown to be essential for efficient viral genome amplification in cells which have re-entered S phase (Nakahara et al., 2005; Wilson et al., 2005; Wilson et al., 2007). In addition, HPV E1^E4 proteins of HPV type 16 and 31 also play a role in suppression of cellular DNA synthesis (Nakahara et al., 2005; Wilson et al., 2005) and expression of HPV1 E1^E4 has been shown to prevent initiation of cellular DNA replication by inhibiting licensing of replication origins (Roberts et al., 2008). Considering the role of the E1^E4 protein during different stages of the cell cycle, phosphorylation may act to differentially regulate the activity of the protein, enabling the protein to perform different functions at different phases of HPV infection.

3.13 Phosphorylation of E1^E4 proteins may regulate protein structure

A recent study has demonstrated that phosphorylation of HPV16 E1^E4 by extracellular signal-regulated kinase (ERK)/MAPK has been shown to alter the structure of the protein (Wang et al., 2009). The T57 phospho-acceptor residue of ERK, lies within a negatively charged region of the HPV16 E1^E4 protein (Figure 3.11 A), which in the folded protein forms one side of a flexible loop, in contact with a positively charged region (Figure 3.11 B).

A

27 amino acid loop region peptide

HPV16 ^{23}TPP**R**IP**K**PSPWAP**KKH RR**LSS**D**Q**D**QSQTP**E**TPATPLSCCT**E**TQW^{67}

HPV18 ^{23}TPPH**R**IPAPCPWAQ**RP**TA**RRR**LLH**D**L**D**TV**D**S**RR**SSIV**D**LSTHFSV^{68}

HPV31 ^{24}TPPH**R**IP**K**PAPWAPV**K**VCGG**RRR**LLS**D**Q**E**QSQST**E**TPTPTSCC**E**A^{68}

HPV11 ^{19}TPPH**R**PPPLQCPPAP**RK**TAC**RRR**L**E**S**E**HV**DR**PLTTPCVWPTL**D**PWT^{65}
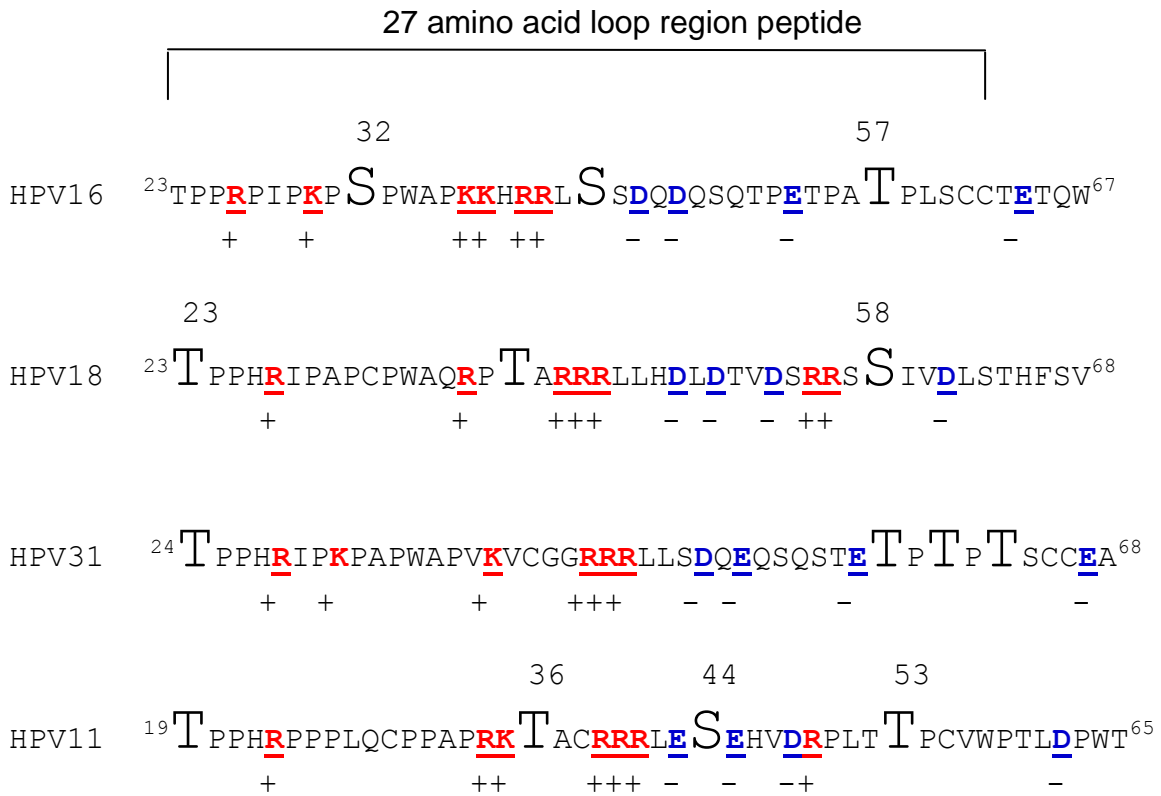
B
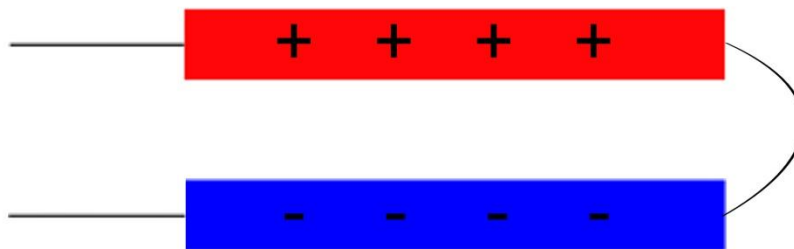
Figure 3.11: Conservation of charged residues within a central region of HPV E1^E4 proteins. **(A)** Primary amino acid sequence within the central region of HPV E1^E4 proteins are shown, with the location of positive (+) and negatively (-) charged residues highlighted in red and blue respectively. The location of a 27 amino acid loop region characterised in HPV16 E1^E4 (Wang et al., 2009) is indicated. Putative phosphoacceptor residues are highlighted in large typeface with numbers above the sequences indicating positions of known phosphorylation. **(B)** Schematic representation of the looped region of HPV16 E1^E4 proteins which may be conserved between E1^E4 proteins. Adapted from Wang et al 2009.

Phosphorylation of T57 increases the negative charge of one side of the loop region which is suggested to enhance folding and leads to a more compact structure of the loop region of the E1^E4 protein. Substitution of T57 with a negatively charged asparagine acid (T57D) to mimic phosphorylation, increased the stability of the HPV16 E1^E4 protein compared with the wild type protein and enhanced keratin binding and reduced proteosomal degradation (Wang et al., 2009). It is hypothesised that the structural change of HPV16 E1^E4 induced by ERK phosphorylation enhances exposure of the leucine-rich motif of the protein which mediates keratin association, and that increased keratin association prevents degradation of the E1^E4 protein (Wang et al., 2009).

Comparison of E1^E4 protein sequences of HPV types 18, 31 and 11 with the loop region of the HPV16 E1^E4 protein sequences indicates that there is a similar distribution of positive and negative charges within this region of E1^E4 proteins suggesting that these proteins may fold in a similar manner (Figure 3.11 A). However, whilst clusters of negatively charged residues are located within the region of the HPV18 E1^E4 protein corresponding to the negatively charged half of the HPV16 E1^E4 loop, immediately adjacent to this region are two positively charged arginine residues which alter the charge distribution within this region and may alter folding of the HPV18 E1^E4 protein (Figure 3.11 A). Although the position of both known and putative phosphorylation sites is not highly conserved between E1^E4 proteins (Figure 3.10) there is similarity in the overall position of phosphorylation sites which are mostly located with a central region of E1^E4 proteins (Figure 3.11 A) corresponding to the defined loop region of the HPV16 protein (Wang et al., 2009).

The sites of MAPK phosphorylation of HPV16 and 11 E1^E4 proteins occur within the sequences: [55]PA<u>T</u>P[58] and [50]PLT<u>T</u>P[54] respectively which match optimal consensus sites for the

kinase (PX<u>S/T</u>P and PXX<u>S/T</u>P) (Figure 3.10). Whilst the E1^E4 protein of HPV type 18 contains a ^{23}TP^{24} di-peptide which matches the minimal requirement of a MAPK consensus site (<u>S</u>/TP), a GST-HPV18 E1^E4 fusion protein was not a substrate for this kinase *in vitro*. This does not, however, exclude the possibility that this protein undergoes phosphorylation within the physiological environment of the host keratinocyte and it is possible that the presence of the GST tag may prevent access of the kinase to the phospho-acceptor residue. Considering that E1^E4 proteins of other anogenital HPV types are targets for MAPK, it was somewhat surprising that HPV18 E1^E4 does not appear to be a substrate for this kinase. However it is possible that introduction of a negative charge in this region of the protein may be compensated for by phosphorylation by another kinase. The PKA phospho-acceptor serine 58 residue of HPV18 E1^E4 lies in a near identical position to the ERK target site of the HPV16 protein (Figure 3.11 A) and it is possible that PKA phosphorylation of HPV18 E1^E4 may act to increase folding in an analogous manner HPV16 E1^E4 phosphorylation.

It is hypothesised that whilst phosphorylation of HPV16 E1^E4 by ERK acts to compact the central region, CDK phosphorylation of the protein at S32 within a positive region of the protein serves to antagonise this effect by reducing folding of the charged regions (Wang et al., 2009). Considering that HPV18 E1^E4 undergoes phosphorylation by CDKs in a similar region of positively charged residues and that there is CDK consensus site conservation between E1^E4 proteins of a number of anogenital HPV types (Figure 3.10) it is possible that CDK phosphorylation may regulate the structure and function of E1^E4 proteins. The report that CDK phosphorylation of HPV16 E1^E4 occurs following ERK phosphorylation within upper layers of the epithelium (Wang et al., 2009), may support a role for CDK phosphorylation of E1^E4 proteins at later stages during the virus life cycle.

Whilst ERK phosphorylation acts to increase the stability of the HPV16 E1^E4 protein through enhanced association with the keratin network (Wang et al., 2009), mutation of the similarly positioned PKA phospho-acceptor site of HPV18 E1^E4 did not alter the steady state levels of the protein. Furthermore, examination of localisation of an HPV18 S58D E1^E4 phospho mimic within an SV40-immortalised cell line did not reveal a higher propensity to co-localise with the keratin IF network in comparison with the wild type protein. However it is conceivable that phosphorylation may act to modulate the structure of the HPV18 E1^E4 protein in a similar manner and that this may serve to regulate the biological activity of the protein, possibly through altered ability to form protein interactions.

The finding that MAPK phosphorylation is not common to all E1^E4 proteins and that modulation of protein stability and cellular localisation of these proteins by phosphorylation differs, indicates that important type-specific differences may exist in the regulation of E1^E4 proteins between the virus types. However considering the finding that HPV18 E1^E4 undergoes phosphorylation by CDK and PKA in a polarised region of the protein in which consensus phosphorylation sites are also located for a number of anogenital E1^E4 proteins, it is proposed that phosphorylation of E1^E4 proteins may be a shared mechanism to regulate the function of the protein.

Since CDK activity in HPV infection impinges upon several HPV functions, the role of HPV18 E1^E4 phosphorylation by CDK-cyclin complexes is examined further within Chapter 4.

CHAPTER 4

HPV18 E1^E4 ASSOCIATION WITH CDK ACTIVITY AND THE EFFECT OF E1^E4 ON CDK-DEPENDENT FUNCTIONS OF HPV PROTEINS

Introduction

A number of HPV proteins can be modified by CDK phosphorylation, which often alters the function of these viral proteins and in addition, several HPV proteins have been shown to modulate kinase activity within cells.

4.1   HPV E7 modifies CDK-cyclin activity

The HPV E7 protein associates with a number of CDK-cyclin complexes and has been shown to either directly, or indirectly modulate their kinase activity. An association has been demonstrated between E7 proteins of both high and low risk HPV types and the CDK2-cyclin A kinase, in complex with Rb1 and p107 or p130 (Tommasino et al., 1993; Ciccolini et al., 1994). Association between HPV18 E7 and CDK2-cyclin E has also been observed, dependent upon p107 (McIntyre et al., 1996), and more recently a direct association has been reported between HPV16 E7 and both cyclin A- and cyclin E-containing CDK2 complexes (Nguyen & Munger, 2008).

The E7 protein can also associate with CDK inhibitors (CKI) and affect their ability to inhibit CDK-cyclin complexes. HPV16 E7 can associate with both p21 and p27$^{\text{cip1}}$ CKIs and prevent CKI-mediated inhibition of CDK2-cyclin E (Zerfass-Thome et al., 1996; Funk et al., 1997). However p21-directed CDK2-cyclin E inhibition was not alleviated by the low risk HPV6 E7, consistent with the observation that low risk E7 proteins do not overcome G1 arrest in response to DNA damage. This observed function of high risk E7 proteins was also reported by another study in which levels of CDK2 activity remained high in HPV16 E7-expressing keratinocytes, despite the induction of p21 and association of the inhibitor with CDK2 (Jones et al., 1997). It is hypothesised that maintenance of CDK2-cyclin E activity by E7 enables a bypass in growth arrest signals thus allowing re-entry into the cell cycle and enabling viral

DNA replication within differentiating keratinocytes. However association between E7 and p21 was only observed at low levels in another study using primary keratinocytes, which contradicted the finding that E7 prevented p21-mediated inhibition of CDK2-cyclin E (Ruesch & Laimins, 1997).

Whilst modulation of CDK activity by high risk HPV E7 proteins has been observed through inhibition of CKI activity, both high and low risk E7 proteins have been shown to directly activate CDK2-cyclin E and CDK2-cyclin A complexes (He et al., 2003). Addition of GST-E7 during an *in vitro* kinase assays resulted in a significant increase in the phosphorylation of both Histone H1 and Rb substrates by these CDK-cyclin complexes.

4.2   Phosphorylation regulates HPV E1 function by modulation of sub-cellular localisation

An interaction between cyclin E and the HPV18 E1 protein has been reported *in vivo* and in complex with E2, HPV11 E1 interacts *in vitro* with cyclins A, B, E and F, dependent upon the presence of an RXL (^{124}RRL^{126}) cyclin binding motif within the E1 protein (Ma et al., 1999). HPV11 E1 undergoes CDK phosphorylation within epithelial cells (Deng et al., 2004) and both proteins of an E1/E2 complex are substrates for *in vitro* phosphorylation by cyclin A and E kinases (Ma et al., 1999). A transient replication assay, in which co-expression of E1 and E2 plasmids results in replication of HPV11 ori-containing plasmids, demonstrated that efficient replication is dependent upon the presence of the E1 RXL motif (Ma et al., 1999). Several S/TP CDK phosphorylation consensus sites have been identified within HPV11 E1 and substitution of these residues also impaired replication, suggesting that association and subsequent phosphorylation of E1 by CDKs is required for efficient viral DNA replication. Further analysis indicated that CDK phosphorylation plays a role in E1 localisation as

inhibition of CDKs, mutation of putative phosphorylation sites or the cyclin-binding (Cy) motif reduced nuclear localisation of E1 (Deng et al., 2004). A classical nuclear export sequence (NES) was identified within the C-terminus of the E1 protein, which is a target site for CRM1-mediated transport of the protein from the nucleus into the cytoplasm. Mutation of the E1 predicted CDK phosphorylation sites resulted in nuclear export whilst further mutation within the NES of these mutants reduced export and significantly restored the replicative activity of these proteins. CDK phosphorylation of E1 therefore serves to inactivate the NES, retaining the E1 protein within the nucleus to enable viral DNA replication.

A bi-partite nuclear localisation sequence (NLS) flanks the NES of HPV11 E1 and has been shown to be important for nuclear import of the protein. In addition, substitution of serine residues 89 and 93 which are targets for ERK and JNK phosphorylation, or inhibition of these MAPK kinases, reduces nuclear import of the E1 protein demonstrating that MAPK phosphorylation is required for efficient E1 import (Yu et al., 2007). HPV11 E1 associates with the MAPK proteins through MAPK docking sequences and reduced MAPK binding abrogates E1 import and replication in transient assays. It is hypothesised that MAPK phosphorylation activates the NLS in order to import the E1 protein into the nucleus to embark upon DNA replication. The sub-cellular localisation and therefore replicative function of the HPV E1 protein is regulated by the actions of MAPK and CDK. Import of the E1 protein into the nucleus is dependent upon MAPK phosphorylation activation of the NLS but for the protein to be retained in this cellular compartment, the NES must be inactivated by CDK phosphorylation. With the E1 phosphorylation status determining the localisation of the protein this ensures that the activity of the helicase can be tightly regulated to prevent unwinding of the viral DNA before the host DNA replicative machinery is present. During S or G2 stages of the cell cycle it is hypothesised that phosphorylation by A and E cyclin

associated CDKs would ensure that the E1 protein remains nuclear. Upon exit of G2 the protein would then undergo de-phosphorylation by protein phosphatase and return to the cytoplasm ensuring that DNA replication occurs at the correct stage of the cell cycle.

4.3 Association of E1^E4 proteins with CDK-cyclin complexes

The E1^E4 protein of HPV type 16 can associate with and sequester CDK1-cyclin B complexes to the cytoplasm of the cell (Davy et al., 2005). Recently, this E1^E4 protein has been shown to be a substrate for phosphorylation by CDK2-cyclin A during *in vitro* kinase assays and is also able to retain this CDK-cyclin complex within the cytoplasm (Davy et al., 2006). It is hypothesised that these interactions result in inhibition of cell cycle progression by preventing nuclear accumulation of these complexes during the G2 stage of the cell cycle, and thus induce a G2 cell cycle arrest. Association with cyclin proteins and induction of a G2 arrest has been mapped to a ^{22}TT^{23} di-peptide within a proline-rich region of the HPV16 E1^E4 protein which forms part of a CDK consensus site, ^{23}TPPR^{26} (S/TPXR/K), although it is a serine residue within a second ^{32}SP^{33} site that is the target of CDK2-cyclin A phosphorylation *in vitro*. However, whilst the ability to induce a G2 arrest function is conserved between E1^E4 proteins, the significance of the E1^E4 G2 arrest function for the virus life cycle is not understood and may just reflect over-expression studies within immortalised epithelial cells.

In HPV-infected cells, interaction between E1^E4 and CDK-cyclin complexes and phosphorylation of the viral protein may impinge upon the function of other HPV proteins that interact with CDKs, including E1 and E7. The sequences required for interaction of

HPV18 E1^E4 with CDKs were identified and are discussed within this chapter, to enable the function of this association to be investigated.

Results

4.4 Phosphorylation of HPV18 E1^E4 by CDK-cyclin complexes is dependent upon RXL cyclin binding motifs

Interactions between CDK-cyclin complexes and their substrates are mediated by a bi-partite recognition mechanism. The consensus motif, S/TPXR/K, is recognised by the CDK catalytic domain, whilst the cyclin binding (Cy) motif with the sequence, RXL, is located at a site within the substrate that is remote from the phospho-acceptor sequence, usually located downstream, and is recognised by the interacting cyclin partner. The presence of both elements of this bi-partite substrate recognition motif are necessary for efficient phosphorylation by the CDK-cyclin complex. Examination of the protein sequence of HPV18 E1^E4 reveals the presence of two overlapping RXL motifs at [43]RRLL[46] which are located downstream from the phosphorylation consensus site, [23]TPPH[26]. The first motif composed of [43]RRL[45], corresponds to an optimal Cy motif, and overlaps with a second motif at, [44]RLL[46] (Figure 4.1 A). Interestingly, the presence of these bi-partite motifs is a conserved feature of E1^E4 proteins of alpha HPV types (Figure 4.1 A) with these motifs present within E1^E4 proteins of HPV61, 72, 81, 83 (species 3); HPV18, 39, 45, 59, 68 (species 7); HPV43 (species 8); HPV16, 31 (species 9); HPV6, 11, 13, 54 (species 10); and HPV54 (species 13). In addition to HPV18, the presence of two Cy motifs is also observed in HPV31, 43, 45 and 59 whilst the other HPV types contain only a single motif.

A

```
        S/TPX(K/R)-------------RXL

HPV18   STPPHRIPAPCPWAPQRPTA RRRLLH
HPV16   TTPPHPIPKPSPWAP K     KHRRLSS
HPV31   TTPPHRIPKPAPWAPVKVCGGRRRLLS
HPV45   NTPPRRPPKPHPWAPQNPTS RRRLLS
HPV54   TTPQKQSRA            RRRLEN
HPV11   TTPPRPPPLQCPPAPRKTAC RRRLGS
```

B

```
E1^E4   ²²STPPH²⁶------⁴³RRLL⁴⁶
T23I    ²²SIPPH²⁶
RXL1                  ⁴³HRRL⁴⁶
RXL2                  ⁴³RPLR⁴⁶
RXLD                  ⁴³HPRR⁴⁶
```

C

CDK2-cyclin E

D

CDK2-cyclin A

E

CDK1-cyclin A
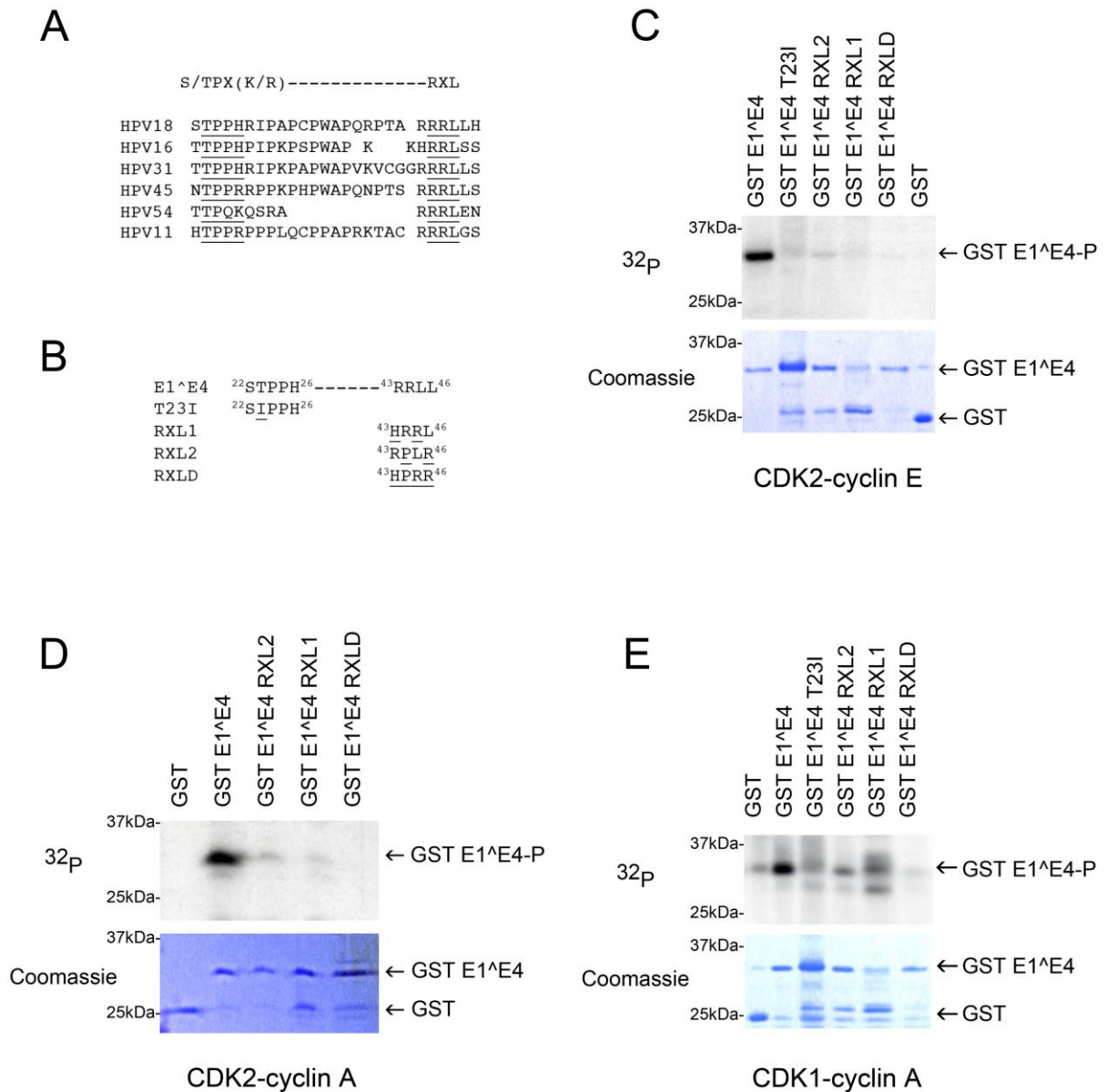
Figure 4.1: Phosphorylation of HPV18 E1^E4 by CDK-cyclin complexes is dependent upon the presence of RXL motifs. **(A)** The presence of a bi-partite CDK recognition site is conservation between E1^E4 proteins of the alpha genus. The putative CDK phosphorylation sites (S/TPX[R/K]) and the RXL cyclin binding (Cy) components of the bi-partite motif are underlined. **(B)** The amino acid substitutions made within the bi-partite CDK recognition motif of HPV18 E1^E4 with substituted residues underlined. **(C-E)** *In vitro* kinase phosphorylation of GST-HPV18 E1^E4, T23I and RXL mutant proteins by: **(C)** CDK2-cyclin E, **(D)** CDK2-cyclin A, **(E)** CDK1-cyclin A. GST-HPV18 E1^E4 is efficiently phosphorylated by CDK-cyclin complexes whilst RXLD mutant fusion proteins are not substrates for CDK-cyclin phosphorylation. The extent of phosphorylation of the RXL1 and RXL2 mutant fusion proteins varies. GST protein levels were visualised by Coomassie stained gels and each experiment is representative of at least three independently performed kinase assays.

To establish whether either, or both of the overlapping Cy motifs are required for phosphorylation of HPV18 E1^E4 by CDK-cyclin complexes, mutations of the RXL motif were introduced which are silent within the E2 ORF (Figure 4.1 B and Appendix I), and the GST-fusion proteins assayed for their ability to co-precipitate with CDK activity. The first of the two RXL motifs, [43]RRL[45], was mutated to HRR and has henceforth been assigned the nomenclature RXL1, whilst the second motif, [44]RLL[46], was altered to PLR and will be referred to as RXL2. However, due to the overlapping nature of these to Cy motifs it is impossible to mutate one motif without altering the structure of the second binding site. A third construct was also generated in which both RXL motifs were mutated by all of the afore-mentioned substitutions, to generate a double RXL mutant (RXLD) in which [43]RRLL[46] was mutated to HPRR. The CDK phospho-acceptor, threonine 23 mutant (T23I) has been previously discussed within Chapter 3. *In vitro* kinase assays were then carried out with various GST-HPV18 E1^E4 mutant proteins and CDK-cyclin complexes to establish whether the RXL motifs are required for E1^E4 phosphorylation.

Phosphorylation of E1^E4 GST-fusion proteins by CDK2-cyclin E *in vitro* was severely reduced upon mutation of either RXL motif (RXL1 and RXL2), or both motifs (RXLD) in comparison with that of the wild type GST-HPV18 E1^E4, and no phosphorylation of the GST control protein was observed (Figure 4.1 C). The extent of phosphorylation of each of the various RXL mutant proteins by CDK2-cyclin A and CDK1-cyclin A kinases was similar, with severely impaired phosphorylation of the RXL single mutations, and virtually undetectable phosphorylation of the RXL double mutation, as would be predicted due to the two kinases comprising the same cyclin subunit (Figure 4.1 D-E). Phosphorylation of RXL1 and RXL2 mutant GST-fusions proteins by both kinases was severely reduced and was completely abrogated in the RXLD reaction. In conclusion, these findings indicate that an

intact [43]RRLL[46] sequence is required for *in vitro* phosphorylation of HPV18 E1^E4 by CDK1 and CDK2 in complex with cyclin A or by CDK2-cyclin E.

4.5   The cyclin binding motifs of HPV18 E1^E4 mediate an association with CDK activity *in vitro*

The observation that HPV18 E1^E4 phosphorylation at threonine residue 23 by CDK-cyclin complexes *in vitro* is dependent upon intact RXL motifs, led to examination of whether E1^E4 can associate with these cellular complexes via the RXL motif. To establish whether E1^E4 is able to associate with active CDKs, a co-precipitation from HeLa cell lysate was carried out using immobilised GST proteins which were then subjected to an *in vitro* kinase assay with exogenous Histone H1, a known CDK target, as a substrate (Figure 4.2 A). Whilst addition of the GST co-precipitate to an *in vitro* kinase assay did not result in phosphorylation of Histone H1, a co-precipitate generated by GST-HPV18 E1^E4 resulted in efficient phosphorylation of the substrate, demonstrating association of HPV18 E1^E4 with a Histone H1-specific kinase activity. To validate that this Histone H1-specific activity is CDK-dependent, the kinase assay was carried out in the presence of the CDK1- and CDK2-specific inhibitor, Roscovitine. Histone H1 phosphorylation was significantly reduced by the addition of Roscovitine to the kinase assay, confirming that the co-precipitated kinase corresponds to a CDK and demonstrating that HPV18 E1^E4 can associate with CDK1 and/or CDK2 *in vitro*.

To establish whether threonine 23 within the CDK consensus site of HPV18 E1^E4 is required for association with CDK activity, an *in vitro* kinase assay was carried out using a co-precipitate obtained with a T23I mutant GST-fusion protein (Figure 4.2 B). Histone H1 was phosphorylated to a similar extent during *in vitro* kinase assays with co-precipitations obtained from both the mutant T23I and wild type GST proteins, demonstrating that the CDK
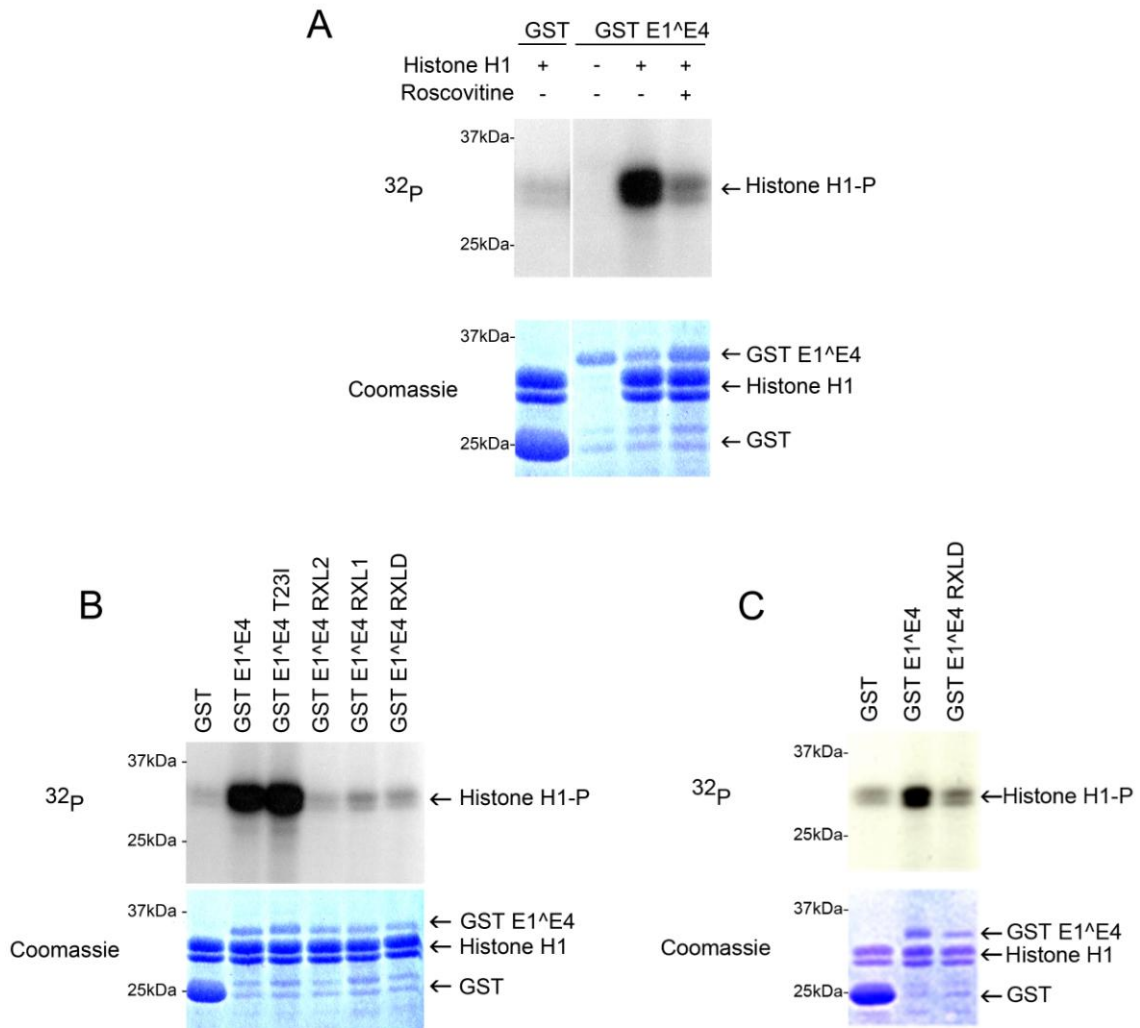
Figure 4.2: Association of HPV18 E1^E4 with CDK-associated kinase activity is dependent upon RXL motifs. (A) HPV18 E1^E4 co-precipitates with a CDK-specific kinase activity. Phosphorylation of Histone H1 with co-precipitates from HeLa cell lysates obtained with GST and GST-HPV18 E1^E4 proteins in the presence and absence of the CDK-specific inhibitor, Roscovitine. Loading of GST and Histone H1 proteins was determined by the Coomassie stained gel. (B) *In vitro* phosphorylation of Histone H1 with co-precipitates from HeLa cell lysates obtained with GST and HPV18 E1^E4 wildtype (WT), T23I and RXL mutant GST-fusion proteins. Visualisation of protein levels by Coomassie staining indicates protein loading. (C) Histone H1 phosphorylation by GST, GST-HPV18 E1^E4 WT and RXLD co-precipitates derived from SAOS-2 cells. Each experiment is representative of three independently performed kinase assays.

phosphorylation site of the bi-partite consensus motif is not required for association of HPV18 E1^E4 with CDK activity. However, co-precipitations obtained by the RXL1, RXL2 and RXLD GST-fusion proteins resulted in a significantly reduced level of Histone H1 phosphorylation, indicating that the sequence: [43]RRLL[46] must be intact for association of HPV18 E1^E4 with CDK activity *in vitro* (Figure 4.2 B).

HeLa cells are derived from a cervical adenocarcinoma and contain chromosomally integrated HPV18 genomes. To establish whether the association of E1^E4 with CDK activity can occur independently of the E6 and E7 HPV proteins expressed within HeLa cells, a co-precipitation was carried out in SAOS-2 cells, which are derived from an osteosarcoma and are HPV-negative (Figure 4.2 C). Whilst GST alone did not co-precipitate with CDK activity, GST-HPV18 E1^E4 associated with a Histone H1 specific kinase in SAOS-2 cells, indicating that this association is independent of other HPV proteins. Consistent with the findings from the co-precipitations carried out in HeLa cells, phosphorylation of Histone H1 was also reduced by the RXLD mutation confirming the observation that an intact RXL motif is required for association of HPV18 E1^E4 with CDK activity.

4.6 Phosphorylation of threonine residue 23 of HPV18 E1^E4 is not required for sequestration of cyclin A to the cytoplasm

Immunofluorescence studies of cells transfected with codon-optimised HPV18 E1^E4 have shown that both cyclin A and B, but not cyclin E, are sequestered to the cytoplasm. To establish whether sequences within the bi-partite recognition motif of HPV18 E1^E4 are required for cyclin sequestration, mutations across this region were generated within the HPV18 coE1^E4 expression construct and following transfection into COS-1 cells, co-localisation with cyclin A was examined by immunofluorescence microscopy (Figure 4.3).
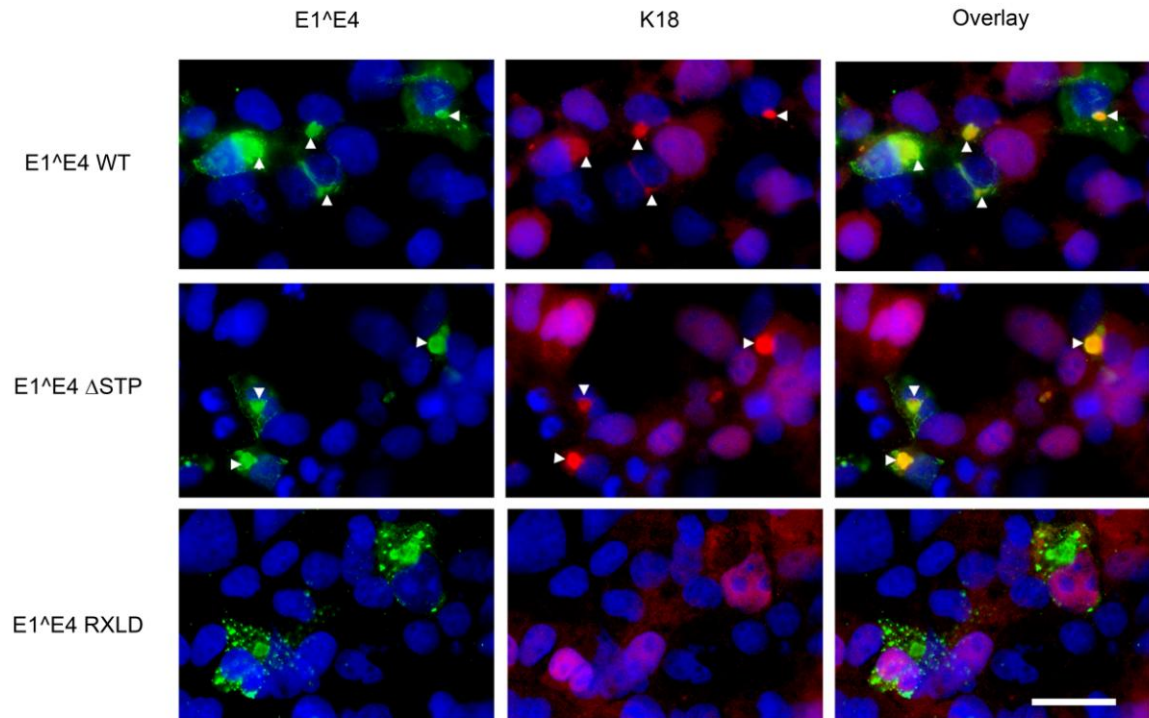
Figure 4.3: The HPV18 E1^E4 target residue for CDK *in vitro* phosphorylation is not required for cytoplasmic sequestration of cyclin A. Immunofluorescence microscopy of COS-1 cells transfected with wildtype E1^E4, an Δ[22]STP[24] E1^E4 deletion mutant and an RXLD E1^E4 mutant. Cells were dual stained for E1^E4 (green) and cyclin A (red) and were counterstained with DAPI for detection of nuclei (blue). Cyclin A is sequested to E1^E4 cytoplasmic inclusions witin cells expressing WT E1^E4 or the STP E1^E4 deletion mutant (cells indicated with arrowheads) but not within cells expressing the RXLD E1^E4 mutant protein. Bar, 30 μm. This result was repeated in two independently performed experiments.

Whilst the wild type E1^E4 protein sequestered cyclin A to a collapsed E1^E4 structure within the cytoplasm, the absence of ^{22}STP^{24} residues containing the CDK phosphorylation site did not affect the ability of the E1^E4 protein to sequester this cyclin protein. The extent of cyclin A association of the Δ^{22}STP^{24} deletion mutant was comparable with that of the wild type protein indicating that the threonine 23 CDK phospho-acceptor residue is not required to mediate this association. This finding is supported by the *in vitro* pull-down assays in which the GST E1^E4 T23I mutant protein retained the ability to associate with active CDK to the same extent as the wild type GST-E1^E4 fusion protein (Figure 4.2 B). However, mutation of the overlapping RXL cyclin binding motifs abrogated association with active kinases (Figure 4.2B), and co-localisation of the E1^E4 with cyclin A was also disrupted by introduction of the RXLD mutation (Figure 4.3).

The immunofluorescence microscopy confirming the sequestration of cyclin A by wild type E1^E4, with abrogation observed upon introduction of mutations to the RXL motif was performed by Dr G.L. Knight whilst the phenotype of the Δ^{22}STP^{24} deletion mutant was characterised by myself.

4.7 The effect of HPV18 E1^E4 CDK association on CDK-dependent functions of HPV E7 and E1 proteins

Given that HPV18 E1^E4 can alter the sub-cellular distribution of cyclin A proteins it is possible that through this interaction, E1^E4 may act to modulate the function of HPV E1 and E7 proteins which are either regulated by, or modulate the activity of, CDKs respectively.

4.7.1 The sub-cellular distribution of E1 is not altered by the presence of the HPV18 E1^E4 protein

The localisation of the HPV E1 protein is regulated by CDK-cyclin complexes, whereby phosphorylation of a NES within E1 acts to retain this protein within the nucleus (Deng et al., 2004). The presence of CDK phosphorylation sites is conserved across HPV E1 proteins and due to the availability of HPV1 flag-tagged E1 expression construct (a kind gift of Professor S. Khan), the localisation of this protein was examined in the presence of HPV18 E1^E4. To establish whether E1^E4 sequestration of cyclin proteins to the cytoplasm may act to prevent CDK-directed nuclear localisation of the E1 protein, HPV1 flag-tagged E1 and HPV18 coE1^E4 were co-transfected into COS-1 cells and their sub-cellular localisation examined by immunofluorescence with an anti-flag rabbit antibody and a mouse monoclonal E1^E4 antibody (Figure 4.4). HPV1 E1 was mostly detected within the nucleus with some cytoplasmic localisation (Figure 4.4 a), consistent with previous studies (Deng et al., 2004), whilst E1^E4 displays a predominantly cytoplasmic filamentous distribution (Figure 4.4 b). Within co-transfected cells, in which the E1^E4 protein was either filamentous (Figure 4.4 c) or collapsed into juxta-nuclear bundles (Figure 4.4 d), the distribution of the E1 protein between the nucleus and cytoplasm was comparable to that of E1 cellular localisation in the absence of E1^E4, with a largely nuclear distribution, indicating that within these cells E1^E4 sequestration of cyclins does not modulate the localisation of the HPV1 E1 protein.
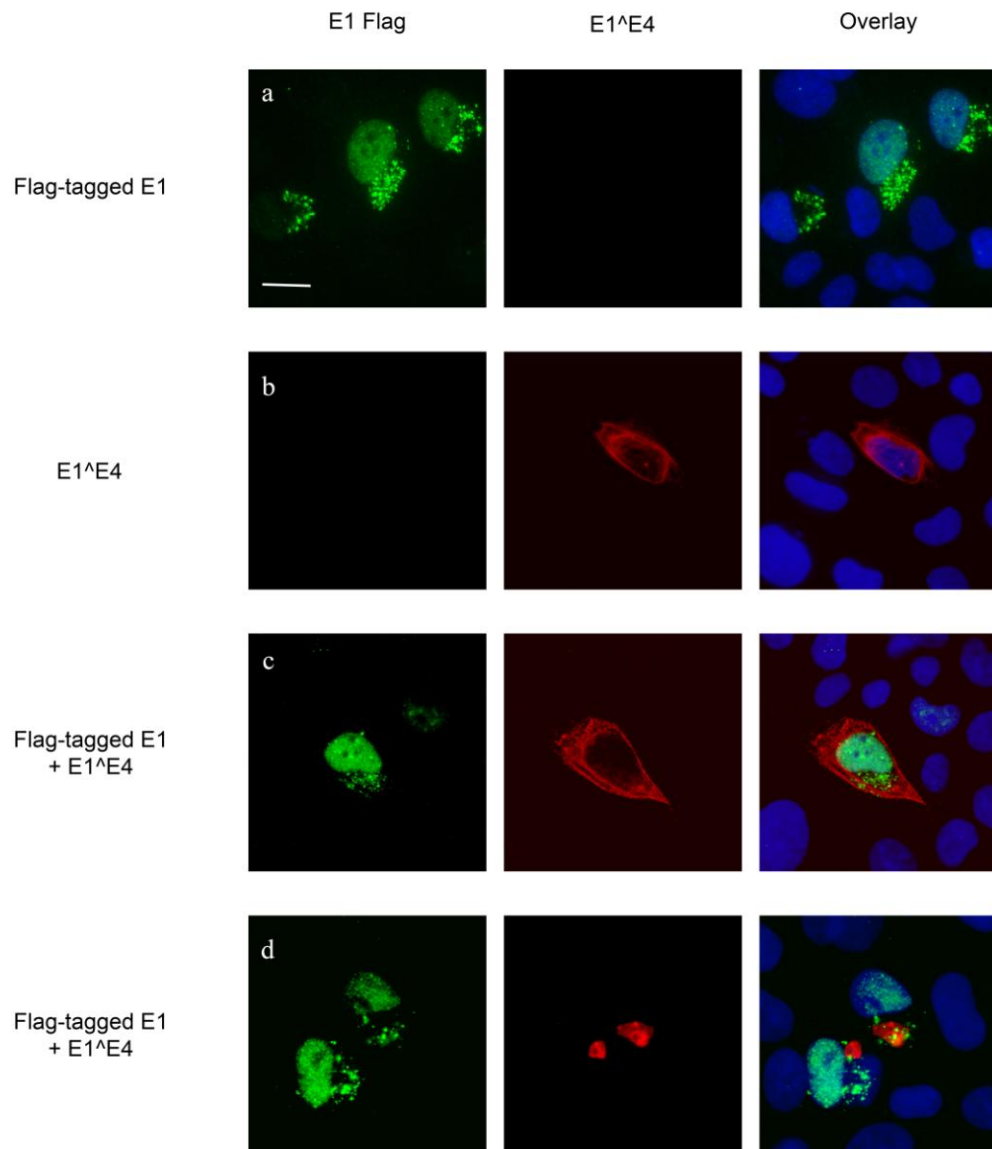
Figure 4.4: The nucleo-cytoplasmic localisation of the E1 protein is not altered by the presence of E1^E4. Immunofluorescence microscopy of COS-1 cells transfected with **(A)** Flag-tagged HPV1 E1 **(B)** codon-optimised HPV18 E1^E4 **(C-D)** both Flag-tagged HPV1 E1 and codon-optimised HPV18 E1^E4. Cells were dual-stained with antibodies to E1^E4 (red) and Flag-tag (green) and counter-stained with DAPI for detection of nuclei (blue). Bar, 10 μM. Results are representative of two independently performed experiments.

4.7.2   HPV18 E1^E4 inhibits CDK2-cyclin A in the presence of the HPV18 E7 protein
in vitro

4.7.2.1   Modulation of CDK2-cyclin A activity by HPV18 E1^E4

Association of HPV18 E1^E4 with CDK-cyclin complexes may act to modulate the activity

of these kinases. Analysis focussed on the effect of E1^E4 on the activity of CDK2-cyclin A

due to the ability of HPV18 E1^E4 to sequester the majority of cyclin A to the cytoplasm,

whilst sequestration of cyclin B complexes by E1^E4 is reduced (G.L. Knight and S. Roberts,

unpublished data). To examine whether the activity of purified CDK2-cyclin A is altered by

the presence of the E1^E4 protein, an *in vitro* kinase assay was performed using Histone H1

as a substrate. Prior to the kinase assay, CDK2-cyclin A was incubated on ice for 30 min with

GST or a GST-E1^E4 fusion protein and the level of incorporation of radioactive isotope was

analysed by phosphor imaging as a measure of CDK activity (Figure 4.5 A). Upon addition

of increasing amounts of GST-E1^E4 fusion protein the phosphorylation of Histone H1 by

CDK2-cyclin A decreased whilst the activity of the kinase was unaffected by the presence of

GST protein indicating that increasing concentrations of HPV18 E1^E4 inhibits the activity of

CDK2-cyclin A activity *in vitro*. This was repeated three times and 1 µg of GST-E1^E4

fusion protein inhibited CDK2-cyclin A activity by an average of 40% relative to CDK

activity in the presence of the GST control (Figure 4.5 B).

4.7.2.2   Modulation of CDK activity by HPV18 E1^E4 in the presence of HPV18
**   E7**

It has been reported that the S phase specific CDK2-cyclin A complex is activated in the

presence of HPV 6, 16 and 31 E7 proteins, with phosphorylation of Histone H1 by the kinase

reported to increase by between 12 and 18 fold when kinase assays were performed in the
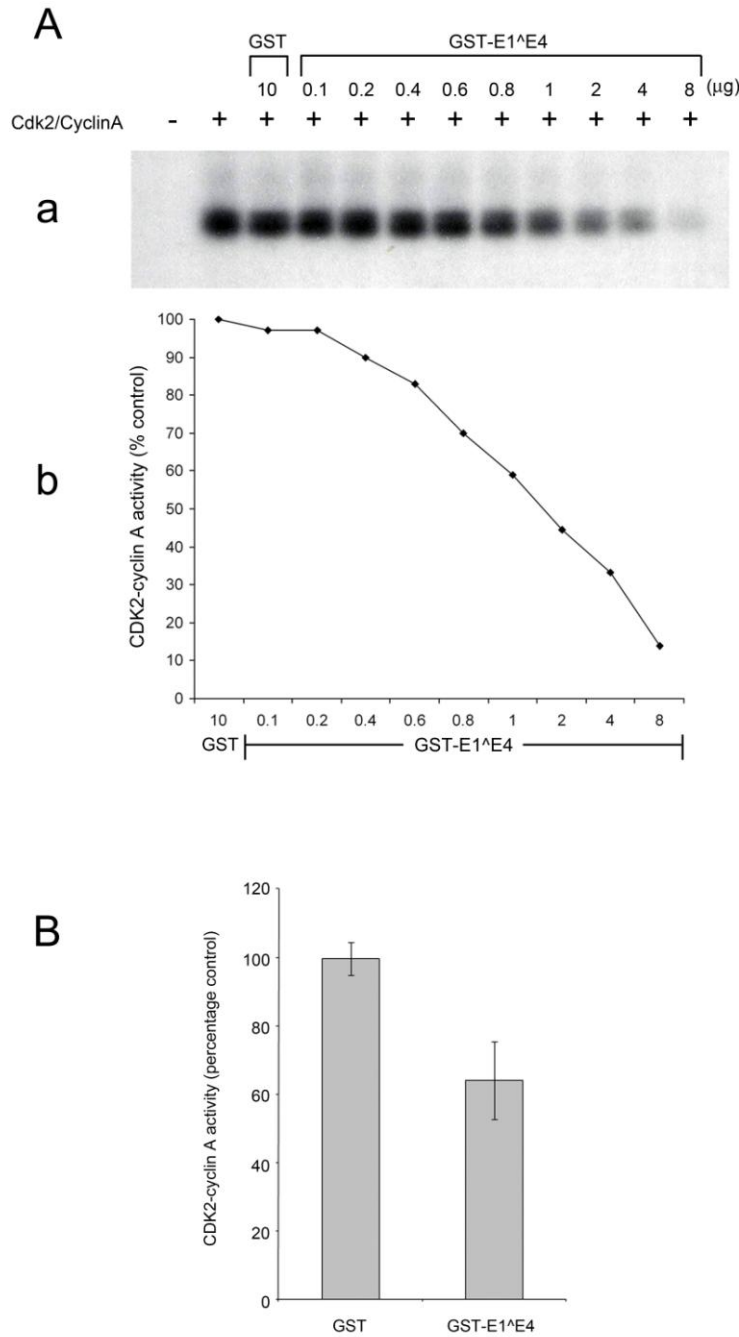
Figure 4.5: HPV18 E1^E4 inhibits purified CDK2-cyclin A activity in vitro. **(A)** (a) Autoradiograph demonstrating *in vitro* phosphorylation of Histone H1 by purified CDK2-cyclin A following prior incubation on ice for 30 min with bacterially-expressed GST or GST-HPV18 E1^E4 proteins. Whilst pre-incubation with GST did not alter the extent of Histone H1 phosphorylation by CDK2-cyclin A, incubation of the kinase with increasing amounts of GST- E1^E4 resulted in progressive reduction in the levels of Histone H1 phosphorylation. (b) Line chart showing data derived from phosphorimaging data in (a) with data normalised to that of the kinase reaction performed in the absence of GST protein. **(B)** Bar chart showing CDK2-cyclin A activity in the presence of GST or GST-E1^E4, calculated as a percentage of Histone H1 phosphorylation by the kinase in the absence of GST protein. Data was derived from phosphorimaging data of three independent experiments using three matched sets of GST and GST-E1^E4 fusion protein and was normalised to that of the kinase reaction performed in the absence of GST protein. Inhibition of CDK2-cyclin A activity by GST-E1^E4 is significantly reduced (One-sided Student T-test: GST alone [p=0.43], GST-E1^E4 [p=0.02]).

presence of the GST-E7 fusion proteins (He et al., 2003). Whilst the E7 protein is predominantly localised to the nucleus, it is in possession of a NES and has been shown to be capable of shuttling between the nucleus and the cytoplasm (Knapp et al., 2009). It is therefore possible that it may enter the vicinity of the E1^E4 protein which is largely cytoplasmic. The inhibition of CDK2-cyclin A activity by HPV18 E1^E4 may act to reduce or prevent the activation of the kinase by HPV E7. To enable expression of HPV18 E7 within bacteria, the E7 ORF was cloned from the pGEMII-HPV18 vector into the pGEX-3X vector to permit expression of a GST-E7 fusion protein within bacteria (Chapter 2: Materials and Methods, 2.2.8.1).

In vitro kinase assays were performed with purified CDK2-cyclin A and Histone H1 as a substrate, with prior incubation of the kinase with: GST, GST E1^E4 or GST-E7 fusion proteins. Pre-incubation of CDK2-cyclin A with the GST-E7 fusion protein or with the GST control protein did not alter the ability of the kinase to phosphorylate the Histone H1 substrate protein *in vitro*, indicating that during this assay HPV18 E7 does not activate CDK2-cyclin A (Figure 4.6 A-B). However incubation of CDK2-cyclin A with the GST-E1^E4 fusion protein resulted in a reduction in Histone H1 phosphorylation as observed previously (Figure 4.5). Upon co-incubation of the GST-E1^E4 fusion protein with GST-E7 and CDK2-cyclin A, the level of Histone H1 phosphorylation reduced to levels comparable with GST-E1^E4 alone (Figure 4.6 A-B). GST and GST-E7 proteins combined did not elicit this affect, indicating that GST-E1^E4 is able modulate the activity of the kinase in the presence of the HPV18 E7 protein.

As a further control, a cellular protein was also used, the SRp20 splicing factor protein, which was kindly made available to me as a GST-SRp20 fusion protein by Ms Emma Prescott.
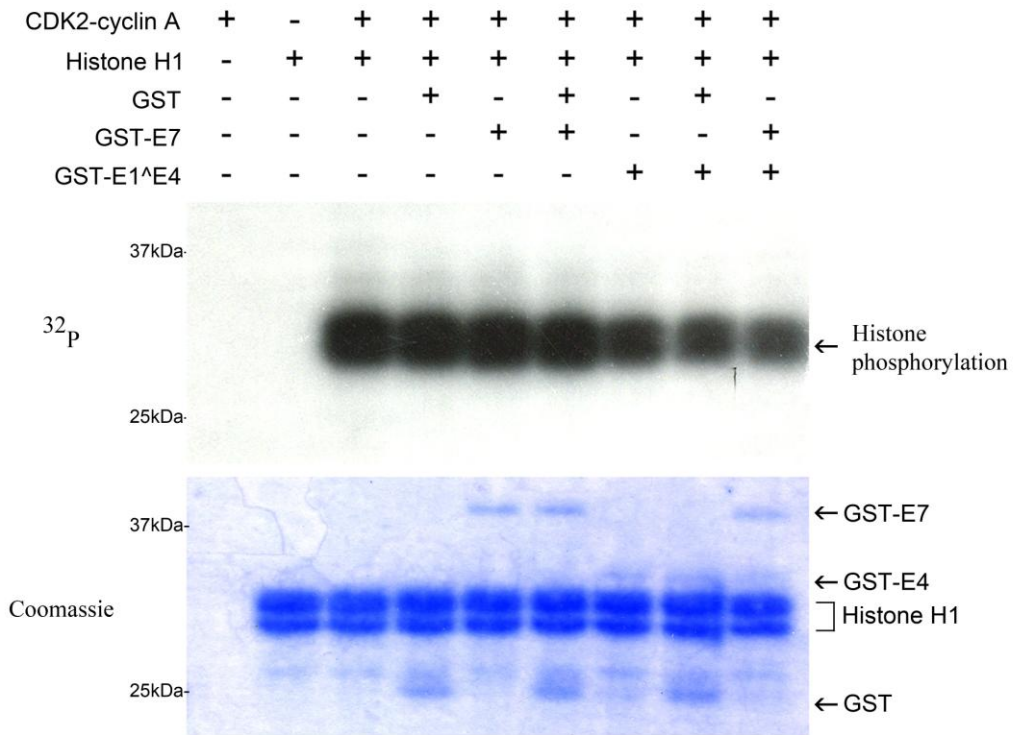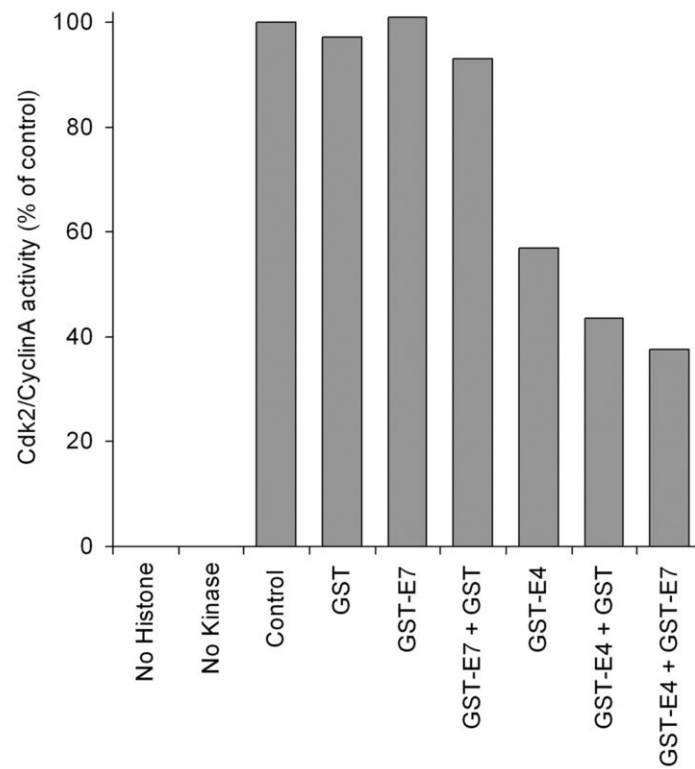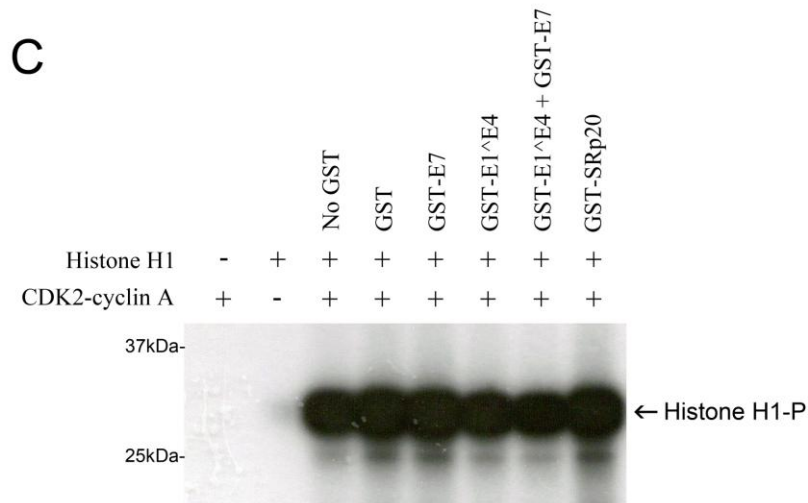
Figure 4.6: Inhibition of CDK2-cyclin A by HPV18 E1^E4 occurs *in vitro* in the presence of HPV18 E7. (A) *In vitro* phosphorylation of Histone H1 by purified CDK2-cyclin A following prior incubation with bacterially-expressed GST, GST-HPV18 E1^E4 or GST-HPV18 E7 fusion proteins. The Coomassie stained gel shows equal loading of Histone H1 and GST-fusion proteins. (B) Line chart showing data derived from phosphorimaging data in (A) with data normalised to that of the kinase reaction performed in the absence of GST protein. (C) *In vitro* CDK2-cyclin A phosphorylation of Histone H1 following prior incubation with GST, GST-HPV18 E1^E4, GST-HPV18 E7 or GST-SRp20 fusion proteins. Experiments are representative of two independently performed experiments.

Whilst GST-E1^E4 inhibited CDK2-cyclin A activity, upon incubation of the kinase with GST-SRp20, Histone H1 phosphorylation was comparable to the GST negative control, indicating that inhibition is specific to the GST-E1^E4 protein (Figure 4.6 C).

The ability of HPV18 E1^E4 to inhibit the activity of CDK2-cyclin A led to the hypothesis that the E1^E4 protein may act to modulate the activity of other CDK-cyclin complexes. The activity of a number of kinases including CDK2-cyclin E, CDK1-cyclin A and CDK1-cyclin B, was assayed in the presence of GST-E1^E4, GST-E7 and control proteins for their ability to phosphorylate Histone H1. However whilst the inhibition of CDK2-cyclin A by GST-E1^E4 but not by the GST control was a consistently reproducible result (Figure 4.5 B), variation in results were observed with other CDK-cyclin complexes (data not shown).

Since viral genome amplification occurs in cells that have re-entered S phase, it is possible that phosphorylation of the E1^E4 protein by S phase CDKs may act to coordinate the activity of E1^E4 in viral replication with the presence of the cellular replication machinery. The HPV E1 protein is also a target for phosphorylation by CDK-cyclin complexes and inability of the E1 protein to undergo phosphorylation by this kinase results in impaired replication of viral genomes, indicating a role for CDK targeting of this protein for viral genome amplification (Deng et al., 2004). Considering the role of both E1 and E1^E4 proteins in amplification of viral genomes, phosphorylation of these viral proteins by CDKs may be a mechanism by which viral genome amplification could be coordinated. However threonine residue 23 of the HPV18 E1^E4 protein has been shown to be dispensable for this viral function, demonstrated by the ability of HPV18 genomes mutated at this residue, to undergo differentiation-dependent viral genome amplification (G.L. Knight and S. Roberts unpublished data).

Discussion

During this study, *in vitro* phosphorylation of HPV18 E1^E4 by CDK2-cyclin E, CDK2-cyclin A and CDK1-cyclin A complexes has been shown to be dependent upon the integrity of both elements of a bi-partite CDK recognition motif, requiring threonine residue 23 as a CDK phospho-acceptor residue and the two overlapping Cy motifs within the sequence: [43]RRLL[46]. HPV18 E1^E4 has also been shown to associate *in vitro* with active CDK-cyclin complexes, and residues important for this interaction have been mapped within the E1^E4 sequence. Whilst threonine residue 23 is dispensable for the interaction between HPV18 E1^E4 and CDK activity, the integrity of the cyclin binding motifs has been shown to be critical. HPV18 E1^E4 can sequester cyclins A to the cytoplasm and association is dependent upon the Cy motifs. This study has demonstrated that cyclin association does not require the presence of CDK phospho-acceptor residue.

4.8   The cyclin binding motif constitutes an important functional domain of the HPV18 E1^E4 protein

Given the conserved nature of RXL Cy motifs between alpha HPV types (Figure 4.1 A) it is likely that these residues mediate important functional roles of E1^E4 proteins. The [43]RRLL[46] sequence of HPV18 E1^E4 enables association with active CDKs, mediates sequestration of CDK-cyclin complexes to the cytoplasm of the cell and is required for phosphorylation of the E1^E4 protein by these CDK-cyclin complexes. The presence of two RXL motifs is not unique to the HPV18 E1^E4 protein, and whilst HPV43 and 59 E1^E4 proteins contain two distinct Cy motifs, E1^E4 proteins of HPV45 and 31 also contain an RRLL sequence comprising two overlapping motifs. The relevance of each RXL motif of HPV18 E1^E4 to the ability of the protein to associate with and undergo phosphorylation by

CDK-cyclin complexes could not be analysed separately due to the overlapping nature of these motifs, whereby one Cy motif could not be altered by mutagenesis without affecting the other. It is therefore possible that both Cy motifs may be functionally active and mediate interactions with CDK activity or alternatively it is also feasible that one of these motifs may be functional redundant.

4.9 Association of HPV18 E1^E4 with active CDK may serve to modulate the activity of cellular or viral proteins regulated by CDK phosphorylation

The possible functional relevance of HPV18 E1^E4 CDK phosphorylation has been discussed previously within Chapter 3 and E1^E4 phosphorylation may serve to regulate the role of the protein within the cell by altering the stability, localisation, or ability of the protein to form interactions. The finding that the E1^E4 protein is able to associate with CDK activity and relocate CDK-cyclin complexes to the cytoplasm suggests that this interaction may be functionally significant. It is possible that in addition to enabling phosphorylation of threonine residue 23 of E1^E4, sequestration of CDK-cyclin complexes to the cytoplasm may also act to enable E1^E4 to direct CDK activity to other proteins to modulate their activity. The number of known interacting partners of the E1^E4 protein are fairly small at present and it is possible that interaction with CDKs may serve to alter an as yet undefined interacting protein.

4.9.1 HPV E1^E4 cyclin sequestration and the localisation of the HPV E1 protein

The ability of E1^E4 to associate with and alter the subcellular distribution of CDK-cyclin complexes may also act to subvert CDK activity from target HPV proteins within the nucleus. CDK phosphorylation within the LRR of HPV11 E1 inactivates the NES and the E1 protein is

subsequently retained within the nucleus (Deng et al., 2004). Comparison of the LRR sequences of E1 proteins of HPV1 and 18 with the NES and known sites of CDK phosphorylation within the LRR of the HPV11 E1 protein (Deng et al., 2004), revealed considerable homology between the E1 proteins (Figure 4.7), suggesting that the nuclear localisation of HPV E1 proteins may be regulated by a common mechanism. Considering the role of CDK phosphorylation in the localisation of the E1 protein, it was hypothesised that cytoplasmic sequestration of CDK-cyclin complexes by E1^E4 may act to prevent E1 CDK phosphorylation and prevent nuclear accumulation of this protein. Within this study, the nucleo-cytoplasmic distribution of the HPV1 E1 protein was comparable to that of the HPV11 E1 protein (Deng et al., 2004) and the localisation of HPV1 E1 was unaltered by co-expression with the HPV18 E1^E4 protein (Figure 4.4). It is possible however that examination of the localisation of the HPV18 E1 protein in the presence of HPV18 E1^E4 may reveal an altered E1 distribution.

Inability of the E1 protein to undergo CDK phosphorylation, and remain in the nucleus has adverse effects on its ability to replicate viral genomes (Ma et al., 1999). Therefore, given the role of E1^E4 proteins during viral genome amplification (Nakahara et al., 2002; Wilson et al., 2005; Wilson et al., 2007), the observation that HPV1 E1 is retained in the nucleus, despite co-expression of HPV18 E1^E4 appears to be consistent with the function of both the E1 and E1^E4 proteins in viral genome amplification. It is possible that the observed nuclear localisation of HPV1 E1 in the presence of HPV18 E1^E4 is due to CDK phosphorylation of the E1 protein within these cells prior to the sequestration of cyclin complexes by E1^E4 to the cytoplasm, and therefore the E1 protein is able to remain nuclear. Indeed, during the virus life cycle, the E1 protein is expressed prior to the onset of high-level E1^E4 expression and it is therefore conceivable that within differentiating cells the ability to sequester CDK-cyclin

HPV11 E1   L*KRK*YLG**S**PYV**S**PISNVANAVESEI<u>**S**PRLDAIKL</u>TTQPKKVK**RRL**

HPV18 E1   VL*KRK*FAGGSTEN**S**PLGERLEVDTEL<u>**S**PRLQEISL</u>NSGQKKAK**RRL**

HPV1 E1   ELNAL*KRK*LLY**S**PQARSADETDIASI<u>**S**PRLETISI</u>TKQDKKRYR**RQL**

Figure 4.7: HPV E1 proteins contain conserved sequences within the localisarion regulatory region (LRR). Sequences of the LRR of HPV11, 18 and 1 are shown with putative nuclear localisation sequence (NLS) indicated in blue italics, putative CDK phosphorylation sites indicated in large typeface, nuclear export sequences (NES) are underlined and RXL Cy motifs highlighted in red. Adapted from Deng et al 2004.

complexes does not impede upon the nuclear localisation and replicative ability of the E1 protein.

4.10   The role of E1^E4 inhibition of CDK2-cyclin A activity

Within this study it was proposed that association of the HPV18 E1^E4 protein with CDK-cyclin complexes may, in addition to altering their sub-cellular localisation, alter the kinase activity of these complexes. This study has demonstrated that the presence of HPV18 E1^E4 inhibits the activity of CDK2-cyclin A *in vitro*. In addition to its role during S phase DNA replication, CDK2-cyclin A also plays a role in regulating the G2/M transition and is thought to induce chromosome condensation (Furuno et al., 1999). It is possible that inhibition of CDK2-cyclin A by the HV18 E1^E4 protein, may contribute together with the ability of E1^E4 to sequester CDK2 and cyclin A in the cytoplasm, towards the G2 arrest function of the E1^E4 protein. Indeed, it has been shown that the integrity of the RXL motif in HPV18 E1^E4 is essential for the G2 arrest function of this protein (G.L. Knight and S. Roberts, unpublished data). Inhibition of CDK2-cyclin A may serve to enhance the G2 arrest function by preventing nuclear activity of the kinase. It is also possible that inhibition may act to prevent aberrant phosphorylation of proteins by this kinase as a result of sequestration to the cytoplasm.

The HPV1 E1^E4 protein has been shown to inhibit cellular DNA replication by blocking the recruitment of cellular licensing factors onto chromatin (Roberts et al., 2008). It is hypothesised that inhibition of cellular DNA replication by E1^E4 is acts to generate an environment in which viral DNA replication is favoured over cellular DNA replication to support amplification of viral genomes. The CDK2-cyclin A complex plays an essential role in cellular DNA replication (Coverley et al., 2002) and the ability of the HPV18 E1^E4

protein to inhibit CDK2-cyclin A activity during *in vitro* kinase assays may be a strategy to prevent cellular DNA replication thereby enhancing replication of viral genomes. Although a previous study has indicated a role for E7 proteins of HPV types 6, 16 and 31 in direct activation of CDK2-cyclin A (He et al., 2003), this effect was not observed with the E7 protein of HPV18 and inhibition of CDK2-cyclin A by HPV18 E1^E4 was maintained in the presence of HPV18 E7. It is possible that the ability to directly activate the CDK2-cyclin A complex may not be conserved between all E7 proteins.

4.11 Conclusions

Considering that the bi-partite CDK recognition motif is highly conserved between E1^E4 proteins of alpha HPV types and that the ability to associate with cyclin proteins is also common to a number of E1^E4 proteins, it is highly probable that this interaction mediates an important E1^E4 function. Identification of further E1^E4 interacting proteins may enable the relevance of this interaction to be established. The combined ability of HPV18 E1^E4 to both redistribute CDK2-cyclin A to the cytoplasm together with the ability to inhibit associated kinase activity indicates that E1^E4 may display a dual mechanism for inactivation of this kinase complex and suggests that prevention of CDK activity is an important function of the E1^E4 protein.

4.12 Note

The work detailing the role of the CDK bi-partite recognition sequence in the association of HPV18 E1^E4 with active CDK complexes (Chapters 3 and 4) forms part of a manuscript submitted to the Journal of Virology (Gillian L. Knight[†], Alice G. Pugh[†], Emma Yates, Ian

Bell, Regina Wilson, Laimonis L. Laimins and Sally Roberts. [†]These authors contributed equally to this work).

CHAPTER 5

N-TERMINAL PROTEOLYSIS OF HPV18 E1^E4 IS

DEPENDENT UPON SEQUENCES WITHIN AN

N-TERMINAL LEUCINE-RICH MOTIF

Introduction

5.1   N-terminal proteolysis of the HPV E1^E4 protein

E4 gene products were first identified from examination of HPV1 induced papillomas by Western blot analysis using an antibody raised against the product of the E4 ORF. This revealed that the E4 protein is expressed as a major protein doublet of 17 and 16 kDa, in addition to a minor doublet of 11 and 10 kDa proteins (Doorbar et al., 1986). Two minor protein doublets of 21/23 kDa and 32/34 kDa, were also observed together with an E4 protein of 45 kDa, which may correspond to multimeric E4 species. Subsequent studies examined the expression profile of these multiple E4 species within serial cross sections taken throughout the layers of HPV1 induced warts (Breitburd et al., 1987). The processed E4 proteins are observed in a progressive manner, as the sections of the wart proceed from the deepest basal cells to the uppermost layers. The 17 kDa E4 protein predominates within the parabasal layers, with coincident expression of the 16 kDa product within more differentiated cells. The smaller E4 proteins, the 11 kDa protein, then the 10 kDa protein, become successively more abundant within superficial layers, whilst expression of the 17 kDa protein, then 16 kDa E4 protein subsequently diminishes. Higher molecular mass E4 species of 21/23 kDa, 32/34 kDa and 45 kDa were most abundant in the central layers of the wart. Within this study, it was postulated that the 16 kDa E4 species is derived from the 17 kDa full-length E1^E4 protein by post-translational proteolytic cleavage with proteolysis of the 17/16 kDa doublet resulting in production of the 11 and 10 kDa E4 proteins.

Subsequent analysis using a panel of epitope-mapped monoclonal and polyclonal antibodies revealed that the 16 kDa protein lacked ~15 N-terminal residues whilst the smaller E4 11/10 kDa doublet was not detected with antibodies specific to the N-terminal half of the E4 protein,

indicating that these proteins are lacking further N-terminal residues (Doorbar et al., 1988). A later study, confirmed that the smaller HPV1 E4 proteins are derived by N-terminal proteolysis of the full-length E1^E4 protein, as demonstrated by sequencing of the N-terminal residues of these E4 species (Roberts et al., 1994). The N-terminus of the 16 kDa E4 protein corresponds to tyrosine residue 16 within the full-length 17 kDa E1^E4 protein whilst the 10/11 kDa proteins both commence with alanine residue 59 of the E1^E4 sequence, suggesting that they arise from N-terminal cleavage at an identical sequence (Roberts et al., 1994). The variation between these smaller proteins is likely to be as a result of C-terminal proteolysis or alternatively through another distinct post-translational modification.

N-terminal proteolytic processing appears to be a conserved feature amongst E1^E4 proteins, with multiple species observed in a number of HPV types. Examination of the expression profile of E4 proteins within papillomas derived from the cutaneous HPV types 2 and 4 by Western blot analysis, identified E4 protein doublets with molecular masses of 16.5/18 kDa and 20/21 kDa respectively (Doorbar et al., 1989). These E4 polypeptide pairs are likely to correspond to full-length and processed E4 species, analogous to the 17/16 kDa E4 protein doublet observed with HPV1. Multiple E4 species have also been detected in differentiating cells containing HPV genomes of anogenital types. Western blot analysis of condyloma accuminata (anogenital warts), induced by HPV6 and 11, revealed that the E4 proteins expressed within these tissues are often expressed as an 11/10 kDa doublet (Tomita et al., 1991). The proportion of the major E4 11 kDa protein to the minor 10 kDa protein varied however between tissues and in some specimens only the 11 kDa E1^E4 protein could be detected. Upon incubation of condyloma tissue at 37˚C, partial conversion of the 11 kDa E4 protein to the 10 kDa species was observed, indicative of proteolytic cleavage. Within HPV31 genome-containing keratinocytes a full-length 11 kDa E4 protein could be observed

at low levels within cells grown in monolayer culture, whilst a smaller 8 kDa E4 species was only detected following differentiation of these cells utilising an organotypic raft expression system (Pray & Laimins, 1995).

5.2 N-terminal sequences contribute towards the cellular localisation of E1^E4 proteins

Several studies have demonstrated that sequences at the N-terminus of E1^E4 are important for the cellular distribution of the protein. Over-expression of the full-length HPV1 E1^E4 17 kDa protein in a rabbit keratinocyte cell line VX2R, assembled in a primarily cytoplasmic distribution with formation of E1^E4 associated inclusions observed with a ring-shaped morphology surrounding a bright core (Rogel-Gaillard et al., 1992; Rogel-Gaillard et al., 1993). Inclusions formed by an E4 protein lacking the five N-terminal E1 residues however, displayed a distinct morphology with formation of round bright granular inclusions. An association was observed between the outer edges of inclusions formed by both of these E4 proteins and tonofilaments, which are comprised of keratin intermediate filaments (Rogel-Gaillard et al., 1993).

Further insight into the contribution of N-terminal residues to E1^E4 localisation, was obtained following the revelation that N-terminal sequences of HPV1 and HPV16 E1^E4, mediate association between these protein and the keratin intermediate filament (IF) network (Roberts et al., 1994). E1^E4 proteins form filamentous cytoplasmic networks which co-localise with the keratin IF network, with the HPV16 E1^E4 protein capable of inducing collapse of the these filamentous networks into juxta-nuclear fibrous bundles (Doorbar et al., 1991; Roberts et al., 1993; Roberts et al., 1994). The distribution of HPV1 E4 proteins, equivalent to the 16 kDa and 11/10 kDa N-terminally cleaved proteins were cytoplasmically

diffuse, and the ability to associate with the keratin intermediate filaments was abrogated. Mutational analysis of the E1^E4 N-terminus mapped keratin association to the leucine-rich sequence: [10]LLGLL[14] of HPV1 E1^E4. The leucine-rich motif is highly conserved between many but not all E1^E4 proteins and deletion of this sequence ([12]LLGLL[16]) within HPV16 E1^E4 also abrogated the formation of filamentous networks, indicating that this is likely to be a common sequence mediating association between E1^E4 proteins and the keratin intermediate filaments (Roberts et al., 1997).

5.3   N-terminal sequences of E1^E4 proteins mediate a number of functions

In addition to mediating an interaction between E1^E4 proteins and the keratin cytoskeleton, N-terminal residues of E1^E4 have also been shown to be integral for the performance of other functions. The ability of the HPV1 E1^E4 protein to relocate the promyelocytic leukemia (PML) protein, that characterises sub-nuclear ND10 domains, to E1^E4 nuclear inclusions during the productive phase of the life cycle has also been assigned to N-terminal residues of the E1^E4 protein (Roberts et al., 2003). ND10 bodies are thought to be sites of replication for HPV genomes and also for a number of other DNA viruses including adenovirus and herpes virus (Everett, 2001). It has been proposed that ND10 domain disruption may generate an environment suitable for replication during the productive phase of the HPV life cycle by allowing formation of viral replication centres thereby facilitating viral genome amplification (Swindle et al., 1999). It has also been suggested that ND10 domains may play a role in anti-viral mechanisms (Regad & Chelbi-Alix, 2001), therefore E1^E4 redistribution of these domains may act to counteract this affect (Roberts et al., 2003).

The N-terminal leucine-rich cluster motif has also been implicated in the association of the HPV16 E1^E4 protein with mitochondria (Raj et al., 2004). As a result of this interaction a

redistribution of mitochondria from the microtubule network to collapsed E4 clusters adjacent to the nucleus was observed, and caused a marked reduction in mitochondrial membrane potential with eventual induction of apoptosis. Deletion of the leucine-rich motif of HPV16 E1^E4 prevents the reorganisation of this cytoplasmic organelle and apoptotic effect which is hypothesised to render the cell more fragile and a may aid the release of progeny virus.

Given that a number of functions of E1^E4 proteins have been assigned to the N-terminus of these proteins, removal of these residues by proteolytic cleavage may be a mechanism for regulation of these functions to ensure events occur in a timely manner.

5.3.1   A distinct function for a proteolytically cleaved HPV1 E4 protein

An important development in understanding of E1^E4 function was the assignment of a distinct function to the cleaved product of the HPV1 protein (Knight et al., 2004). Loss of N-terminal residues has been shown to result in a gain of function, with the HPV1 E4 N-terminally truncated protein displaying a distinct function from that of the full-length E1^E4 protein (Knight et al., 2004; Knight et al., 2006). E1^E4 proteins of a number of HPV types have been shown to induce a G2 arrest within over-expression studies. Whilst expression of the full length 17 kDa E1^E4 protein of HPV1 does not perturb progression of the cell-cycle, the truncated 16 kDa E4 species, derived by N-terminal proteolysis of the full-length protein, is adept at inducing a G2 arrest, either alone or when co-expressed with the full-length E1^E4 protein (Knight et al., 2004). The cooperation between the full-length and a truncated HPV1 E4 demonstrates unique roles for the truncated version of the E4 protein and highlights the importance of these proteolytic processing events for the function of the HPV1 E1^E4 protein.

5.4   **Rationale for investigation of HPV18 E1^E4 proteolysis**

The formation of multiple E4 species of smaller molecular masses than the full-length E1^E4 protein, appears to be a common event observed with diverse HPV types and suggests that proteolytic cleavage may be a common mode of regulation employed by the virus to modify the role of the E1^E4 protein during the virus life cycle. Whilst the generation of an N-terminally truncated E4 protein with a distinct function from that of the full-length HPV1 E1^E4 protein has been well characterised, little is known about these events in anogenital viruses therefore proteolysis of E1^E4 of a high risk virus, HPV18 was investigated to gain a greater insight into the role of this protein during the virus life cycle.

Results

5.5   **Development of HPV18 E1^E4 expression systems for characterisation of**
             E1^E4 proteolysis

Expression of E1^E4 proteins of several anogenital types can prove to be problematic as considerably low levels of the these proteins have been reported to be generated by transient transfection (Roberts et al., 1993; Raj et al., 2004). Inadequate expression can impede analysis of the biological activity of the E1^E4 proteins but can be partially overcome by expression of E1^E4 proteins by recombinant viruses which efficiently deliver genes into a cell and can result in high level expression of E1^E4 protein (Roberts et al., 1993; Knight et al., 2004; Raj et al., 2004).

Whilst a recombinant SV40 virus expressing HPV18 E1^E4 was successfully generated (Appendix II), the level of protein expressed in COS-1 cells was too low to be useful for the

analysis of HPV18 E1^E4 proteolysis therefore alternative expression systems were investigated.

5.5.1 HPV18 E1^E4 expressed with a recombinant adenovirus expression system undergoes N-terminal proteolysis

Recombinant adenoviruses (rAd) have been used effectively in the study of HPV1 E1^E4 proteolysis (Knight et al., 2004; Knight et al., 2006) and a rAd-HPV18 E1^E4 virus was subsequently generated by Dr. G. L. Knight. At the beginning of my studies this was the only method available within our research group for expression of the HPV18 E1^E4 protein so initial studies were performed using rAd-HPV18 E1^E4. Following infection of the SCC-12F human squamous cell carcinoma line with rAd-HPV18 E1^E4 a major E1^E4 species was detected by 72 h post infection (PI) by Western blot analysis with a polyclonal E1^E4 antibody, which was not detected in cells infected with the rAd-βgal control (Figure 5.1 A). The molecular mass of the E1^E4 protein was estimated to be 11.7 kDa. There was a progressive accumulation of this major E1^E4 species at later times PI coincident with the detection of additional faster migrating E4 polypeptides of 10.5 kDa and 9.5 kDa. These lower molecular mass E4 species however, whilst being recognised by a polyclonal E1^E4 antibody, were not detectable with the N-terminal specific antibody, suggesting that they lack N-terminal residues. It was hypothesised that these E4 species are derived by N-terminal proteolytic cleavage of the full-length HPV18 E1^E4 protein. Higher molecular mass E4 species were also observed which may represent E4 multimers of both full-length and processed E4 species.

A higher level of E1^E4 expression was observed upon infection of the adenovirus 5 E1A-transformed epithelial cell line, 911 with rAd-HPV18 E1^E4 for 96 h and a similar profile of
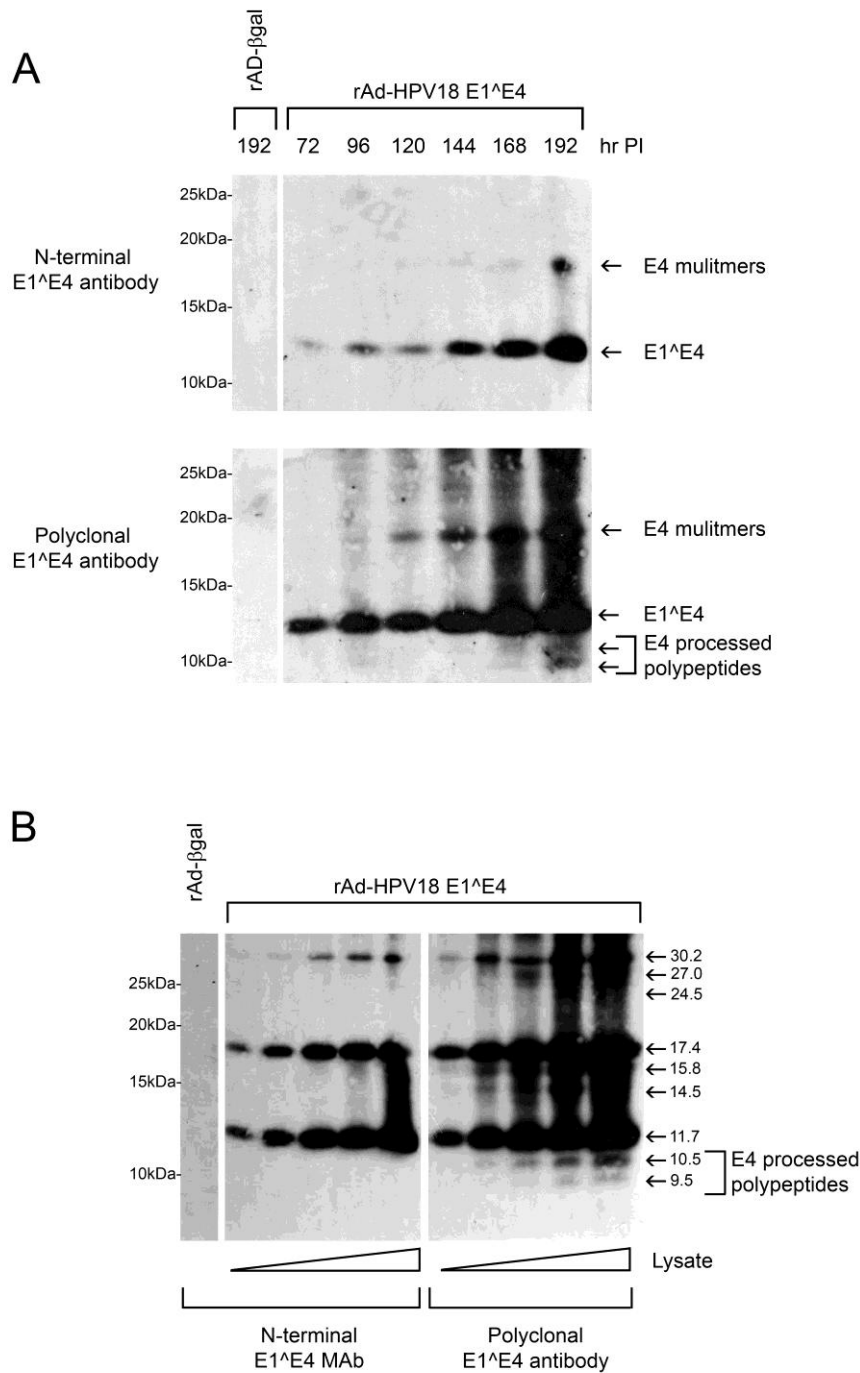
Figure 5.1: N-terminal truncations of HPV18 E1^E4 are detected following infection of epithelial cells with a recombinant HPV18 E1^E4 adenovirus. **(A)** The expression profile of HPV18 E4 species in SCC-12F cells was examined by Western blot analysis, following infection with rAd-HPV18 E1^E4. Full-length E1^E4 is detected by both a polyclonal E1^E4 antibody and an N-terminally specific E4 antibody, whilst faster migrating E4 species are detectable by 168 h post-infection (PI), only with the latter antibody. **(B)** Following infection of 911 epithelial cells with rAd-HPV18 E1^E4 for 96 h, increasing amounts of cellular lysate, from 2-30 μg, were analysed by Western blot analysis. Faster migrating E4 species were detectable in addition to the full-length E1^E4 protein. In addition, multiple higher molecular weight E4 proteins were observed with the polyclonal E1^E4 antibody, whilst only a subset of these proteins are detectable with an N-terminal specific E4 antibody. Western blots are representative of three independently performed experiments.

E4 species were observed, with full-length E1^E4 and faster migrating E4 species (Figure 5.1 B). In addition to the major E1^E4 species of 11.7 kDa and the faster migrating E4 species lacking N-terminal residues, numerous higher molecular mass species were also observed. Two major higher molecular mass E4 species of 17.4 and 30.2 kDa, are likely to correspond to dimeric and trimeric complexes although there is however a disparity between the observed molecular mass, and the predicted sizes of 23.4 and 35.1 kDa respectively. This discrepancy has also been observed with the analysis of multimeric complexes of HPV1 E1^E4 (Doorbar et al., 1996; Ashmole et al., 1998). It is probable that the minor higher molecular mass species of 27.0, 24.5, 15.8 and 14.5 kDa, observed only with the polyclonal E1^E4 antibody, equate to multimeric E4 complexes formed by various permutations of N-terminally truncated E4 species or E4 proteins with differing combinations of other post-translational modifications such as phosphorylation.

However by expressing the HPV18 E1^E4 protein using a recombinant adenovirus it must be acknowledged that is possible that this may affect the expression and post-translational modification of the E1^E4 protein. The expression profile of the E1^E4 protein observed within these experiments may therefore be affected by the use of the adenovirus vector.

5.5.2 High level expression of HPV18 E1^E4 from a codon-optimised cDNA results in the production of smaller E4 species derived by N-terminal proteolysis of the full-length protein

During my studies, a codon-optimised HPV18 E1^E4 expression construct was generated by Dr. I Bell in which a number of codons were converted within HPV18 E1^E4 cDNA (coE1^E4) to those more commonly used in genes of the human genome (Appendix III). A greatly enhanced expression of the E1^E4 protein was observed within epithelial cells

transfected with coE1^E4, in comparison with unmodified E1^E4 cDNA, with the greatest expression of coE1^E4 observed in 293T cells, a human embryonic kidney cell line containing the SV40 Large-T antigen. This enabled analysis of E1^E4 expression by transient transfection of cDNA, an approach that has a number of advantages over recombinant virus expression systems. Mutations can be directly introduced into the E1^E4 cDNA, and the addition of epitope tags to the protein can be carried out conveniently, enabling analysis of E1^E4 function without the lengthy requirement of generating new recombinant virus. Following generation of the codon-optimised HPV18 E1^E4 expression construct, subsequent experiments to analyse expression of the HPV18 E1^E4 protein by Western blot analysis were preferentially performed by transfection of 293T cells with coE1^E4.

Expression of coE1^E4 was analysed by Western blot analysis following transfection of 293T cells. In addition to the major E1^E4 observed at 11.7 kDa, the faster migrating E4 processed polypeptides of 10.5 and 9.5 kDa were also observed with a polyclonal E1^E4 antibody, which were not detected in pcDNA control transfected cells (Figure 5.2 A). In the absence of the reducing agent β-Mercaptoethanol (βME), multiple higher molecular mass E4 species can be observed (Figure 5.2 B). However following reduction of the cell lysate with βME only the major E1^E4 species and the processed E4 are observed indicating that the higher molecular mass species are E4 multimeric proteins.

To verify that proteolysis removes N-terminal sequences, a coE1^E4 construct was generated in which the HPV18 E4 protein is tagged at the N-terminus with a Haemagglutinin (HA) epitope tag (N-HA E1^E4). Following expression in 293T cells, multiple E4 species were observed by Western blot analysis with a polyclonal E1^E4 antibody. However, only the
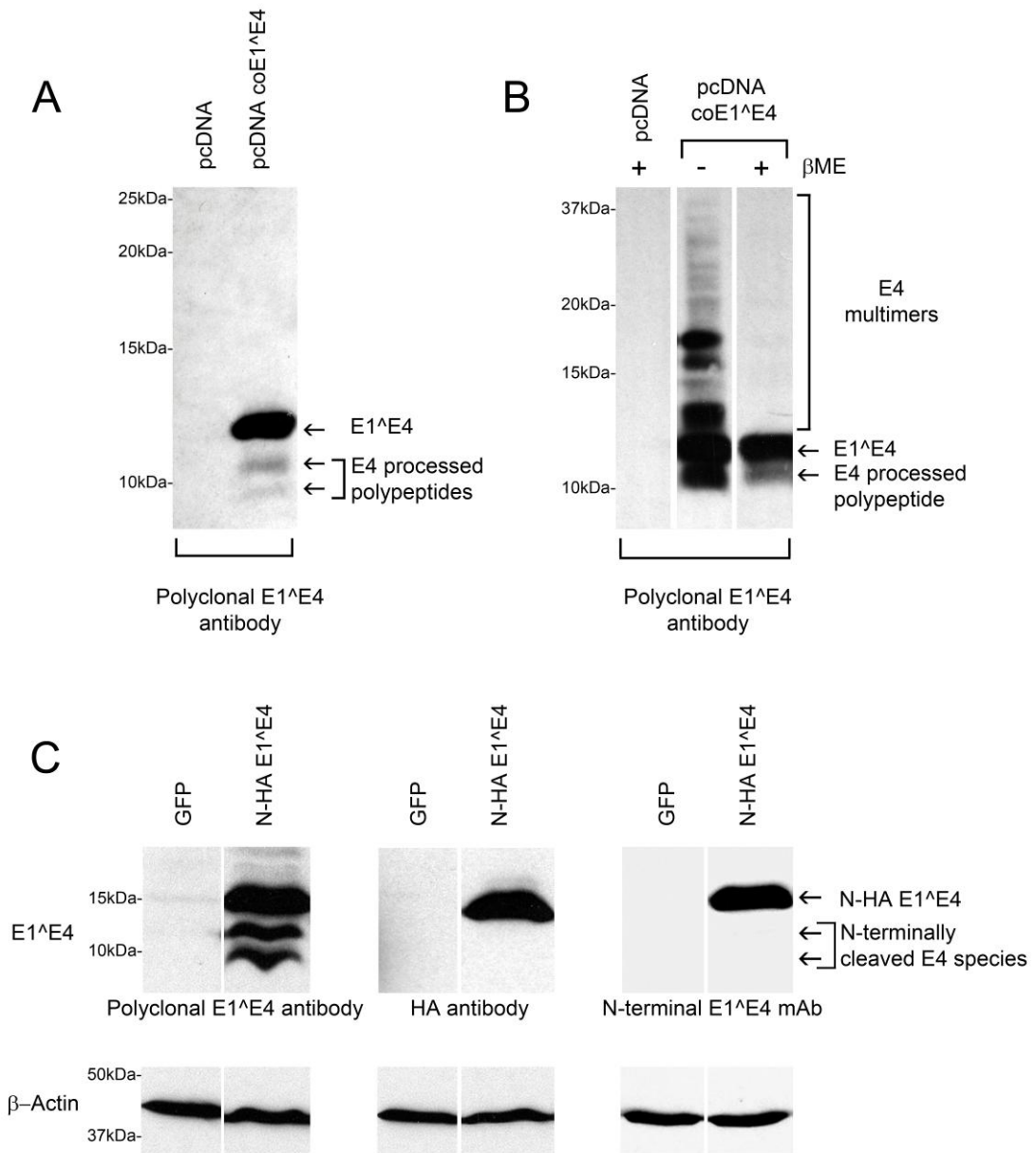
Figure 5.2: HPV18 E1^E4 expressed from codon-optimised cDNA, undergoes N-terminal proteolysis and forms multimeric complexes. **(A)** Following transfection of 293T epithelial cells with codon-optimised HPV18 E1^E4 (coE1^E4), a major E1^E4 species of 11.7 kDa was observed by Western blot analysis, in addition to lower molecular weight species of 10.5 and 9.5 kDa. **(B)** 293T cells transfected with coE1^E4 were lysed in the presence or absence of β-Mercaptoethanol (βME). In the absence of βME higher molecular weight E4 species were observed, whils in the presence of the reducing agent, these E4 species are no longer detected indicating that they correspond to E4 multimeric complexes. **(C)** Following expression of N-terminally, HA tagged E1^E4 in 293T cells, full-length E1^E4 is detected with a HA antibody, an N-terminal specific E1^E4 antibody and a polyclonal E1^E4 antibody. Lower molecular weight E4 species are detectable only with the latter antibody, indicating that they are N-terminally truncations. β-Actin demonstrate equal loading of protein samples. Representative Western blots of three independently performed experiments.

major E1^E4 species was detected with an HA and N-terminal specific E4 antibody (Figure 5.2 C), indicating that both the HA tag and N-terminal E1^E4 residues have been released from the smaller E4 species by proteolysis. These smaller E4 species lack N-terminal sequences and are therefore derived from N-terminal proteolytic cleavage of the full-length E1^E4 protein.

5.6 A functional role for HPV18 E1^E4 N-terminal sequences

5.6.1 Expression of an N-terminally truncated HPV18 E4 protein, based on prediction of the site of E1^E4 proteolysis

To determine whether loss of these N-terminal sequences alters the localisation of the HPV18 E1^E4 protein, a truncated protein species was generated. The position of proteolysis within the HPV18 E1^E4 species is unknown, however identification of the N-termini of the proteolytically processed 16 kDa protein of HPV1 E4 by sequencing, revealed that the full-length HPV1 E1^E4 protein undergoes proteolysis following glutamine residue 15 (Roberts et al., 1994). Sequence comparisons of anogenital HPV E1^E4 proteins including HPV18, together with the HPV1 E1^E4 sequence, reveals that there is considerable homology at the point at which the latter protein is cleaved to generate the truncated species (Figure 5.3 A). A conserved sequence of LLXLLX (S/T/Y) (S/T/P) is located within this region of E1^E4 proteins in which a leucine-rich sequence is followed by two hydroxyl containing residues. Based on homology with the type I protein, a potential cleavage site was identified within HPV18 E1^E4, following asparagine residue 19 (Figure 5.3 A). A deletion mutant was generated in the HPV18 coE1^E4 expression vector, based on this predicted point of cleavage, in which the N-terminal residues up to this residue, excluding the initiation codon, are deleted (Δ2-19).

A

```
                                      Position of
                                       cleavage

                                          ↓

     HPV 1    MADNK APQG      LLGLL   Q  *  YTPTTQPYPRVT

     HPV 18   MADPE VPVTTRYP  LLSLL   N     SYSTPPHRIPAP
     HPV 11   MADDS ALYEKYP   LLNLL   H     TPPHRPPPLQCP
     HPV 16   MADPA AATKYP    LLKLL   G     STWPTTPPRPIP
     HPV 31   MADPA AVTKYP    LLGLL   Q     SYQQPTTPPHRI
     HPV 45   MADPE VPVTTRYP  LLRLL   D     SYNTPPRRPPKP

            Conserved residues:   LLXLL    X     SP
                                                 TT
                                                 YY
```

B
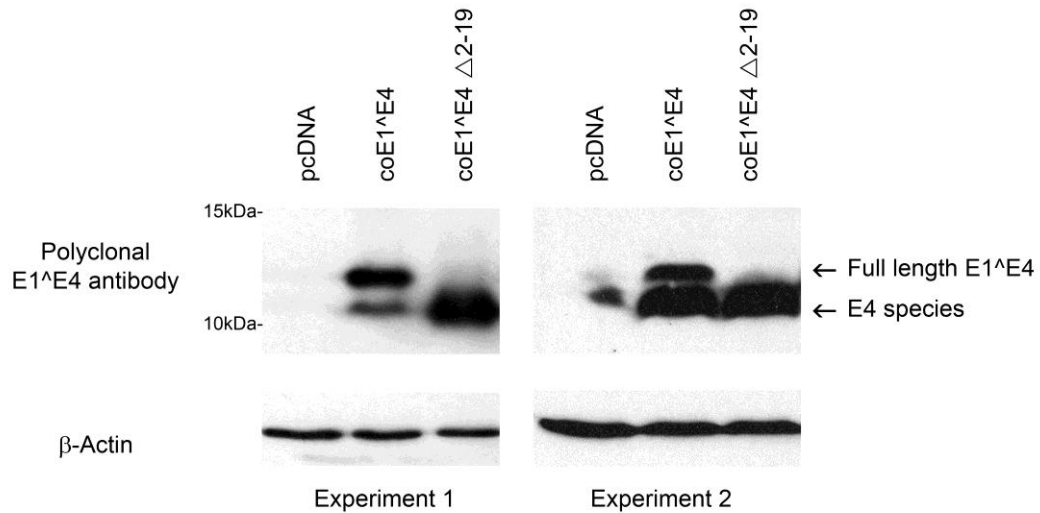
Figure 5.3: Prediction of the site of proteolytic cleavage in the HPV18 E1^E4 amino acid sequence and expression of a corresponding N-terminal deletion mutant protein. **(A)** Sequence comparisons between HPV1 E1^E4 and E1^E4 proteins of several anogenital HPV types were used to predict a putative site of proteolysis in the HPV18 E1^E4 sequence, in accordance with the known point of cleavage in HPV1 E1^E4. **(B)** Western blot analysis with a polyclonal E1^E4 antibody, of 293T cells transfected with codon-optimised (co) E1^E4, or coE1^E4 Δ2-19. β-Actin demonstrates equal loading of protein samples. The processed E4 polypeptide derived by N-terminal proteolysis of the full-length E1^E4 migrates at a similar molecular weight as the Δ2-19 E4 protein. Two representative experiments are shown, demonstrating the variation in the extent of E1^E4 proteolysis observed between experiments.

Following Western blot analysis of lysate derived from 293T cells transfected with coE1^E4 Δ2-19, the N-terminal truncation mutant was detected with a molecular mass of 10.5 kDa (Figure 5.3 B). The processed E4 polypeptide, derived by N-terminal cleavage of the full-length E1^E4 protein, migrates at a similar molecular mass, suggesting that the N-termini of the cleaved protein is likely to be near to serine residue 20. The N-terminally truncated Δ2-19 deletion mutant is therefore likely to closely resemble, and may emulate, the naturally occurring processed E4 protein. It is interesting to note that the proportion of N-terminally processed E4 to full-length E1^E4 can vary significantly between experiments.

5.6.2 An N-terminally truncated E4 protein associates with the full-length E1^E4 polypeptide

Given that higher molecular mass E4 species are detectable, which correspond to E4 multimeric complexes it is possible that full-length HPV18 E1^E4 is able to associate with the N-terminally truncated E4 protein. Following transfection of 293T cells with an N-terminally tagged HA E1^E4 construct or a Δ2-19 E4 construct, cells were harvested at 48 h post-transfection and immunoprecipitation was performed with an HA antibody. Western blot analysis of HA immune complexes with a polyclonal E1^E4 antibody revealed that whilst the full-length HA-tagged E1^E4 protein was successfully isolated; an extremely low level of the untagged Δ2-19 E4 protein was non-specifically precipitated (Figure 5.4). However in co-transfected cells, both the full-length E1^E4 and smaller E4 protein were immunoprecipitated in equal amounts, indicating that HPV18 E1^E4 is able to associate with an N-terminally truncated E4 protein to form a heteromeric complex.
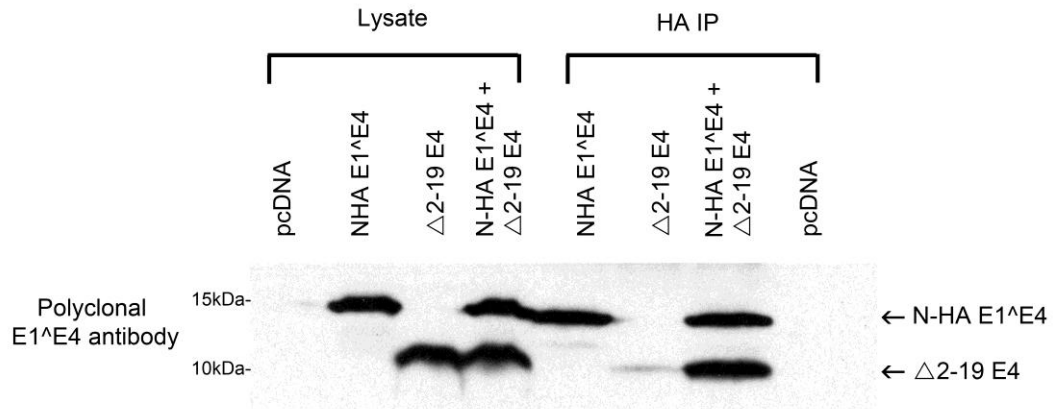
Figure 5.4: HPV18 full-length E1^E4 and N-terminally truncated Δ2-19 E4 protein associate *in vitro.* Codon-optimised E1^E4, tagged at the N-terminus with a HA tag (N-HA E1^E4) and untagged coE1^E4 Δ2-19 expression constructs were transfected into 293T epithelial cells. At 48 hours post-transfection, HA immunoprecipiates were isolated and and lysate and immune complexes were analysed by Western blot analysis with a polyclonal E1^E4 antibody. N-HA E1^E4 co-precipitated with the Δ2-19 E4 protein. This result is representative of two independently performed experiments.

5.6.3 Loss of the N-terminus of HPV18 E1^E4 does not abrogate association with keratin intermediate filaments

To investigate the importance of N-terminal residues in determining the intracellular distribution of the HPV18 E1^E4 protein, the full-length E1^E4 and the Δ2-19 E4 truncation mutant were transfected into COS-1 cells and their localisation was examined by immunofluorescence microscopy with a polyclonal E1^E4 antibody (Figure 5.5). COS-1 cells were used to allow examination of E1^E4 localisation within epithelial cells and transfection efficiences proved to be good within this cell line. HPV18 E1^E4 localisation has been characterised in Chapter 3 and typically displays a filamentous distribution, co-localising with the keratin intermediate filament network, resulting in collapse into juxta-nuclear bundles (Figure 5.5). The Δ2-19 deletion mutant E1^E4 protein possesses a similar phenotype to the wildtype E1^E4 protein with co-localisation with the keratin IF network observed and collapse of E1^E4 filaments into fibrous bundles (Figure 5.5). Loss of N-terminal sequences of HPV18 E1^E4 therefore does not abrogate K18 co-localisation, confirming the findings of a previous study (Nakahara et al., 2002). Whilst within a number of cells expressing wild type E1^E4, the localisation of the E1^E4 protein is marked with punctate, bright cytoplasmic foci, there is a greater frequency of this phenotype observed within cells expressing the Δ2-19 deletion mutant E1^E4 protein (Figure 5.5), indicating that N-terminal residues contribute towards the cellular localisation of the HPV18 E1^E4 protein.

5.7   Identification of residues required for proteolytic cleavage within the primary sequence of HPV18 E1^E4

To examine the physiological importance of E1^E4 proteolysis during the HPV18 virus life cycle, it is necessary to identify residues required for N-terminal cleavage of the full-length
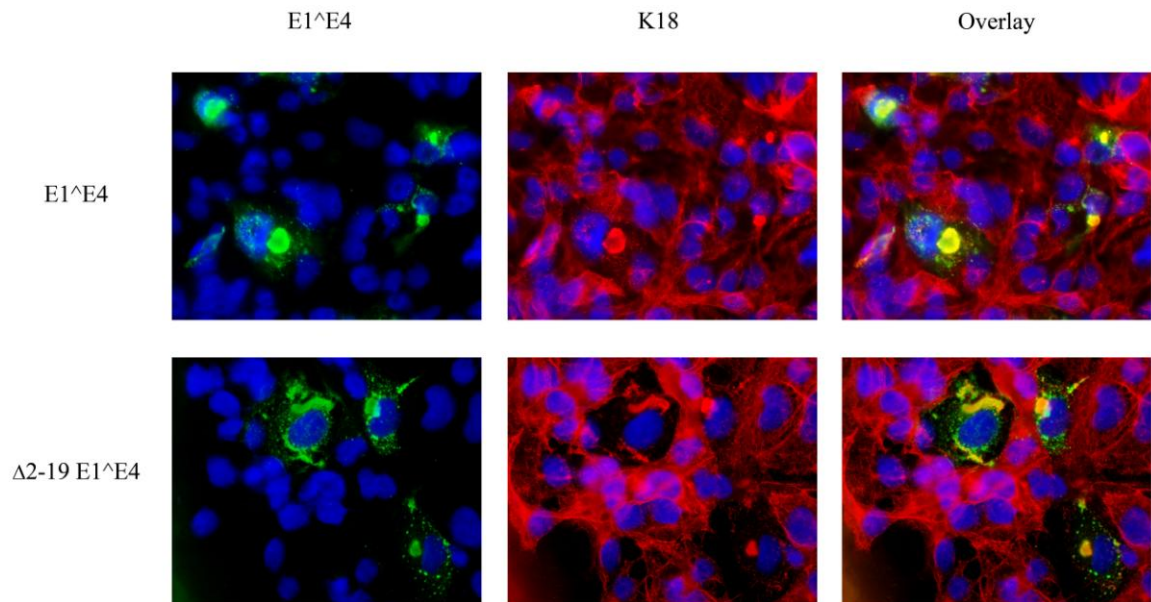
Figure 5.5: N-terminal residues are not essential for co-localisaton of HPV18 E1^E4 with the keratin network. Following transfection of COS-1 cells with the Δ2-19 E1^E4 deletion mutant, cells were fixed at 36 h post transfection in 4% paraformaldehyde for 5 min, prior to permeabilisation in acetone for 10 min. Dual-staining was performed with a polyclonal E1^E4 anti-rabbit antibody and a keratin 18 anti-mouse antibody. Immune complexes were visualised with an anti-rabbit Alexa Fluor 488 antibody (green) and an anti-mouse Alexa Fluor 596 antibody (red) and counter-stained with Dapi for detection of nuclei (blue).

protein. These sequence requirements can subsequently be utilised to enable the generation of an E1^E4 protein which is unable to undergo proteolytic cleavage in the context of the HPV18 genome. The role of E1^E4 proteolysis during the virus life cycle can then be investigated by introducing mutant HPV18 genomes into primary human foreskin keratinocytes (HFK) to generate a stable cell line and vegetative viral functions can be examined within these cell lines following cellular differentiation.

5.7.1 Isolation and sequencing of HPV18 E1^E4 proteolytic species

Identification of the position of proteolytic cleavage within the primary amino acid sequence of HPV18 E1^E4 is required to allow generation of an E1^E4 protein unable to undergo proteolysis. N-terminal Edman protein sequencing of the smaller E4 polypeptides would provide the N-terminal sequence of these proteins and reveal the location of proteolytic cleavage sites within HPV18 E1^E4. To prepare the smaller E4 species for Edman degradation sequencing, sufficient amounts of these cleaved products must be isolated to enable the protein bands to be detectable by Ponceau or Coomassie staining. Initially the recombinant adenovirus expression system was utilised, with infection of 911 cells by rAd-HPV18 E1^E4, however, immunoprecipitation of the processed E4 proteins with a polyclonal E1^E4 antibody proved to be poor (data not shown). To improve the efficiency of E4 immunoprecipitation, HPV18 E1^E4 was cloned into a range vectors to enable expression of E1^E4 proteins C-terminally tagged with: FLAG, 6 X HIS and HA epitope tags, (the latter vector was a kind gift of Dr. I. Bell) to allow isolation of E4 with commercially-generated antibodies. Following expression of these E1^E4 epitope-tagged proteins in 293T cells, immunoprecipitation conditions were optimised and a number of different lysis buffers were tested for the recovery efficiency of E4 proteins by immunoprecipitation (data not shown).

Immunoprecipitation of C-terminally tagged HA-E1^E4 proteins in an NP40 buffer with an HA antibody proved to be the most efficient, and isolated the greatest amount of both the full-length and an N-terminally processed E4 protein (Figure 5.6 A).

To enable sufficient quantity of E4 proteins to be isolated for N-terminal sequencing, a large-scale experiment was carried out with transfection of approximately 1×10^8 293T cells. Following transfection of the E1^E4-HA, or the RAC-HA control, an epitope-tagged RAC protein of the Rho family GTPases (a kind gift of Dr. D. Powner), constructs into 293T cells and subsequent immunoprecipitation with an HA antibody, the resultant precipitates were resolved by a 10-17.5% acrylamide gradient gel, to ensure good separation of proteins. Subsequent silver-staining of the gel allowed detection of a major band of approximately 20 kDa, derived from immunoprecipitation of RAC-transfected 293T cells, corresponding to RAC-HA, whilst a protein band, although extremely faint in comparison, was observed at approximately 13 kDa in the E1^E4-HA experiment (Figure 5.6 B, indicated *3). This silver-stained band is likely to correspond to full-length E1^E4-HA protein, and a second band of approximately 11 kDa (*4), to the proteolytically cleaved E4-HA protein, as these are not detected in the RAC control experiment. In addition two other bands were detected in the E1^E4-HA experiment which were either more intense than band present, or undetectable in the RAC-HA control lane (*1 and *2), which may therefore contain E4 proteins.

All four bands (*1-4) were extracted and sequenced by Fourier transform ion cyclotron resonance mass spectrometry (FTICR) MS/MS. Whilst E1^E4 sequences were not detected in bands 1 or 2, the HPV18 E1^E4 peptides: [44]RLLHDLDTVDSRR[56] and [77]KDGNSVVVTLRL[88] were detected in both lower bands 3 and 4 which correspond to C-
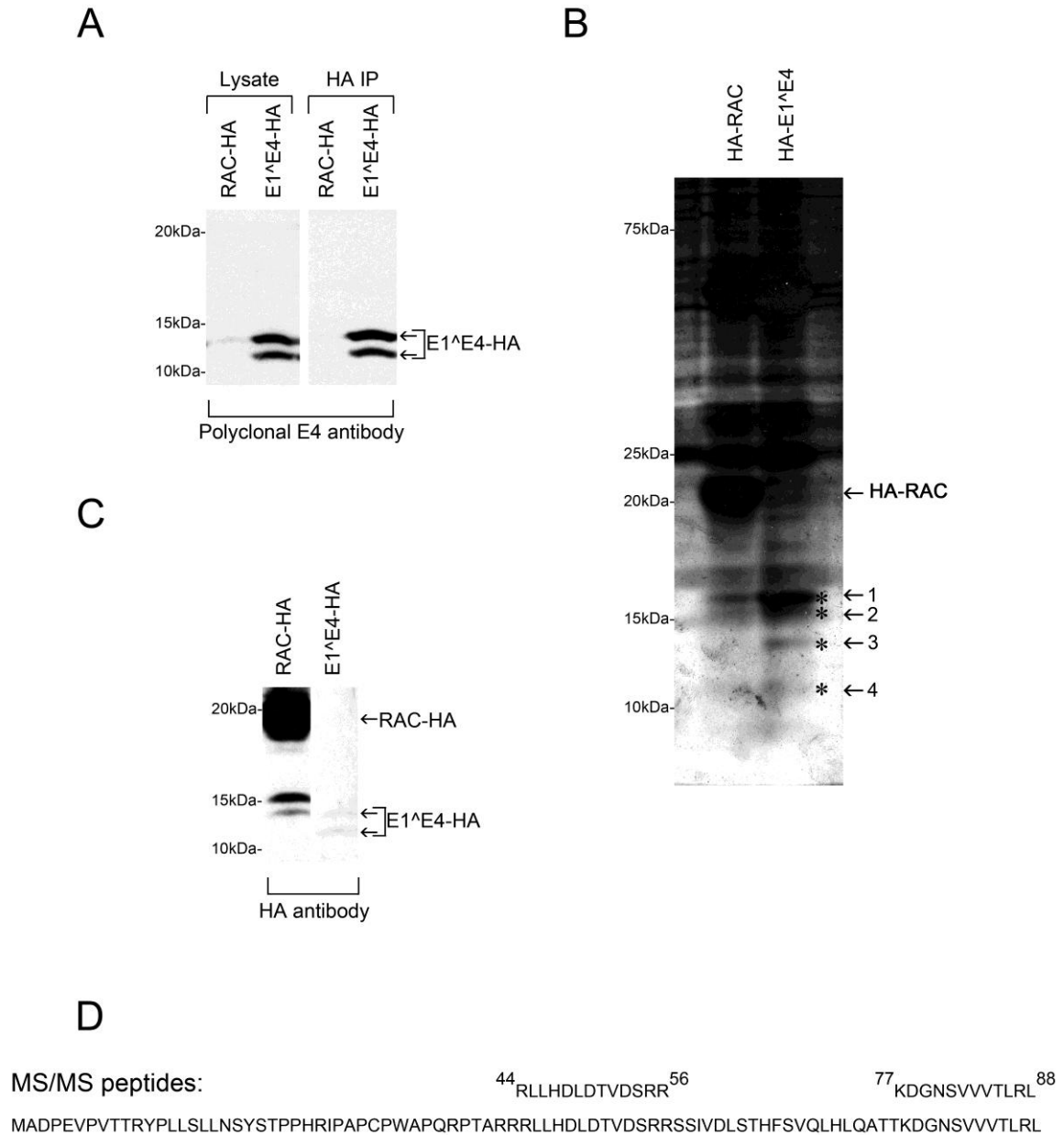
Figure 5.6: Immunoprecipitation and sequencing of C-terminally tagged HA-E1^E4 proteins. E1^E4 or RAC, HA-tagged expression vectors were transfected into 293T cells. 48 hours post-transfection, immunoprecipitation of cellular lysates was performed in an NP40 lysis buffer, with an HA antibody. **(A)** 10% of the immune complexes were analysed by Western blot analysis with a polyclonal E1^E4 antibody. **(B)** The remaining 90% of the immunoprecipitation was separated with a 10-17.5% acrylamide gradient gel. Proteins were visualised by silver staining of the gel and bands 1-4 were extracted for sequencing by FTICR MS/MS. **(C)** The relative levels of RAC and E1^E4 present in the input lysate are shown by Western blot analysis with an HA antibody, indicating that E1^E4 expression is exceedingly low in comparison with RAC protein levels. **(D)** HPV18 E1^E4 peptides detected in bands 3 and 4 by FTICR MS/MS.

187

terminal sequences of the E4 protein (Figure 5.6 D). However these peptides were not the major protein component of these bands detected by MS/MS due to non-specific immunoprecipitation of cellular proteins. Silver-staining is a more sensitive method of protein detection than Coomassie, and although this experiment was repeated with 2×10^8 cells, protein bands corresponding to E4 proteins were not detectable with Coomassie (data not shown), therefore the quantity of N-terminally processed E4 isolated in these experiments, is insufficient for protein sequencing.

During immunoprecipitation experiments, whilst a large and intense band is detectable by silver stain, corresponding to RAC-HA, the quantity of protein present in the E4-containing bands is extremely poor in comparison, despite reasonable immunoprecipitation (Figure 5.6 C). Examination of RAC-HA and E1^E4-HA expression in these experiments by Western blot analysis with an HA antibody revealed that the level of E1^E4-HA is exceedingly low compared to RAC-HA. The type 18 E1^E4 protein is a largely insoluble, making immunoprecipitation in non-denaturing conditions difficult, and the yield of N-terminally processed E4 generated and recovered by this method is low.

5.7.1.1 Investigation of the ability of HPV18 E1^E4 to undergo proteolysis *in vitro*

In an attempt to improve the yield of processed E4 obtained by immunoprecipitation, efforts were made to increase the proportion of the full-length E1^E4 undergoing proteolysis to enable a greater recovery of the processed E4 protein. Whilst transfection times of 48 h were optimal for production of the processed E4 species, it was envisioned that incubation of the full-length E1^E4 protein in cellular lysate in which the protease was contained, could conceivably result in increased proteolysis. Incubation of HPV6 and HPV11 condyloma acuminate tissue at 37°C, resulted in a partial conversion of the E4 11 kDa protein to a 10

kDa protein, indicating that the E4 protein had undergone proteolysis *in vitro* (Tomita et al., 1991).

To establish whether the full-length HPV18 E1^E4 polypeptide can be converted to the processed E4 product *in vitro*, cellular lysate, derived from 293T cells transiently expressing E1^E4, underwent three cycles of freeze/thawing before incubation at 25 or 37°C for various periods of time, as indicated (Figure 5.7). Following incubation at 25°C, at which protease activity would be expected to be sub-optimal, the levels of the full-length E1^E4 and processed E4 proteins were consistent over the duration of up to 240 min (Figure 5.7 A). Whilst it was expected that the level of the processed E4 species would increase, relative to the full-length protein, upon incubation at 37°C, at which proteases are active, a slight decrease in the processed species, occurred over the duration (Figure 5.7 B). The processed E4 species may therefore have a greater instability than that of the full length E1^E4 protein. Whilst this observation contrasts with the expected outcome, if the protease targeting E1^E4 for proteolysis was inactivated during the process of repeated freezing and thawing, this approach would not generate a greater yield of processed E4 for purification.

Knowledge of the protease responsible for inducing N-terminal proteolysis of HPV18 E1^E4 could prove to be valuable in identification of the point of cleavage within the protein. By utilising the protease to proteolytically cleave the E1^E4 protein a greater yield of the processed E4 could be obtained. A range of protease inhibitors were utilised in an attempt to identify the target protease however these studies proved to be unsuccessful (Appendix IV).

In conclusion, it was deemed unlikely that identification of N-terminal residues of the proteolytically cleaved E4 species would be elucidated by the method of sequencing of this truncated protein and a second approach was therefore instigated.
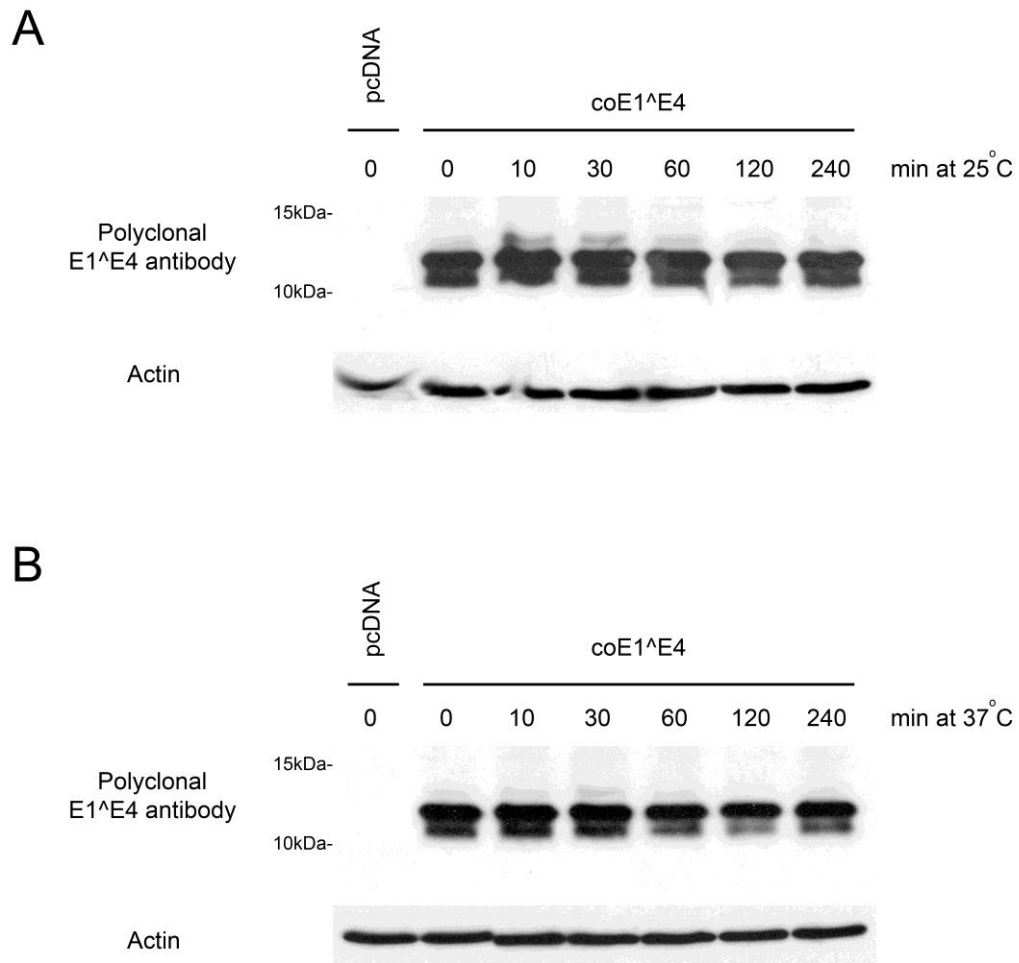
Figure 5.7: Incubation of HPV18 E1^E4 at 37°C in cellular lysate does not increase N-terminal proteolysis. 293T cells transfected with codon-optimised (co) E1^E4 for 48 h, were subjected to three cycles of freezing at -80°C followed by thawing on ice. Lysates were subsequently incubated at **(A)** 25 or **(B)** 37°C for indicated times. Western blot analysis was carried out with a polyclonal E1^E4 antibody, indicating that the proportion of processed E4 relative to the full-length protein did not increase over this period, and was slightly reduced at later time points following incubation at 37°C. β-Actin demonstrates equal loading of protein samples and Western blots are representative of two independently performed experiments.

5.7.2   Identification of E1^E4 mutant proteins that are unable to undergo N-terminal proteolysis

A mutagenesis approach could identify the residues required for E1^E4 cleavage, which although may only predict the cleavage site, would allow mutant HPV18 genomes to be generated in which E1^E4 is unable to undergo proteolysis.   However this approach would not conclusively locate the N-termini of the processed E4 species.

5.7.2.1   Deletion of the leucine-rich motif abrogates proteolysis of the HPV18 E1^E4 protein

Following expression of the Δ2-19 E1^E4 N-terminal mutant in 293T cells, Western blot analysis revealed that the protein has a similar molecular mass to that of the processed E4 polypeptide (Figure 5.3 B), suggesting that N-termini of the processed polypeptide is likely to be near to serine residue 20. Extensive mutagenesis analysis was therefore conducted within this region of the protein, with a series of deletion mutants generated in the coE1^E4 expression vector and the proteins were expressed within 293T cells (Figure 5.8). The cells were lysed in a urea-containing buffer to solubilise E4 and Western blot analyses were performed with both a polyclonal E1^E4 antibody and an N-terminal specific antibody, to establish whether these deletion mutants of E1^E4, retain the ability to undergo N-terminal proteolysis (Figure 5.8 B).

Deletion of a series of overlapping tri-peptides across E1^E4 residues 18-25 did not prevent proteolytic cleavage of the full-length E1^E4 protein, with smaller E4 polypeptides detectable with a polyclonal E1^E4 antibody. The N-terminal specific antibody detected only the major E1^E4 species, although not all of the deletion mutants were detectable due to mutation of the epitope that the antibody recognises. Whilst some variation between the extent of E4

A

E1^E4	^{1}MADPEVPVTTRYPLLSLLNSYSTPP^{25}
Δ2-19	^{2}ADPEVPVTTRYPLLSLLN^{19}
ΔYP	^{12}YP^{13}
ΔLLSLL	^{14}LLSLL^{18}
ΔLNS	^{18}LNS^{20}
ΔNSY	^{19}NSY^{21}
ΔSYS	^{20}SYS^{22}
ΔYST	^{21}YST^{23}
ΔSTP	^{22}STP^{24}
ΔTPP	^{23}TPP^{25}

B
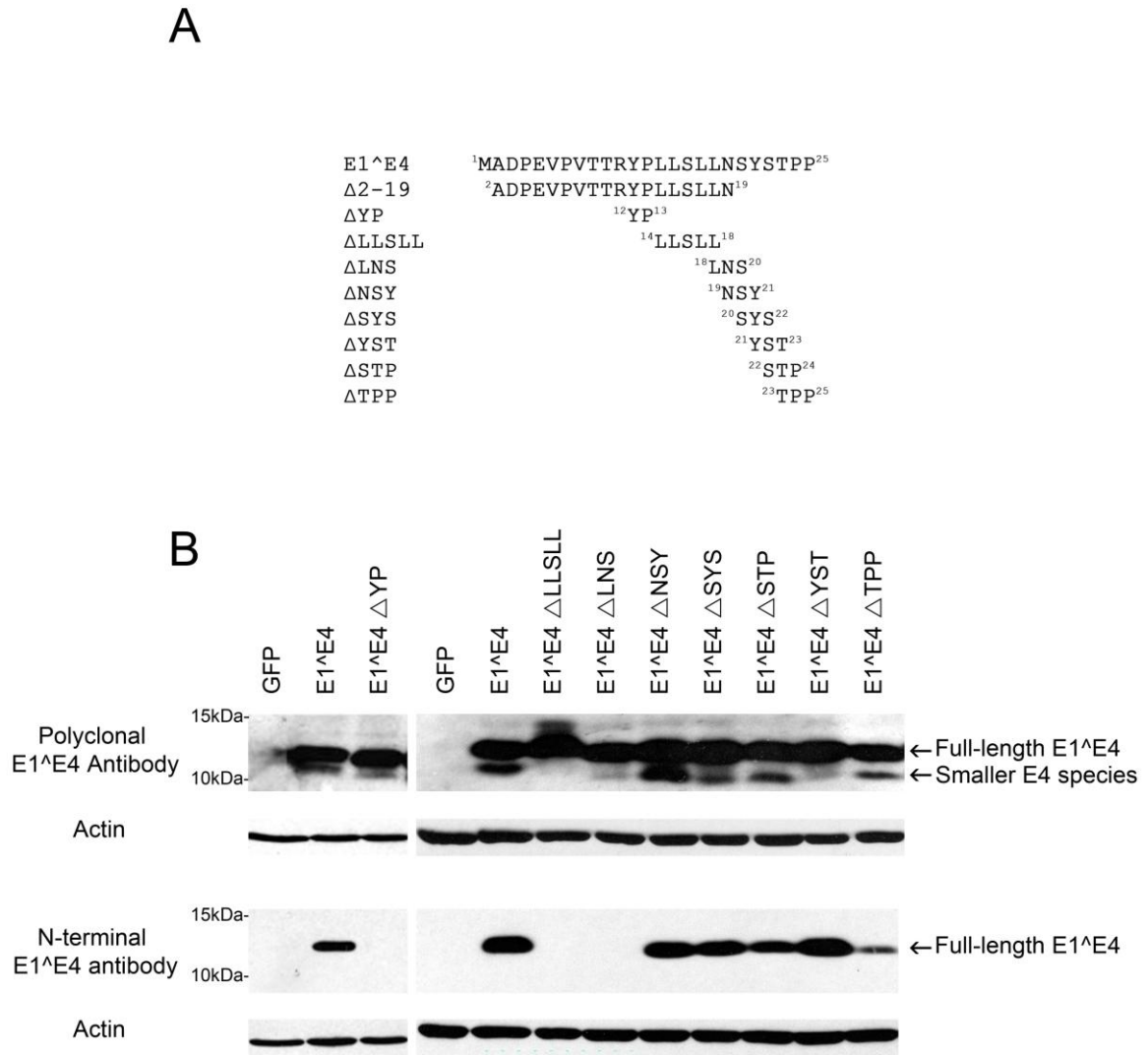
Figure 5.8: The leucine-rich motif of HPV18 E1^E4 is required for N-terminal proteolysis. **(A)** Depiction of HPV18 E1^E4 deletion mutants generated within the codon-optimised E1^E4 expression construct, showing the positions of deleted sequences. **(B)** 293T cells transfected with indicated deletion mutants, were harvested 48 h post transfection in urea lysis buffer. Western blot analysis was performed with a polyclonal E1^E4 antibody and a monoclonal N-terminal E1^E4 antibody. β-Actin demonstrates equal loading of protein samples. Deletion of the ^{14}LLSLL^{18} sequence of E1^E4 abrogates the formation of the proteolytically cleaved E4 protein. Obervations are representative of three independently performed experiments.

proteolysis is evident between the deletion mutants, across a number of experiments no significant trend became apparent. However deletion of the leucine-rich sequence: ^{14}LLSLL^{18}, prevented E1^E4 proteolysis, as demonstrated by an absence of the 10.5 kDa E4 polypeptide. A minor higher molecular mass E4 species was observed (Figure 5.8 B), migrating a little above the LLSLL deletion mutant, although in subsequent experiments this slower migrating E4 species was not observed, therefore the significance of this polypeptide is unknown. To establish whether residues N-terminal of this motif are required for proteolysis, E1^E4 residues ^{12}YP^{14} were deleted and the extent of proteolysis assessed, revealing that these residues are not required for N-terminal cleavage of the full-length protein (Figure 5.8 B).

5.7.2.2   The di-leucine sequence: ^{14}LL^{15} of HPV18 E1^E4 is required for N-terminal proteolysis

Due to the overlapping nature of the E2 and E4 ORFs, the leucine-rich motif of E4 cannot be deleted in the total HPV18 genome without disrupting the E2 coding sequence, and possibly the function of the E2 protein. However certain point mutations of E4 residues can be made which are silent within the E2 ORF. It is desirable that mutations are made to minimise disturbance to the overall structure of the protein or other functions mediated by this motif therefore further mutagenesis of the LLXLL motif was performed with substitution mutations to elucidate the requirement of each of the di-leucine pairs for E1^E4 cleavage.

In the context of the total HPV18 genome, the first di-leucine motif ^{14}LL^{15}, can be substituted by PP (Leucine 14+15 -> Proline, PPLSS) or RR (Leucine 14+15 -> Arginine, RRSLL) whilst the second di-leucine motif ^{17}LL^{18} may be mutated to SS (Leucine 17+18 -> Serine, PPSSS) without altering the ORF of the E2 gene (a schematic is given in Appendix I). To

establish whether these E1^E4 substitutions abrogate proteolysis, all the aforementioned di-peptide mutations were generated in the coE1^E4 expression vector and the contribution of each di-leucine pair to N-terminal proteolysis was assessed by Western blot analysis (Figure 5.9). Following transfection of coE1^E4 into 293T cells a proteolytically processed E4 species was detectable with a polyclonal E1^E4 antibody. However proteolysis was severely reduced by substitution of the second di-leucine residues within the LLSSS mutation, and completely abrogated with both RRSLL and PPSLL, indicating that whilst both di-leucine pairs are required for efficient proteolysis the first di-leucine pair is an absolute requirement.

5.8 Proteolytically-defective E1^E4 mutations retain the ability to associate with the keratin intermediate filament network

Substitutions within the leucine-rich motif of E1^E4 in the total HPV18 genome may impede the generation of N-terminally truncated E4 species. However they may also disturb association between E1^E4 and the cytokeratin network, even though this motif appears not to be critical for this association, since an N-terminal deletion mutant of HPV18 E1^E4 has been shown to not to colocalise to the keratin network in COS-1 cells (Nakahara et al., 2002 and Figure 5.7 B). Therefore to confirm the role of the LLXLL motif of HPV18 E1^E4, COS-1 cells were transfected with an E1^E4 deletion mutant lacking the [14]LLSLL[18] sequence and various substitution mutants of the di-leucine peptides of this motif (Figure 5.10). Deletion of the leucine-rich motif did not prevent association of E1^E4 with the keratin network, and a typical phenotype of juxta-nuclear ΔLLSLL E1^E4 protein bundles were frequently observed co-localising with keratin 18 (Figure 5.10). However, in some cells a more diffuse cytoplasmic staining pattern of E1^E4 ΔLLSLL was observed, marked with punctate spots with an absence of keratin co-localisation (Figure 5.10). Whilst this distribution has also been

A

E1^E4 ¹⁴LLSLL¹⁸

PPSLL ¹⁴PPSLL¹⁸

RRSLL ¹⁴RRSLL¹⁸

RRSSS ¹⁴RRSSS¹⁸

B

C
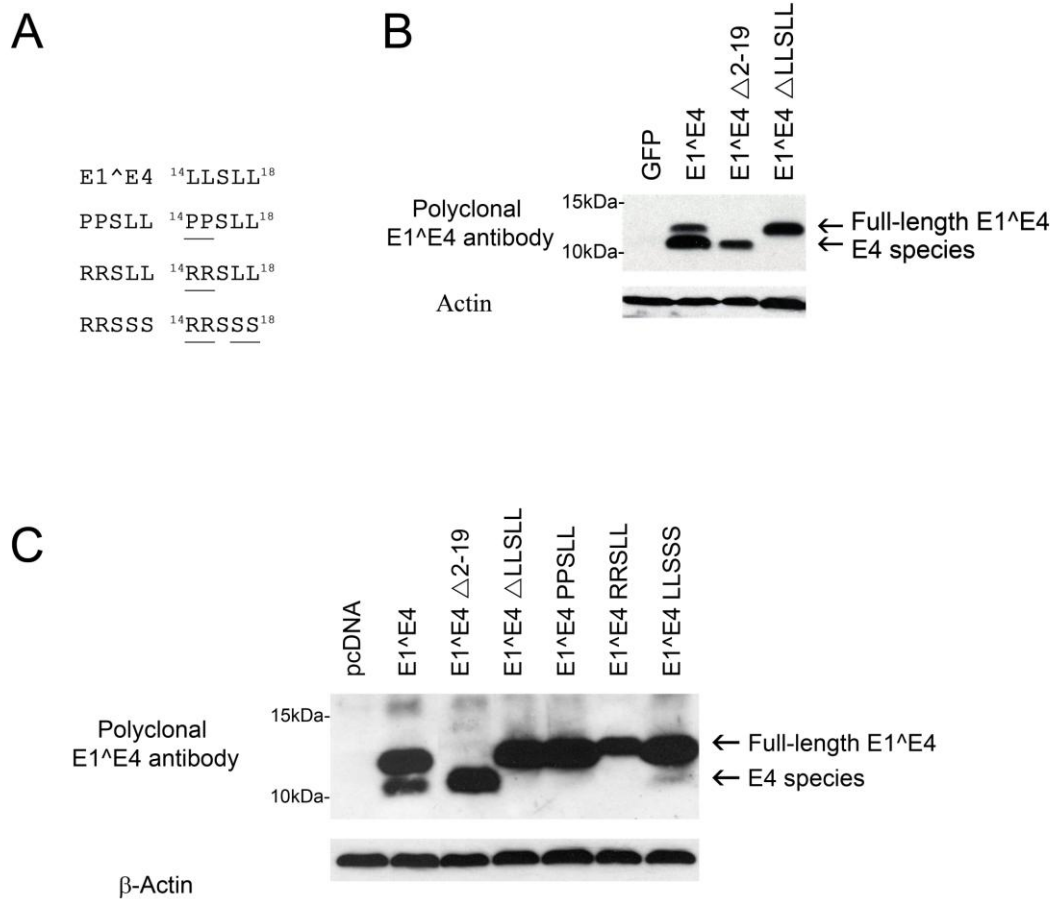
Figure 5.9: Identification of residues within the leucine-rich motif required for proteolysis of HPV18 E1^E4. **(A)** Depiction of HPV18 E1^E4 substitution mutants generated within the leucine-rich motif of the codon-optimised E1^E4 expression construct. Substituted residues are underlined. **(B-C)** Following 48 h transfection, cell extracts from 293T cells, expressing the indicated E1^E4 proteins, were prepared in urea lysis buffer and analysed by Western blot analysis with a polyclonal E1^E4 antibody. β-Actin demonstrate equal loading of protein samples. **(B)** Deletion of the ¹⁴LLSLL¹⁸ sequence of E1^E4 prevents the formation of processed E4 species. **(C)** Representative Western blot of three independently performed experiments. Substitution of the ¹⁴LL¹⁵ E1^E4 di-peptide by PP or RR completely abrogates N-terminal proteolysis whilst some proteolysis is observed upon substitution of ¹⁷LL¹⁸ by SS.
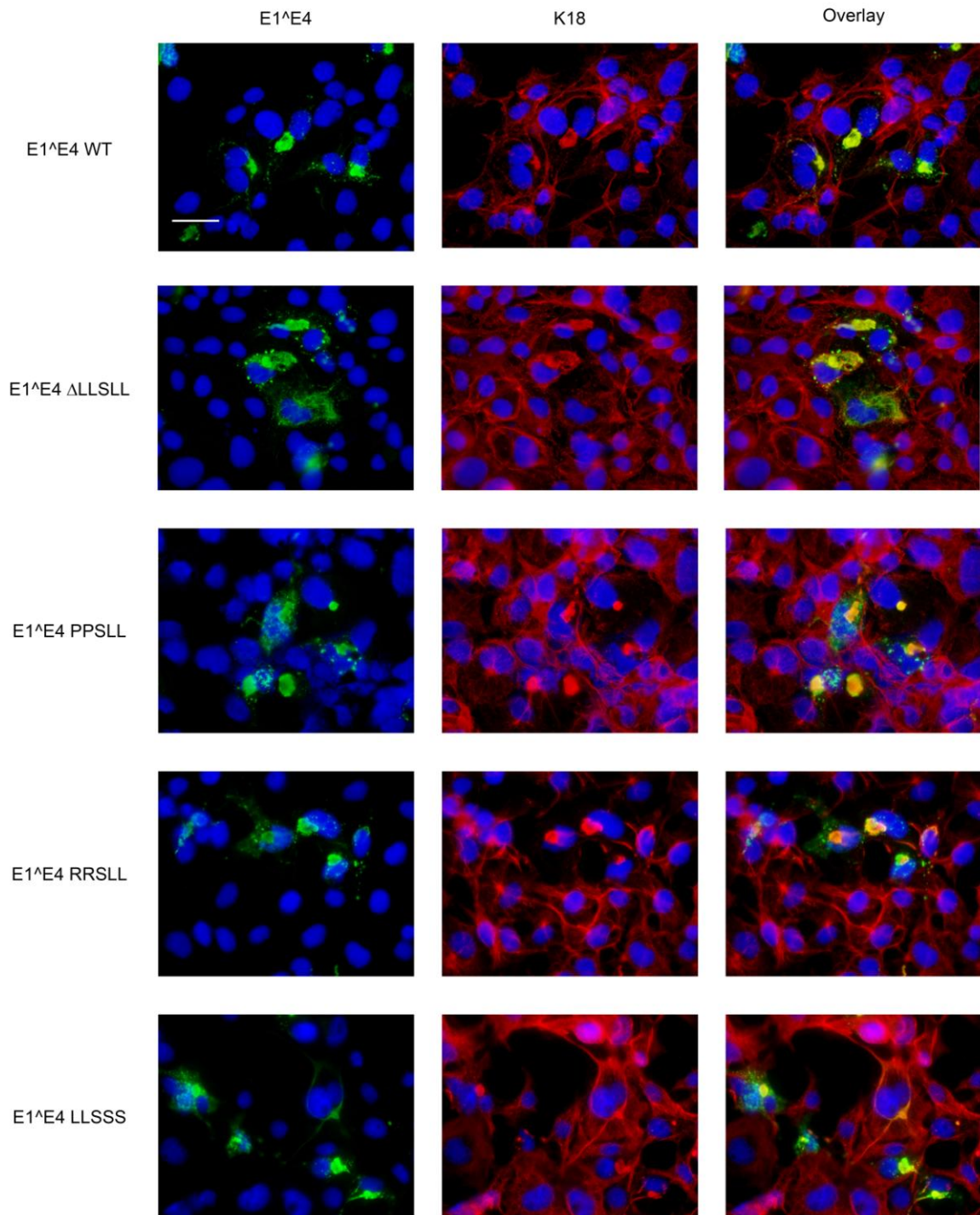
Figure 5.10: The leucine-rich motif is not essential for association of HPV18 E1^E4 with the keratin network. COS-1 cells were transfected for 48 h, followed by fixation in 4% paraformaldehyde and permeabilisation in acetone. Cells were dual-stained with a polyclonal E1^E4 anti-rabbit antibody and a keratin 18 anti-mouse antibody. Immune complexes were visualised with an anti-rabbit Alexa Fluor 488 antibody (green) and an anti-mouse Alexa Fluor 596 antibody (red) and counter-stained with Dapi for detection of nuclei (blue). Panels show cells transfected with: WT, ΔLLSLL PPSLL, RRSLL and LLSSS E1^E4. The leucine-rich motif mutant E1^E4 proteins co-localise with the keratin intermediate filament network, with collapse of these filamentous structures into fibrous bundles. Bar, 30 μm. Observations are representative of at least four independently performed experiments.

noted with wild type E1^E4, it is slightly more frequent with the ΔLLSLL E1^E4 deletion mutant, suggesting that removal of the leucine-rich motif may, to some extent, reduce the efficiency with which E1^E4 can associate with the keratin network.

The localisation of the di-leucine substitution mutants (PPSLL, RRSLL and LLSSS) was also examined, and substitution of the first di-leucine residues of the leucine-rich motif with PP or RR, or the latter residues with SS did not prevent association of these E1^E4 mutant proteins with the cytokeratin network (Figure 5.10). These mutant proteins are observed forming juxta-nuclear fibrous aggregates which co-localise with keratin, as observed with the wild type protein. The presence of bright cytoplasmic spots was often observed following collapse of the keratin network into fibrous clumps with similar frequency to the WT E1^E4 protein. Substitution of the di-leucine pairs of the leucine-rich motif of HPV18 E1^E4 does not appear to alter the cellular distribution of the protein, and these mutations do not affect the ability of the protein to associate with, and collapse the cytokeratin network.

5.9   HPV18 E1^E4 proteolysis is not dependent upon association with the keratin network

Given that the integrity of the leucine-rich motif is required for HPV18 E1^E4 proteolytic cleavage, but that removal of this motif does not prevent association with the keratin network, it is possible that E1^E4 proteolysis is dependent upon association with the keratin network. To assess whether E1^E4 can undergo proteolytic cleavage in the absence of a keratin network, coE1^E4 was expressed in SAOS-2 cells, an untransformed human cell line derived from an osteosarcoma, which lack keratin intermediate filaments (Figure 5.11). E1^E4 proteolysis was retained, despite the lack of keratin, as indicated by detection of smaller E4 species with a polyclonal E1^E4 antibody, demonstrating that E1^E4 proteolysis is
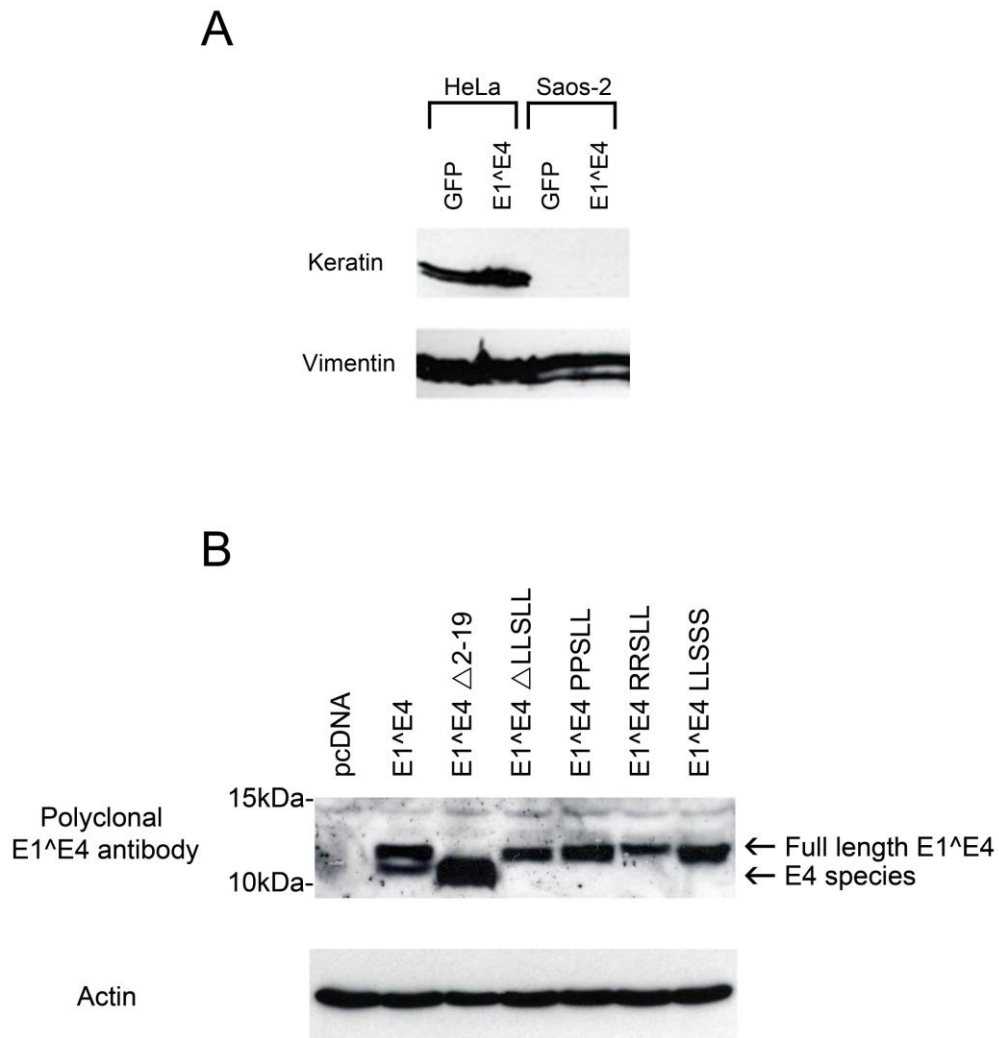
Figure 5.11: HPV18 E1^E4 undergoes proteolysis in the absence of a keratin network. **(A)** SAOS-2 cells lack keratin intermediate filaments. Western blot analysis of lysate derived from SAOS-2 and HeLa cells with a vimentin antibody and a keratin antibody. **(B)** Expression of E1^E4 and deletion mutants in SAOS-2 cells was examined by Western blot analysis with a polyclonal E1^E4 antibody. E1^E4 wild type protein undergoes N-terminal proteolysis within SAOS-2 cells to generate a processed E4 polypeptide, whilst this was abrogated by mutations of the leucine-rich motif. Association of E1^E4 with the cytokeratin network is therefore not required for N-terminal proteolysis. Representative Western blot of two independent experiments.

independent of the ability to associate with the keratin network. Proteolysis of the di-leucine substitution mutations of the first two leucines of E1^E4 (PPSLL and RRSLL) was abrogated demonstrating that these residues are required for proteolytic cleavage of HPV18 E1^E4 in a number of cell types, thus indicating that the protease targeting the protein for N-terminal cleavage is not restricted to epithelial cells.

Substitutions of the latter leucine pair (LLSSS) has previously been shown to undergo a small amount of proteolysis however, within SAOS-2 cells the processed E4 species was not observed, although this may be due to the lower level of protein expressed, preventing detection. Overall the LLSSS mutation severely reduces E1^E4 proteolysis but may not completely abrogate the production of N-terminally cleaved products. The PPSLL and RRSLL substitutions however consistently prevent production of proteolytically cleaved E4 species, detectable by Western blot analysis.

Discussion

The full-length HPV18 E1^E4 protein of 11.7 kDa has been shown to undergo N-terminal proteolysis to generate faster migrating E4 species of 10.5 kDa and 9.5 kDa. Whilst the 10.5 kDa processed E4 protein was routinely observed, the 9.5 kDa protein was only observed within a subset of experiments in which levels of E1^E4 expression are high. This observation suggests that proteolysis of the HPV18 E1^E4 protein occurs in a progressive manner as observed with the HPV1 E1^E4 protein (Breitburd et al., 1987) with initial proteolysis generating the 10.5 kDa E4 protein followed by further proteolysis to generate the 9.5 kDa E4 species. An HPV18 E1^E4 Δ2-19 deletion mutant, predicted to correspond to the 10.5 kDa processed E4 species was generated based on sequence comparisons of E1^E4 proteins and the known point of cleavage of the HPV1 E1^E4 protein (Roberts et al., 1994).

The Δ2-19 E1^E4 protein migrated as a protein of approximately 10.5 kDa therefore the N-terminus of this protein, beginning at serine residue 20 is likely to be close to that of the processed E4 species, thereby indicating that proteolysis is likely to occur within this region of the protein.

Mutational analysis identified residues of the leucine-rich sequence ([14]LLSLL[18]) of the protein as being essential for efficient proteolysis, with mutation of the [14]LL[15] di-peptide completely abrogating production of processed E4 species. The site of proteolysis of the HPV18 E1^E4 protein is therefore proposed to occur within this [14]LL[15] di-peptide or alternatively that mutation of this sequence alters the protein structure to prevent the protein undergoing proteolysis possibly by limiting access of the protease to the site of proteolysis or by preventing recognition of the protein as a substrate for the protease.

The leucine-rich motif is highly conserved between E1^E4 proteins and has been shown to mediate an association between HPV1 and 16 E1^E4 proteins and the keratin IF network (Roberts et al., 1997). However N-terminal residues of the HPV18 E1^E4 protein have previously been demonstrated to be dispensable for association with the keratin 8/18 network (Nakahara et al., 2002) and in this current study an E1^E4 Δ[14]LLSLL[18] deletion mutant was shown to retain the ability to associate with these intermediate filaments, therefore although highly conserved, the role of this motif in protein function may not be common to all E1^E4 proteins. However whilst the leucine-rich motif of the HPV18 E1^E4 protein may be dispensable for association with the K8/K18 network of simple epithelial cells, it may still play a role in association with more complex networks such as the K5/K14 network and differentiation-specific keratins present in HPV infected epithelium including K1/K10 and K4/K13.

N-terminal proteolysis of HPV18 E1^E4 has been shown to be independent of the ability of the protein to associate with the K8/K18 keratin network as demonstrated by the formation of the processed E4 species within cells lacking keratins.

5.10   Proteolysis of HPV18 E1^E4 may modulate protein function

A number of E1^E4 associations are encoded by N-terminal residues, including the ability to interact with PML, mitochondria and keratin (Roberts et al., 1994; Roberts et al., 2003; Raj et al., 2004; Knight et al., 2006; McIntosh et al., 2008). Proteolysis of HPV18 E1^E4 may act to regulate the function of the protein by removing N-terminal sequences which mediate protein-protein interactions and may result in a gain of function as with the HPV1 processed E4 protein (Knight et al., 2004). Whilst N-terminal residues are not essential for association of the HPV18 E1^E4 protein with the keratin network, an Δ2-19 deletion mutant protein displayed a distinct cellular distribution with a more diffuse cytoplasmic location, with reduced filament formation and with punctate and bright cytoplasmic foci often observed. This suggests that N-terminal residues outside of the leucine-rich motif may contribute towards efficient keratin association. In contrast, the full-length E1^E4 protein has a predominantly filamentous distribution, indicating that a loss of N-terminal residues by proteolysis will alter the subcellular localisation of the E4 protein and may act to modulate the function of the protein by enabling or preventing access to other proteins.

The leucine-rich motif may play a role in protein structure and removal of this region of the protein by proteolysis may alter the overall structure of the E1^E4 protein. In a paper published whilst the study of HPV18 E1^E4 proteolysis was ongoing, removal of N-terminal residues of the HPV16 E1^E4 protein was shown to alter the structure of protein and result in

assembly into amyloid-like filaments (McIntosh et al., 2008). Upon deletion of N-terminal residues, the HPV16 E4 protein was shown to adopt a ribbon like morphology which is characteristic of amyloid fibrils within an ordered aggregation. Probing of E4 fibrils with thioflavin T resulted in enhanced fluorescence of the dye, characteristic of binding to beta sheet structures of amyloid fibrils, indicating that N-terminally truncated, but not full-length E1^E4 assembles into amyloid-like fibrils. This study also demonstrated that the HPV16 E1^E4 protein undergoes N-terminal proteolysis to produce a single E4 species lacking N-terminal residues and it is hypothesised that alteration of protein structure in this manner enables accumulation of the protein within the cell. HPV18 E1^E4 is also detectable with amyloid probes indicating that this protein may display a similar structure to the HPV16 protein (McIntosh et al., 2008). Proteolysis of HPV18 E1^E4 may therefore serve to alter the structure of the protein which may prevent or facilitate associations with interacting proteins or alter the stability of the protein, and may thus regulate the activity of E1^E4.

The presence of E4 multimers within cells expressing the HPV18 E1^E4 protein suggests that the protein may function within a complex. Some of these multimers may correspond to homomultimers whilst some may be heteromultimeric complexes comprised of both full-length and truncated E4 proteins and indeed within this study the Δ2-19 E4 deletion mutant was has been shown to associate with the full-length E1^E4 protein of HPV18. It is possible that the complex formed between the full-length and N-terminal truncated E4 species may encode novel functions as has been shown for the HPV1 E4 heterocomplex (Knight et al., 2006). The full-length HPV1 E1^E4 protein and a protein corresponding to an N-terminally proteolytically cleaved E4 protein have been shown to cooperate to induce a G2 cell cycle arrest without coincident cellular replication (Knight et al., 2004; Knight et al., 2006).

5.11   **Is E1^E4 proteolysis important for E4 function during the HPV18 virus life cycle**

Proteolysis of HPV E1^E4 proteins appears to be a common post-translational modification shared between diverse HPV types (Doorbar et al., 1986; Doorbar et al., 1989; Tomita et al., 1991; Pray & Laimins, 1995; McIntosh et al., 2008), and it has now been shown that the HPV18 E1^E4 protein also undergoes N-terminal proteolysis. Proteolytic cleavage has been shown to modify the biological actions and cellular localisation of the full-length E1^E4 protein and proteolytic products encode novel functions (Knight et al., 2004; Knight et al., 2006). Proteolysis may therefore be an important viral mechanism to regulate E4 function during the HPV life cycle. To examine this hypothesis, mutations within the leucine-rich motif which abrogate proteolysis were introduced into the E4 ORF of the HPV18 genome. Transfection of human foreskin keratinocytes (HFK) with HPV18 genomes has previously been performed successfully within our laboratory to generate stable cell lines, which can be induced to differentiate to enable the virus replication cycle to be examined (Wilson et al., 2007). HFK cell lines containing HPV18 genomes with mutations of the E4 leucine-rich motif were subsequently analysed to examine the role of E1^E4 proteolysis during the HPV life cycle and this is discussed further within Chapter 6.

CHAPTER 6

THE ROLE OF N-TERMINAL PROTEOLYSIS OF

THE E1^E4 PROTEIN DURING THE HPV18

VIRUS LIFE CYCLE

Introduction

6.1   The development of expression systems for studying the papillomavirus life cycle

Whilst ectopic over-expression studies have enabled assignment of numerous functions to papillomavirus proteins, they may be inadequate for identification of the complete and genuine functional repertoire of these proteins. Examining the role of viral proteins during the HPV life cycle however, has proven to be difficult and had long been hindered by the lack of an available tissue culture system in which the virus can be propagated. The HPV life cycle is tightly linked to the differentiation programme of the host keratinocyte and difficulties have been encountered in propagation of HPV within monolayer tissue culture due to the dependency upon differentiating keratinocytes for completion of the virus life cycle. To understand the significance of viral proteins in the context of the HPV life cycle it is therefore desirable to utilise an expression system which permits completion of the viral replication cycle, in which viral proteins are expressed in a timely manner and at physiological levels which most closely resemble events of a true infection.

Early attempts focussed on the use of human xenografts, in which infected tissue is implanted into immunocompromised mice (Kreider et al., 1987). These systems produce lesions induced by HPV infection and enable propagation of papillomaviruses *in vitro*. Whilst papillomaviruses are host specific, the use of animal models to study papillomavirus infections have been widespread, notably so for the cottontail rabbit papillomavirus (CRPV) rabbit model. Following delivery of CRPV genomes into rabbit epithelium, papillomas quickly form which can progress to malignancy and this model has parallels with HPV

infections (Brandsma et al., 1991). This system can be used to elucidate the role of viral proteins through the use of mutant CRPV genomes.

Development of cell culture systems in which HPV replication can be supported has allowed the virus life cycle to be examined in tissue culture. Keratinocytes containing HPV genomes can be grown in organotypic raft cultures which induces cellular differentiation and recapitulates the stratified epithelium (McCance et al., 1988). Cells are grown on a collagen matrix with a feeder layer of fibroblasts, before raising of cells to the air-liquid interface which induces production of a fully stratified epithelium. Differentiation of keratinocytes using the organotypic raft culture system supports the complete HPV life cycle and enables production of virions in tissue culture (Meyers et al., 1992). Alternatively, suspension of cells in a semi-solid media containing methylcellulose has been shown to induce differentiation of epithelial cells and supports late viral functions including amplification of viral genomes and expression of late viral transcripts (Ruesch et al., 1998). Although this system does not reproduce the morphology of a stratified epithelium, cells differentiate within semi-solid media as a mostly homogeneous population, allowing discreet stages of the life cycle to be studied.

A number of biopsy derived cell lines have been established that stably maintain HPV genomes, including the HPV31 containing CIN612-9E cell line (Bedell et al., 1991) and the HPV16 W12 cells (Stanley et al., 1989). Whilst differentiation of cell lines harbouring HPV genomes has been effectively used to study the HPV life cycle within tissue culture, genetic analyses cannot be performed utilising these cells and the functional contribution of the viral proteins cannot be examined. Development of transfection systems in which mutant genomes can be analysed has furthered research into the HPV life cycle, thereby allowing manipulation

of viral genes by mutagenetic analysis. Initial studies proved to be unsuccessful, with integration of viral genomes into host DNA, however later studies allowed successful maintenance of viral episomes. Using normal primary human foreskin keratinocytes (HFK) derived from neonatal foreskin as a host cell, HPV genomes are co-transfected together with a drug resistance marker (Frattini et al., 1996). Following drug selection, the population of cells harbouring viral genomes is expanded, generating a cell line which stably maintains HPV genomes as extrachromosomal elements. These cell cultures can subsequently be grown as raft cultures or induced to differentiate following suspension in a semi-solid media composed of methylcellulose. This method has been subsequently used to analyse the life cycle of a number of HPV types to elucidate the role of viral proteins, including E1^E4.

Primary keratinocytes have a limited lifespan which is extended by the presence of high risk HPV genomes. Whilst these systems have been successfully utilised to generate HFK cell lines which stably maintain HPV11 genomes as episomes for extended passages (Thomas et al., 2001) attempts to study other low risk HPV types have proved to be difficult. Normal immortalised keratinocytes (NIKS), a spontaneously immortalised cell line, has been shown to support the virus life cycle of HPV16 within raft cultures (Flores et al., 1999). These cells can stably maintain transfected episomal genomes and can be used to analyse the life cycle of those papillomaviruses which do not extend the lifespan of the host cell. The use of N-Tert-1 cells, a cell line immortalised by the catalytic subunit of human telomerase, has proved useful in the study of the low risk HPV types, with raft cultures able to support the complete virus life cycle of HPV11 (Fang et al., 2006b).

6.2   Rationale for analysis of E1^E4 proteolysis in the context of the HPV18 life cycle

Mutational analysis of the coding sequence of the HPV18 E1^E4 protein during ectopic over-expression studies performed during Chapter 5, established that the presence of a ^{14}LL^{15} leucine di-peptide is require for N-terminal proteolysis of the full length E1^E4 protein within these cells. Substitution of the ^{14}LL^{15} di-peptide with PP or RR, abrogated the expression of faster migrating E4 species lacking N-terminal residues. It is hypothesised that the E1^E4 protein may undergo N-terminal proteolysis to form truncated E4 species during the HPV18 virus life cycle, in a manner similar to that observed during the over-expression studies. Given that E1^E4 proteins of several HPV types including HPV18, have been shown to contribute towards a number of functions during the virus life cycle, including viral genome amplification and late gene expression (Nakahara et al., 2005; Wilson et al., 2005; Wilson et al., 2007), it is proposed that proteolysis of E1^E4 may be necessary for, or may contribute towards the protein's ability to perform these functions.

To address the hypothesis that HPV18 E1^E4 undergoes N-terminal proteolysis during the virus life cycle and to establish whether this is dependent upon the ^{14}LL^{15} di-peptide the primary keratinocyte cell culture system was utilised. This system was chosen since it has previously been applied successfully within our laboratory to generate HFK cell lines containing HPV18 genomes, which upon cellular differentiation are able to recapitulate the HPV18 life cycle (Wilson et al., 2007). Loss of expression of the full length protein was shown to induce defects in cell growth in basal-like cells and following differentiation, defects in viral genome amplification and late gene expression. Generation of a HFK cell line harbouring HPV18 genomes in which the E1^E4 protein is unable to undergo proteolysis

would allow the contribution of E1^E4 proteolytic cleavage to viral genome amplification and late gene expression to be tested.

Results

6.3   Generation of HFK cell lines containing HPV18 genomes with mutations within the E4 sequence that are necessary for proteolysis of the E1^E4 protein *in vitro*

With the aim of assessing the role of E1^E4 proteolysis during the HPV18 life cycle, HFK cell lines were established, containing HPV18 genomes with point substitutions within the E4 leucine-rich motif (Figure 6.1, A). To allow maximal assessment of the contribution of the di-leucine sequence ^{14}LL^{15} to N-terminal proteolysis of E1^E4, both the ^{14}PP^{15} (PPSLL) and ^{14}RR^{15} (RRSLL) mutations were introduced into the total HPV18 genome. If a phenotype were to be observed with both of these mutations, rather than one but not the other, this could suggest that this is due to an inability of the E1^E4 protein to undergo proteolysis, and not due to a structural change introduced by the substituted residues. It was decided that due to the highly conserved nature of the leucine-rich motif between HPV E1^E4 proteins, assessment of the contribution of the complete ^{14}LLSLL^{18} sequence during the HPV life cycle would be valuable to understanding E1^E4 function, therefore both leucine di-peptides (^{14}RRSSS^{18}) were substituted in the HPV18 genome.
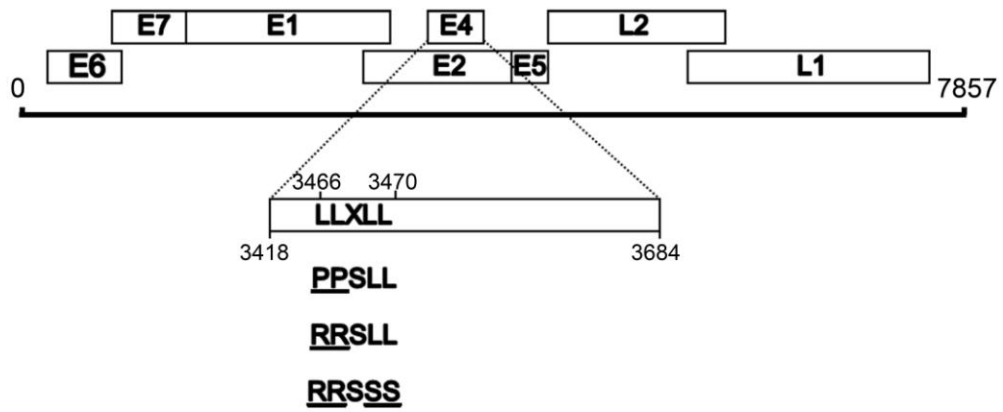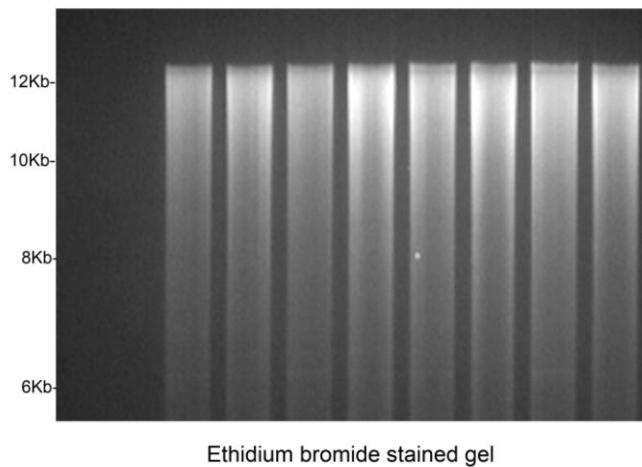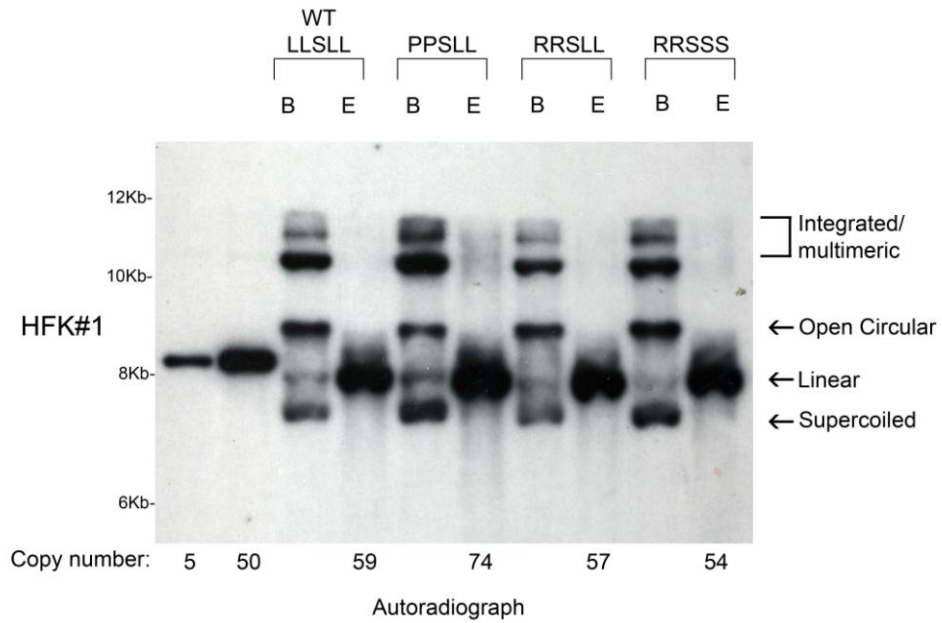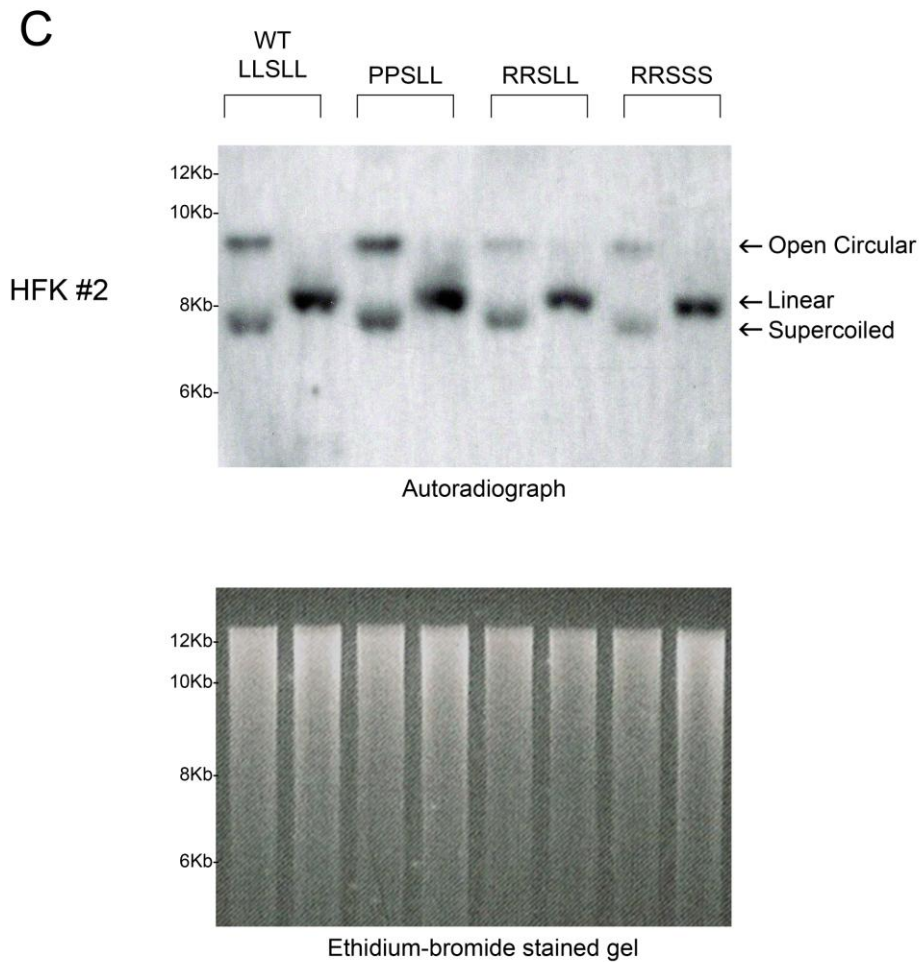
Figure 6.1 HPV18 genomes containing substitution mutations of the leucine-rich motif of the E4 ORF are replicated following transfection into human foreskin keratinocytes (HFK). **(A)** PPSLL, RRSLL and RRSSS substitution mutations were inserted into the LLSLL sequence of the E4 ORF, between nucleotides 3466 and 3470 of the HPV18 genome. These changes are silent within the E2 coding sequence (adapted from (Wilson et al. 2007)). **(B,C)** Total DNA isolated from HFK cell line #1 (B) and #2 (C) transfected with wildtype (WT), or E1^E4 substitution mutant genomes was digested with either *Eco*RI (E) a single cutter, or with *Bgl*II (B) a non-cutter, of the HPV18 genome. DNA was also digested with *Dpn*I to remove residual input DNA and Southern blot analysis was performed with HPV18 genomic DNA as a probe. HPV18 genomes migrate as linear, open-circular and super-coiled episomal DNA. Copy numbers were calculated from phosphoimaging analysis of the autoradiographs compared with a copy number control and are indicated below the autoradiograph. Equal loading is demonstrated by staining of the agarose gel with ethidium bromide.

6.3.1 Introduction of mutations into the E4 leucine-rich sequence of HPV18 genomes

Mutations of the leucine-rich motif [14]LLSLL[18], PPSLL, RRSLL and RRSSS, were introduced into E4 ORF of the pGEMII-HPV18 vector by site-directed mutagenesis. To confirm that only the intended mutations had been inserted, sequencing of the complete genomes of each mutant was performed utilising a panel of primers designed to cover the entire genome. Genomes were subsequently prepared for transfection into HFKs by digestion and re-ligation to release the HPV18 genome from the pGEMII vector backbone before re-circularisation (Chapter 2: Materials and methods, 2.7.5.2).

6.3.2 Transfection of HFKs with HPV18 genomes

HFKs derived from neonatal foreskin of two donors, denoted HFK#1 and HFK#2, were co-transfected with HPV18 wild type or E4 mutant genomes (PPSLL, RRSLL and RRSSS) and a plasmid containing a neomycin resistance gene. Following transfection, cells were grown on a feeder layer of irradiated J2-3T3 fibroblasts and a period of selection with the G418 selection antibiotic ensued, whereby only cells transfected with the neomycin resistance gene are provided with resistance. Surviving clones following selection were pooled and the cells expanded to produce a stable cell line. Sequencing of extracted DNA from these cell lines with a primer covering the E4 ORF confirmed that the cell lines contained the correct HPV18 genomes.

6.4 **Substitutions within the leucine-rich motif of E4 do not prevent viral episomal replication.**

To confirm that the transfected HFK cell lines had established episomal copies of HPV18 genomes, DNA was extracted and analysed by Southern blot analysis using a HPV18 genomic probe (Figure 6.1, B-C). Equal amounts of DNA derived from wild type or E4 mutant cell lines were digested with *Eco*RI or with *Bgl*II. The HPV18 genome does not contain any restriction sites for *Bgl*II and therefore epsiomal genomes remain intact and could be observed migrating as both supercoiled (SC) and open circular (OC) plasmid DNA in both wild type and E4 mutant cell lines. A degree of integrated and multimeric genomes could also be observed within all cell lines. A single restriction site for *Eco*RI is contained within the HPV18 genome and therefore digestion with this enzyme allowed analysis of linear forms of genomic DNA, with no substantial differences in copy number observed between wild type and E4 mutant cell lines. Ethidium bromide stained gels were used to provide an indication of DNA loading in all Southern blots. However it must be acknowledged that probing Southern blots for a housekeeping gene such as GAPDH would provide a more accurate method of determining equal loading of DNA samples.

Transfection of HPV18 genomes containing mutations within the leucine-rich motif of the E4 ORF was repeated in a second cell line (HFK#2) revealing similar results within a different donor background (Figure 6.1, C). Given that a previous study has shown that full-length E1^E4 is not required for establishment of episomal HPV18 genomes (Wilson et al., 2007), substitutions of the leucine-rich motif would not be expected to affect these viral functions in the context of this virus type unless these mutations were gain of function mutations.
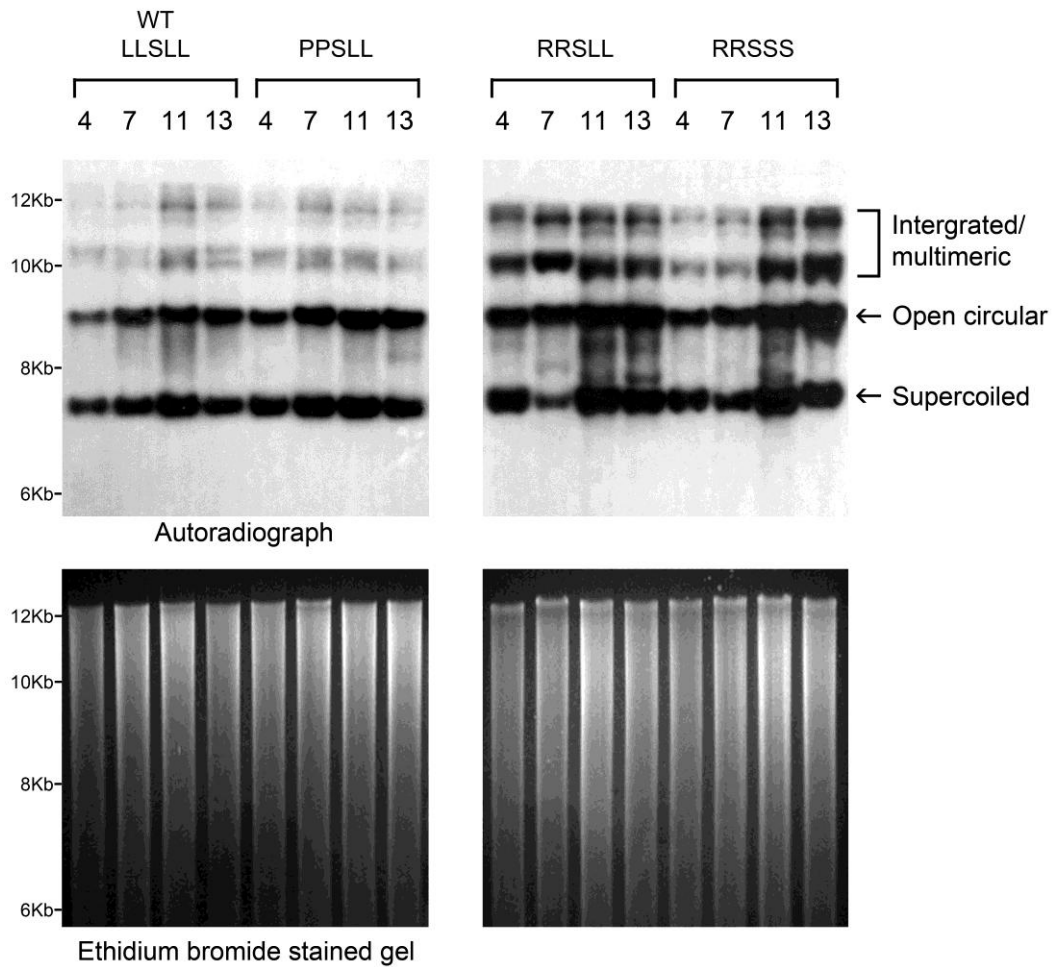
Figure 6.2: HPV18 genomes containing mutations of the leucine-rich motif within the E4 ORF are stably maintained as episomes upon repeated cell passaging. Total DNA from HFKs harbouring: wildtype, PPSLL, RRSLL and RRSSS E1^E4 mutant HPV18 genomes, was extracted at various passages, digested with BglII, a non-cutter of the HPV18 genome and examined by Southern blot analysis with an HPV18 genomic probe. Equal loading is demonstrated by staining of the gel with Ethidium bromide.

To confirm that HPV18 viral episomes are maintained upon extended passaging of cells, DNA was extracted from cells at passages 4, 7, 11 and 13, and subjected to Southern blot analysis (Figure 6.2). Whilst some fluctuations where observed between passages, HPV18 genomes with mutations of the E4 leucine-rich motif were maintained at similar levels to wild type genomes following cellular passaging.

6.5   Mutation of the leucine-rich motif of E4 does not alter the growth of HFKs harbouring HPV18 genomes

Introduction of a stop codon into the E4 ORF (E4M17) of the HPV18 genome, resulting in predicted expression of a truncated E1^E4 protein of 16 aa has previously been shown to enhance the growth of undifferentiated HFKs compared with those cells harbouring wild type HPV18 genomes (Wilson et al., 2007). The altered growth rate of HFKs containing E4M17 genomes suggests a function for the E1^E4 protein within undifferentiated keratinocytes and perhaps proteolytically cleaved HPV18 E4 proteins, may be produced within these cells at low levels and contribute towards early E4 functions.

To assess the contribution of the leucine-rich motif towards the E1^E4 function of growth suppression, growth assays were performed with HFKs containing wild type and E4 mutant genomes. The growth rate of HFK cells was significantly enhanced as expected, following transfection of HPV18 genomes, with un-transfected HFKs growing at a significantly lower rate in comparison with HFKs harbouring wild type HPV18 genomes (Figure 6.3 A). Growth profiles of HFKs containing HPV18 mutant genomes, in which the: PPSLL, RRSLL and RRSSS substitutions had been introduced, were assessed in both HFK #1 (Figure 6.3 B) and HFK #2 (Figure 6.3 C) donor backgrounds. Cellular proliferation of all three E1^E4 mutant cell lines, was comparable to the rates of cell growth of HFKs harbouring wild type genomes
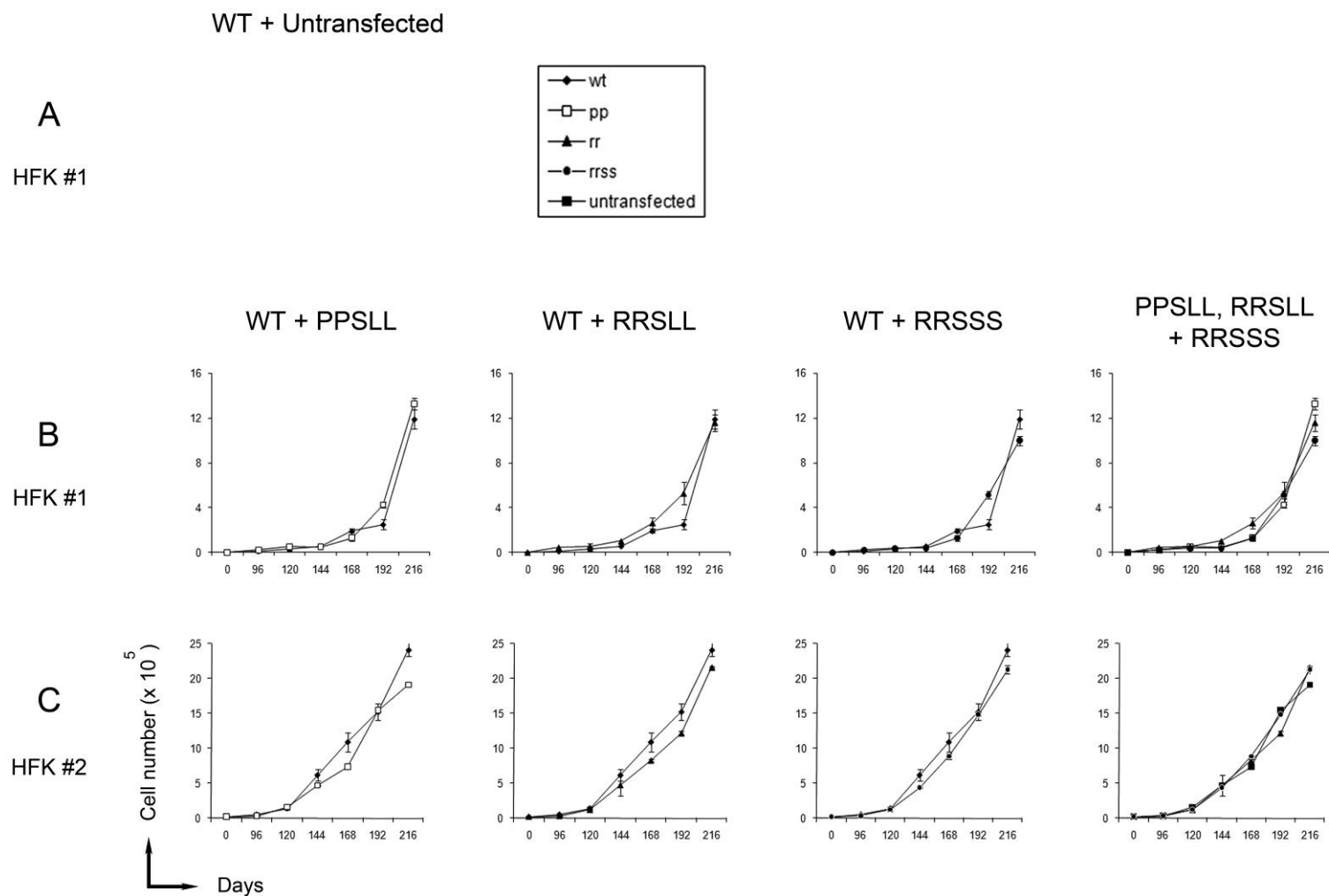
Figure 6.3: HPV18 genomes with mutations of the leucine-rich motif within the E4 ORF show growth rates comparable with wildtype cells. Cellular proliferation profiles of HFKs harbouring: wildtype (♦), PPSLL (□), RRSLL (▲) and RRSSS (●) E1^E4 mutant HPV18 genomes or untransfected (■) cells. Cell growth profiles were performed in duplicate with two donor cell lines: HFK#1 and #2 and error bars represent the standard deviation of the experiments shown. Growth profiles are representative of three sets of independent experiments.

within both backgrounds and indicates that the leucine-rich motif does not contribute towards growth suppression.

6.6   Expression of E1^E4 is reduced upon differentiation in semi-solid media of HFKs with HPV genomes containing mutations of the leucine-rich motif

Whilst N-terminal proteolysis of the HPV18 E1^E4 protein has been shown to occur during over-expression studies within immortalised and transformed cell lines (Chapter 5), it was important to establish whether E1^E4 proteolysis occurs during the virus life cycle and whether this is dependent upon residues within the leucine-rich motif of the protein.

HFK cell lines #1 harbouring HPV18 wild type and mutant genomes were induced to differentiate by suspension of cells in semi-solid media. HFKs grown as monolayer cultures, were transferred to media containing 1.5% methylcellulose prior to harvesting following 24 and 48 h growth in methylcellulose. Protein lysates were extracted, and Western blot analysis performed with keratinocyte terminal differentiation markers to confirm successful cellular differentiation. Expression of involucrin and keratin 1 is restricted to the suprabasal layers of the stratified epithelium and the detection of these proteins at comparable levels within wild type and E1^E4 mutant cell lines, following suspension in semi-solid media, confirmed that cellular differentiation had been achieved to a similar extent (Figure 6.4). However, whilst cellular differentiation levels were comparable, and equal protein loading was confirmed with an anti-β-actin antibody, the levels of E1^E4 detectable with a polyclonal E1^E4 antibody varied greatly between wild type and the E4 mutant cell lines. In HFKs harbouring wild type genomes, high level expression of the E1^E4 protein was observed, and by 48 h post-
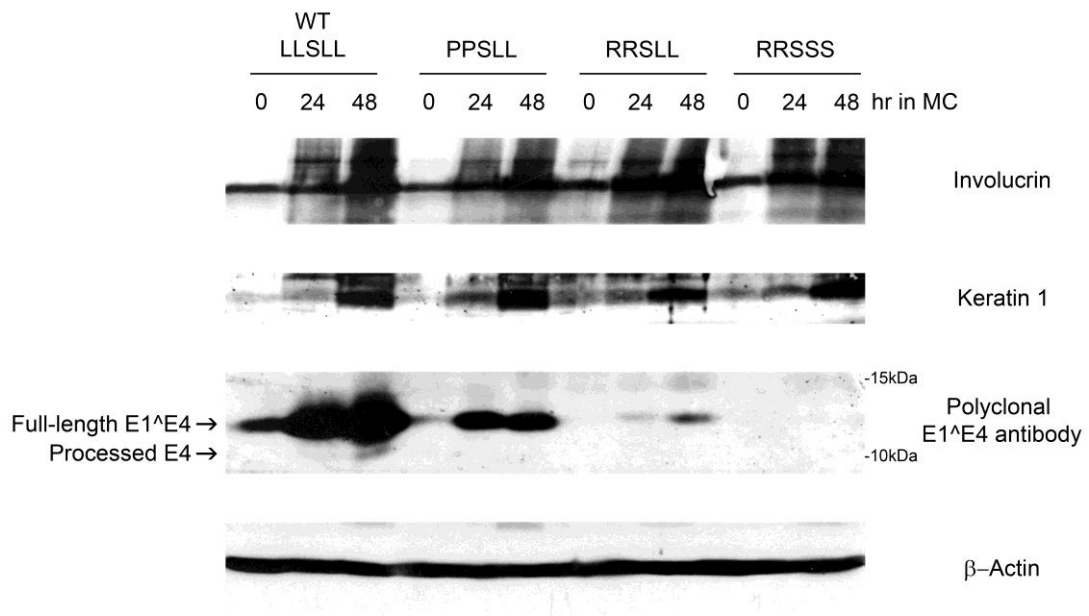
Figure 6.4: Differentiation-dependent expression of the E1^E4 protein is reduced in HFKs containing HPV18 genomes with mutations of the E1^E4 leucine-rich motif. Following suspension of HFKs containing: wildtype (WT), PPSLL, RRSLL and RRSSS E1^E4 mutant HPV18 genomes in semi-solid methylcellulose (MC) media for various times, cells were harvested and lysate analysed by Western blot analysis with a polyclonal E1^E4 antibody and antibodies to involucrin and keratin 1 differentiation markers. β-Actin demonstrate equal loading of protein samples and Western blot is representative of three independently performed experiments.

induction a faster migrating E4 species of 10.5 kDa was detected, indicating that the wild type E1^E4 protein undergoes proteolysis within differentiating keratinocytes.

E1^E4 expression within HFKs harbouring the RRSSS mutant genomes was virtually undetectable after 48 h suspension in semi-solid media, and whilst levels of E1^E4 RRSLL and PPSLL proteins were greater, the levels were well below that observed within wild type cells. Faster migrating E4 proteins could not be detected in any of the E4 mutants containing substitutions of the leucine-rich motif although lack of detection may be due to the low levels of the full-length E1^E4 protein expressed within these cells. These observations were also observed with HFK #2 (data not shown).

The levels of the PPSLL, RRSLL and RRSSS E1^E4 mutant proteins observed within over-expression studies in 293T and SAOS-2 cells were comparable to wild type E1^E4 levels indicating that within these studies protein stability was not affected. The reduced levels of these mutant E1^E4 proteins within differentiated HFKs may be due to a reduction in expression of E1^E4 containing transcripts. Upon cellular differentiation, induction of E1^E4 containing transcripts is activated from a differentiation-dependent late promoter within the E7 ORF from which late genes including the E1^E4 protein are transcribed. Decreases in viral genome amplification leading to reduced production of late viral transcripts from which the E1^E4 protein is translated could be another explanation of the reduced levels of E1^E4 protein. To address this hypothesis the ability of HFK cell lines containing the different mutant HPV18 genomes to undergo viral genome amplification was examined.

6.7   The effect of substitutions within E4 on HPV18 genome amplification

HFK #1 cell lines containing HPV18 genomes with the different mutations within the leucine-rich sequence were induced to differentiate by suspension of cells in 1.5% methylcellulose semi-solid media. Included in the experiments was the E417M mutant genome that is unable to express full length E1^E4 and does not amplify efficiently following differentiation (Wilson et al., 2007). Cells were harvested at 24 and 48 h and viral DNA extracted and analysed by Southern blot analysis (Figure 6.5). By 48 h, wild type viral genomes had amplified by approximately 2 fold, comparable with prior studies (Wilson et al., 2007). Whilst amplification of E417M viral genomes was severely impaired, consistent with previous studies, HPV18 genomes containing the E1^E4 substitutions of the first leucine di-peptide (PPSLL and RRSLL), amplified to levels comparable with wild type genomes indicating that the ^{14}LL^{15} di-peptide of E1^E4 is not required for the role of the protein in viral genome amplification. However, upon differentiation of HFKs harbouring genomes containing the RRSSS mutation, viral genome amplification was severely reduced compared to wild type genomes and suggests that an intact leucine-rich motif of HPV18 E1^E4 is required for efficient viral genome amplification and that the ^{17}LL^{18} di-peptide of E1^E4 may play an important role in this viral function.

This experiment was repeated three times using HFK #1 cell lines and a Student T-test was performed on data from the 48 hr time points (Figure 6.5 B). Fold amplification at the 48 hr time point of the WT, PPSLL and RRSLL mutations were significantly different (at the 90% confidence interval) from that of the E4M17 mutation, whilst the RRSSS mutation was not found to be significantly different and is therefore comparable with the E4M17 (Figure 6.5 B, b). This experiment was repeated once in HFK #2 with amplification of the WT genome observed, whilst the RRSSS mutant genome failed to amplify (Figure 6.5 C).

A

Autoradiograph

Ethidium bromide stained gel
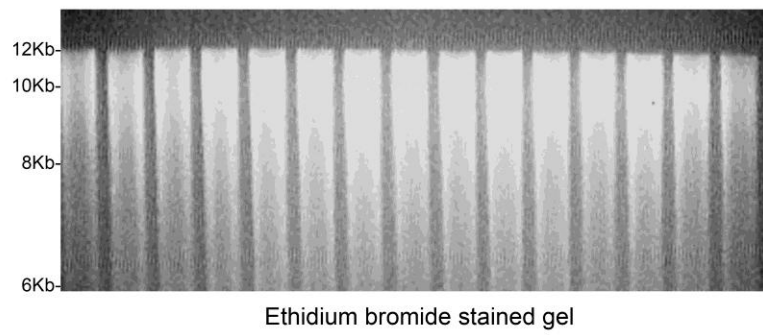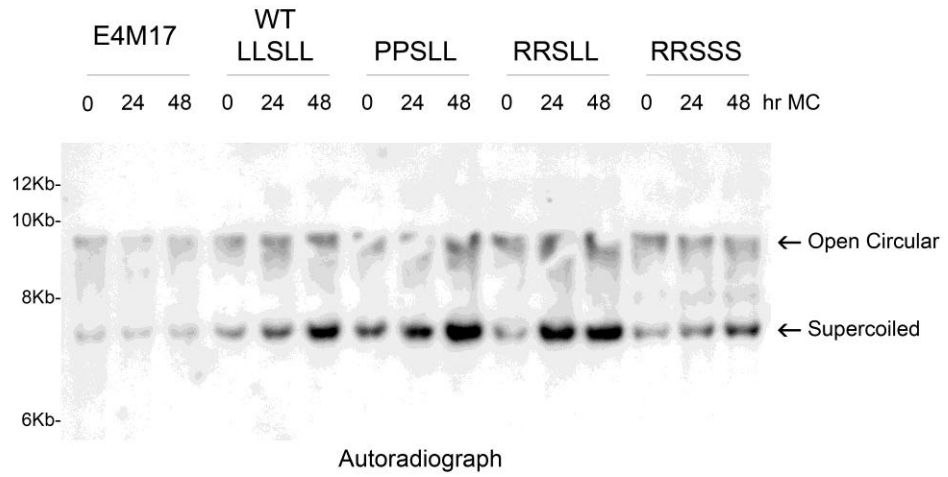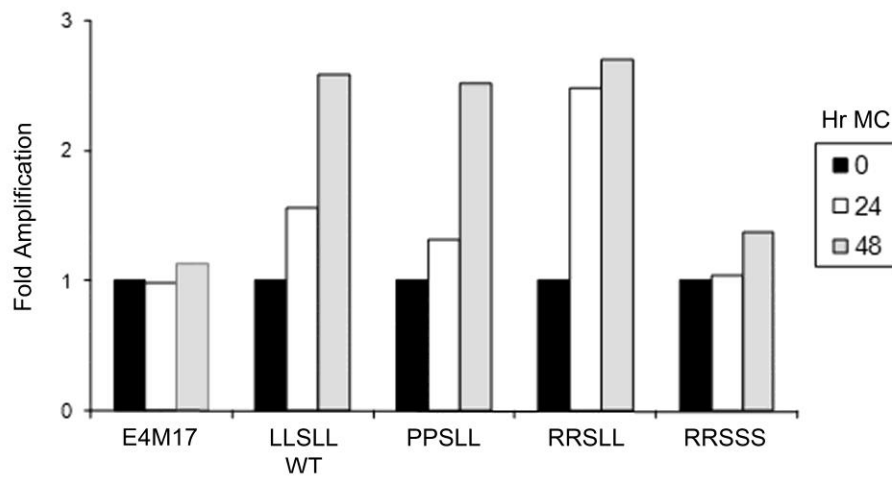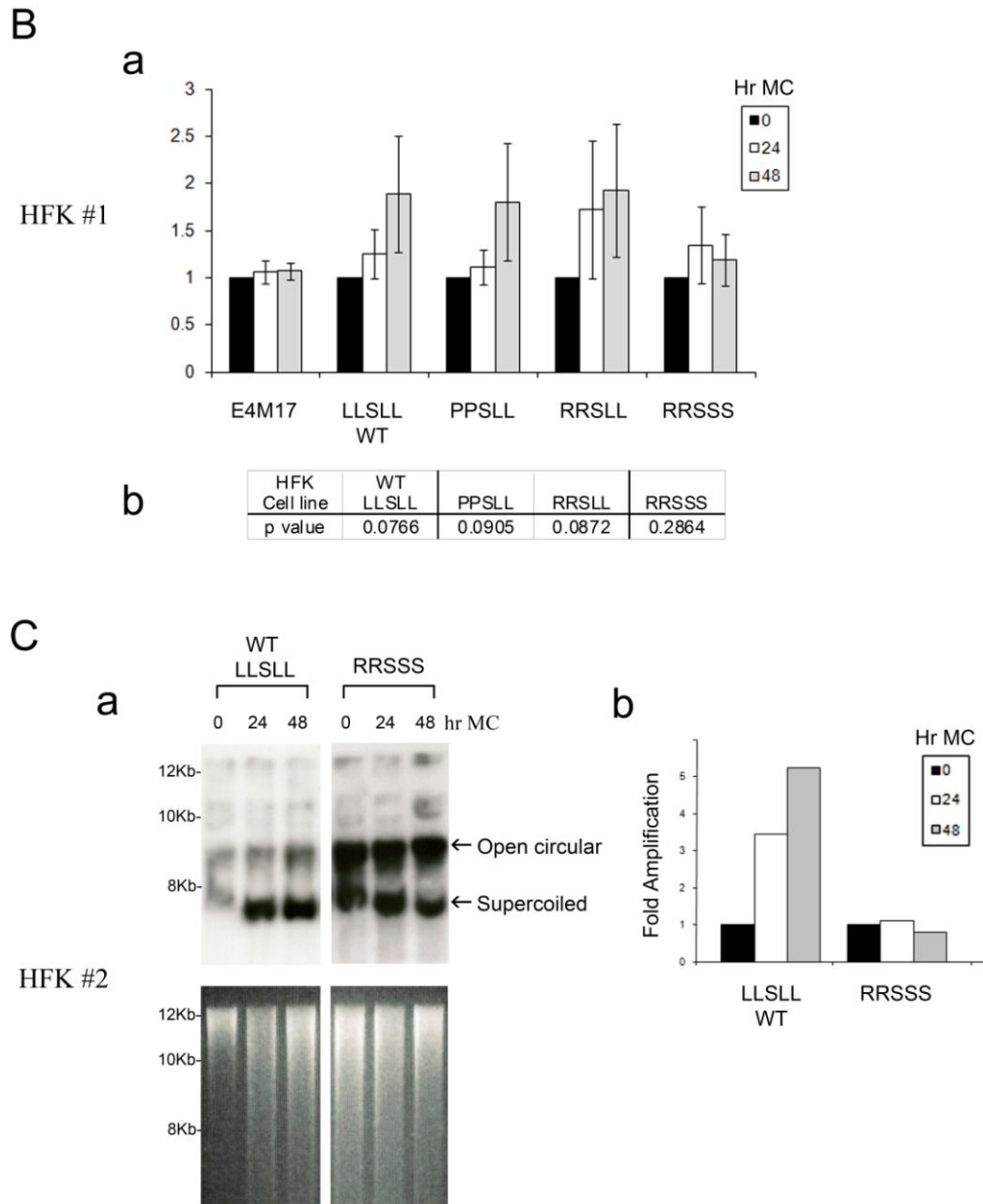
Figure 6.5: Differentiation-dependent viral genome amplification of HPV18 genomes is reduced upon mutation of an E4 leucine-rich motif. **(A)** (a) Total DNA was extracted at various times following cellular suspension in semi-solid methylcellulose (MC) media of HFK #1 containing HPV18 genomes with: E4M17, wildtype (WT), PPSLL, RRSLL and RRSSS mutations of the E1^E4 leucine-rich motif: LLSLL. Equal amounts of DNA from each cell line were digested with *Bgl*II, a non cutter of the HPV18 genome, and subjected to Southern blot analysis with an HPV18 genomic probe. Equal loading is demonstrated by staining of the gel with ethidium bromide. (b) Bar graph shows fold amplification of viral genomes upon HFK differentiation. Data was derived from phosphorimaging data of (a) and was normalised to that of monolayer cells at T = 0 for each cell line. **(B)** (a) Bar graph showing data derived from three independent experiments (as performed in A) using matched sets of wildtype and E1^E4 mutant HPV18 genomes from HFK #1 cell line. (b) P values from a one-sided student T-test performed on data from the 48 hr time point in (a). Amplification of WT, PPSLL and RRSLL genomes at 48 hr were significantly different (at the 90% confidence interval) compared with the E4M17 genome. RRSSS does not differ significantly from E4M17 at 48 hr. **(C)** (a) Southern blot analysis of differentiation-dependent viral genome amplification of wildtype and RRSSS mutant HPV18 genomes in HFK #2 cell line. (b) Bar graph showing data derived from phosphoimaging data of (a).

The impaired ability of HPV18 genomes containing the RRSSS mutation of the leucine-rich motif of the E4 ORF to undergo viral genome amplification, is likely to account for the observed reduction in differentiation-dependent expression of the E1^E4 protein. However the retained ability of HPV18 genomes with PPSLL and RRSLL mutations to undergo efficient viral genome amplification, accompanied by reduced levels in differentiation-dependent expression of these E1^E4 proteins, indicates that reduced E1^E4 levels are unlikely to be the result of fewer genome templates.

Whilst late gene transcription and viral genome amplification coincide with the onset of cellular differentiation, differentiation-dependent production of late viral transcripts is not entirely dependent upon amplification of viral genomes (Spink & Laimins, 2005). E1^E4 proteolysis may therefore be required for activation of late gene transcription or alternatively introduction of these mutations may affect the stability of E1^E4 containing transcripts. The low levels of E1^E4 mutant proteins observed upon differentiation may be the result of impaired E1^E4 transcription.

6.8 The role of the leucine-rich sequence of E4 in the production of HPV18 late gene transcripts

The various cell lines were induced to differentiate by suspension in semi-solid media and cells were harvested after 24 and 48 h. RNA was extracted and analysed by Northern blotting with an HPV18 probe and a Northern blot representative of three experiments is shown in Figure 6.6. Late viral transcripts encoding for the E1^E4, E5 transcript were observed upon cellular differentiation of HFKs containing wild type HPV18 genomes and were also observed within cell lines containing the leucine-rich E4 ORF mutations: PPSLL, RRSLL and RRSSS, indicating that the leucine-rich motif is not required for production of these late viral
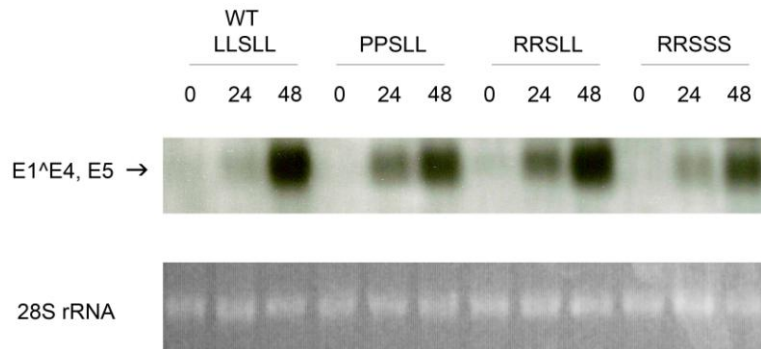
Figure 6.6: Human foreskin keratinocytes (HFK) harbouring HPV18 genomes with substitution mutations of the E1^E4 leucine-rich motif produce E1^E4, E5 late viral transcripts upon cellular differentiation in semi-solid media. Northern blot analysis of total RNA extracted from HFKs containing: wildtype (WT), PPSLL, RRSLL and RRSSS E1^E4 mutant HPV18 genomes, following cellular suspension in semi-solid methylcellulose media for various times. The E1^E4, E5 late viral transcript is indicated and 28S rRNA used as a loading control.

transcripts. However, the level of late viral transcripts observed within HFKs harbouring the RRSSS mutation was slightly lower than in those with the wild type HPV18 genome and this was consistent for 3 experiments performed with HFK cell line #1. The reduction in late viral transcripts is likely to be due to a reduced number of template genomes as a result of impaired viral genome amplification of HPV18 genomes. A similar analysis of late viral gene expression was not performed in HFK #2 due to time constraints.

6.9   An intact leucine-rich motif is required for efficient proteolysis of the E1^E4 protein within HFK organotypic raft cultures

Whilst cellular differentiation using semi-solid media is a useful tool for studying differentiation-dependent stages of the virus life cycle, including viral genome amplification and late gene expression (Ruesch et al., 1998), the complete HPV life cycle is not supported within this system. Organotypic raft culture however recapitulates the spatial separation of differentiated cells within a stratified epithelial and enables completion of the HPV life cycle, resulting in the production of virions (Meyers et al., 1997). The accumulation of E1^E4 and the cleaved E4 species may be enhanced at later stages during the life cycle and the processed protein may encode functions necessary for late events.

To explore this possibility organotypic raft cultures of HFKs containing HPV18 wild type genomes and genomes with mutations of the leucine-rich motif of the E4 ORF were kindly grown on my behalf by Dr Sally Roberts. The raft cultures were harvested in raft lysis buffer and solubilised in Laemelli buffer, then equal amounts of the lysate were analysed by Western blot analysis, performed using a polyclonal E1^E4 antibody or an E1^E4 N-terminal specific MAb, with a GAPDH antibody to ensure equal protein loading (Figure 6.7 A). Within organotypic raft cultures derived from HFKs harbouring HPV18 wild type genomes, the

A

Poly-clonal E1^E4

GAPDH

MAb E1^E4

GAPDH

← Full-length E1^E4
← Processed E4

← Full-length E1^E4

B

Processed E4 protein levels (% of total E4)
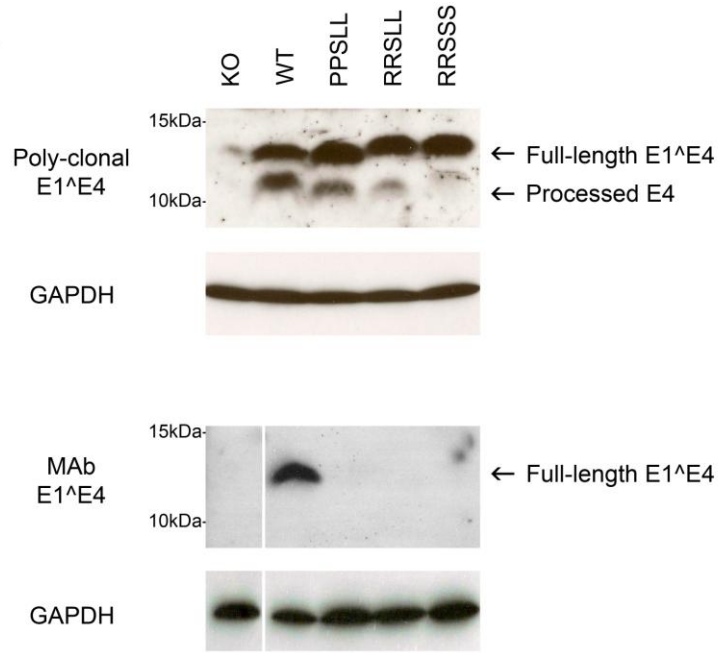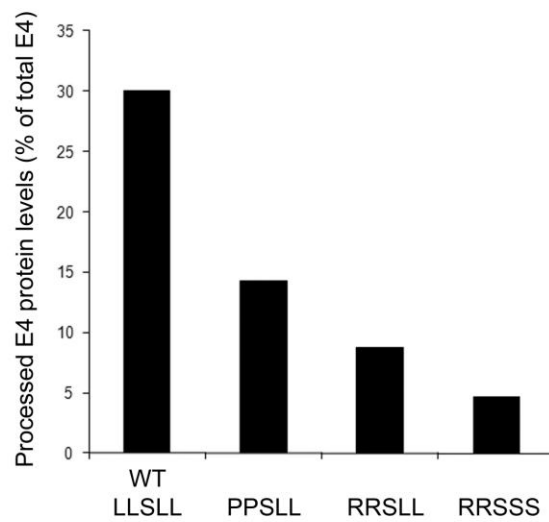
C
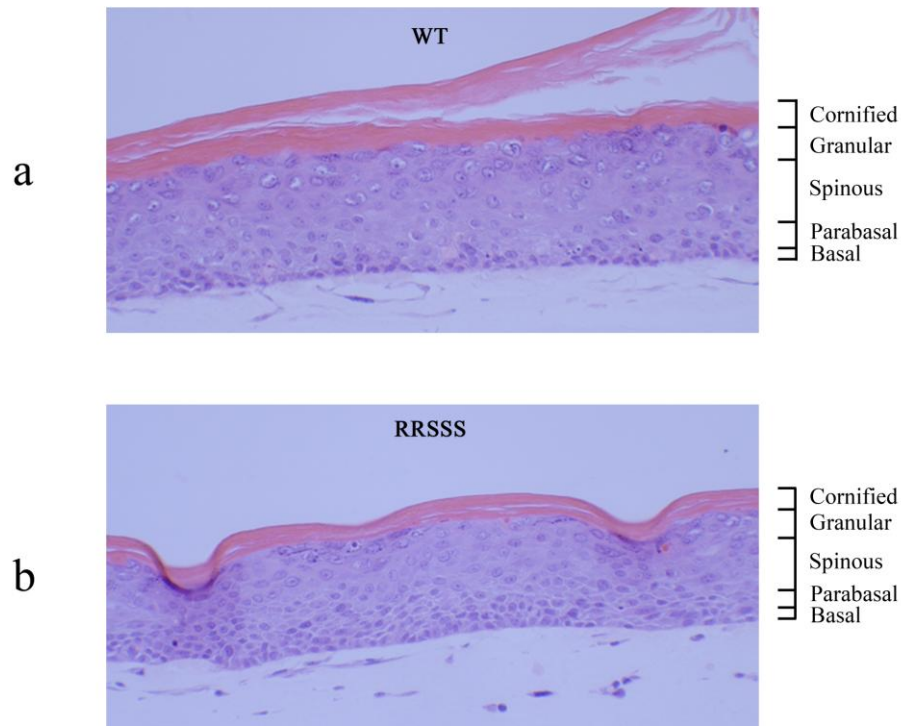
Figure 6.7: Production of an N-terminally processed E4 species upon cellular differentiation in organotypic raft culture is reduced in HFKs harbouring HPV18 genomes with mutations of the E1^E4 LLSLL sequence. **(A)** Following differentiation of HFKs harbouring: E4M17, wildtype (WT), PPSLL, RRSLL and RRSSS E1^E4 mutant HPV18 genomes in organotypic rafts culture, cellular lysate was subjected to Western blot analysis with a polyclonal E1^E4 antibody and a N-terminal specific monoclonal E1^E4 antibody. GAPDH demonstrates equal loading of protein samples. **(B)** Following densitometry scanning of the Western blot (A), the level of processed E4 protein was calculated as a percentage of the total E4 protein level for each cell line. **(C)** Tissue cross sections of organotypic rafts of HFKs containing (a) WT, and (b) RRSSS mutant, HPV18 genomes were stained with hematoxylin and eosin. Layers of the stratified epithelium are indicated.

major E1^E4 protein of 11.7 kDa was observed, and in addition, a faster migrating E4 species was detected of approximately 10.5 kDa, consistent with the E4 proteins observed in over-expression studies and upon differentiation in semi-solid media (Chapter 5 and Figure 6.5). The faster migrating E4 species was not detectable with an N-terminal specific E1^E4 antibody indicating that this 10.5 kDa protein lacks N-terminal residues and demonstrates that the full-length E1^E4 undergoes N-terminal proteolytic cleavage during the HPV18 virus life cycle (Figure 6.7 A).

Substitution of the leucine di-peptide ^{14}LL^{15} (PPSLL and RRSLL) abrogated formation of the faster migrating 10.5 kDa E4 species during over-expression studies (Chapter 5). However accumulation of the 10.5 kDa polypeptide was observed in addition to the 11.7 kDa full-length E1^E4 protein within lysate derived from raft cultures of HFKs containing these mutant HPV18 genomes (Figure 6.7 A). The level of the processed E4 species observed with these mutations was however considerably lower than that of the wild type protein, suggesting that the di-leucine peptide ^{14}LL^{15} is required for efficient N-terminal proteolysis of the E1^E4 protein. A further reduction in the level of the 10.5 kDa species to barely detectable levels was observed in rafts of cells containing the HPV18 genome with the RRSSS substitution. The ratios of the faster migrating 10.5 kDa E4 species relative to the full-length E1^E4 protein were calculated following densitometry analysis of the Western blot film (Figure 6.7 B). Whilst the processed E4 protein constitutes approximately 30% of total wild type E4 protein levels, the level of proteolytically cleaved E4 species generated within organotypic raft cultures of the leucine-rich motif mutants was only 14.3% of PPSLL, 8.7% of RRSLL and 4.6% of RRSSS E1^E4 protein. An intact leucine-rich motif is therefore required for efficient N-terminal cleavage of E1^E4 during the HPV18 life cycle. This was repeated with consistent results observed and whilst significantly lower levels of the leucine-rich motif

mutants were observed in comparison to wild type E1^E4 protein following cellular differentiation in semi-solid media, the protein levels observed in organotypic raft cultures were equal to or greater that wild type levels.

.

Discussion

6.10   Efficient E1^E4 N-terminal proteolysis occurs during the HPV18 virus life cycle and is dependent upon the leucine-rich motif

Expression of E1^E4 by transient transfection or recombinant viral infection of a number of immortalised and transformed cell lines has demonstrated that the HPV18 E1^E4 protein undergoes post-translational modification by N-terminal proteolysis to generate smaller E4 species (Chapter 5). Mutagenesis studies established that the E1^E4 di-peptide ^{14}LL^{15} within an N-terminal leucine-rich motif of the protein was essential for proteolysis within these cells. To investigate the role of proteolysis of the E1^E4 protein within the HPV18 life cycle, and its dependency upon the leucine-rich motif, HFK cell lines were generated which stably maintained HPV18 genomes containing mutations within the leucine-rich motif of the E4 ORF.

Differentiation of HFK cells harbouring HPV18 genomes in organotypic raft culture demonstrated that the E1^E4 protein undergoes proteolysis to generate a faster migrating E4 species lacking N-terminal residues. This truncated E4 protein is of the same size as the protein species observed within over-expression studies, therefore what was observed within the ectopic system is relevant. Substitution of the first two leucines of the ^{14}LLSLL^{18} E1^E4 sequence in the HPV18 genome resulted in a reduction of 50-65% in the formation of this faster migrating E4 species, indicating that the ^{14}LL^{15} di-peptide of E1^E4 contributes

towards proteolysis of the E1^E4 protein within the replication cycle of HPV18. Disruption of both di-leucine peptides within this motif, with the introduction of an ^{14}RRSSS^{18} mutation, resulted in virtually un-detectable levels of the proteolytically cleaved form of the E4 protein, demonstrating that an intact leucine-rich motif is required for efficient E1^E4 proteolysis. Given that proteolysis of the E1^E4 protein is not entirely abrogated by disruption of the leucine motif it is likely that the site of proteolysis lies outside of this motif. Disruption of the leucine-rich motif of the E1^E4 protein may induce a structural change that limits access of the protease to the recognition site for proteolysis or may alter the recognition site for the protease.

Upon differentiation of HFKs containing wild type HPV18 genomes within organotypic raft culture, the ratio of processed E4 to full-length E1^E4 had increased significantly in comparison with the ratio observed within cells induced to differentiate within semi-solid media. Given that organotypic raft cultures support the complete life cycle, including virion morphogenesis, this suggests that the processed E4 protein accumulates at later stages of the replication cycle.

6.11 The leucine-rich motif of E1^E4 is not required for early stages of the HPV18 life cycle

Whilst E1^E4 proteolysis is perhaps an event that is pertinent to late stages of the virus life cycle, mutations with the E4 leucine-rich motif of HPV16 result in defective viral DNA replication in undifferentiated NIKS cells (Nakahara et al., 2005). Similar changes to the HPV18 E4 sequence however did not compromise replication of HPV18 genomes. The necessity of the leucine-rich sequence within the E4 ORF of HPV16, but not HPV18, for

effective replication of viral genomes during early stages of the virus life cycle may reflect the differences in the role of the E1^E4 protein in viral replication within these HPV types.

Whilst E1^E4 proteolysis did not contribute towards the E1^E4 role in regulating the growth of the basal-like monolayer HFKs, it is possible however, given that some proteolysis of the E1^E4 protein is observed with these mutations, that a low level of processed E4 may be sufficient to contribute towards the growth suppressive function of the HPV18 protein.

6.12 The leucine-rich motif contributes towards the steady-state levels of the E1^E4 protein during the HPV18 replication cycle

The steady-state levels of E1^E4 proteins with substitutions of the leucine-rich motif were significantly reduced from that of the wild type protein upon differentiation of HFKs harbouring HPV18 genomes in semi-solid media. However upon growth of these cells in organotypic raft culture the levels of the mutant E1^E4 proteins were comparable to that of the wild type protein. Given that organotypic raft cultures support later stages of the life cycle, the difference observed between these two methods of differentiation suggests that mutations of the leucine-rich motif may delay accumulation of the E1^E4 protein until later stages in the virus replication cycle.

6.12.1 Mutations of the leucine-rich motif may alter the structural stability of the E1^E4 protein

The leucine-rich motif may play a key role in the structure of the E1^E4 protein and substitution of these residues could inherently alter the structure and stability of E1^E4, resulting in reduced levels of the protein within less differentiated cells. During later stages of the life cycle, as recapitulated within raft cultures, the protein may be stabilised by other

factors or interactions. However, whilst the same [14]LL[15] di-peptide is mutated within the PPSLL and RRSLL mutant proteins, consistently lower levels of the latter protein were observed following differentiation in semi-solid media. Given that this region of the E1^E4 protein is predicted to be α-helical in nature (Roberts et al., 1997; McIntosh et al., 2008), introduction of a proline residue at this site would be predicted to be more, not less disruptive due to steric interference, destabilising protein structure to a greater extent than that of an arginine residue and it therefore seems improbable that this accounts for the observed differences in protein levels.

Due to the overlapping nature of the E2 and E4 ORFs, mutations of E4 are limited to those which introduce silent mutations to the E2 coding sequence. Introduction of mutations can have the unintentional consequence of resulting in a gain of function and mutations of the leucine-rich motif unfortunately introduce consensus motifs into the E1^E4 protein for a number of kinases including: Abl tyrosine kinase, DNA-dependent protein kinase (DNA-PK), PKC, PKA and calmodulin-dependent protein kinase (CaMKII). Phosphorylation within this region of the E1^E4 mutant proteins may result in structural changes, or reduce protein interactions, acting to reduce the stability of the protein. Indeed, phosphorylation has been shown to enhance keratin association and the stability of the HPV16 E1^E4 protein (McIntosh et al., 2008; Wang et al., 2009). However, the various leucine substitutions introduce different putative consensus sites into the HPV18 E1^E4 protein therefore it seems improbable that all three mutant proteins would be targeted by different kinases to mediate this effect in reduced E1^E4 protein stability.

Alternatively, these mutations may change the solubility of the E1^E4 protein, and the observed discrepancy in protein levels of these mutant proteins between the two methods of

differentiation may be due to differences in cellular lysis. Although the raft cultures were solubilised in Laemelli loading buffer, resulting in analysis of total lysate, cells harvested from semi-solid media were solubilised in a buffer containing SDS and urea with the soluble fraction analysed.

6.12.2 The leucine-rich motif may mediate interactions which stabilise the E1^E4 protein

The leucine-rich motif of E1^E4 may be required for formation of protein-protein interactions with cellular factors or viral proteins that are required for the stability of the E1^E4 protein within less differentiated cells. However within the upper layers of the organotypic raft, at later stages in the replication cycle, the importance of these interactions may be negated, possibly due to other interacting protein partners requiring residues outside of the leucine-rich motif. Mutation of the leucine-rich motif of HPV18 E1^E4 may act to prevent stabilisation of the protein through reduced or abrogated association with the keratin network. However the leucine-rich motif of HPV18 E1^E4 has previously (Nakahara et al., 2002), and within this study (Chapter 5), been shown to be dispensable for association of the protein with the simple keratin 18 network, although the contribution of this motif to association with keratin proteins within differentiated keratinocytes has not been addressed. It is possible that residues of the leucine-rich motif may be required to mediate an interaction with complex keratins, or with other interacting protein partners, during differentiation-dependent stages of the virus life cycle to stabilise the protein.

6.13 An intact leucine-rich motif is required for efficient E1^E4 proteolysis and viral genome amplification during the HPV18 virus life cycle

The correlation between a severely impaired ability to produce processed E4 species and the impaired ability to undergo viral genome amplification within HFKs containing HPV18 genomes with the E4 RRSSS mutation, suggests that formation of the processed E4 species is required for efficient viral genome amplification. However, unfortunately it is not possible to determine a conclusive role for E1^E4 proteolysis in viral genome amplification. Mutations within the leucine-rich motif of E1^E4 affect the ability of the protein to undergo efficient proteolysis within over-expression studies and within organotypic raft cultures of HFKs containing HPV18 genomes. However the severely reduced levels of these mutant E1^E4 proteins within cells stimulated to differentiate in semi-solid media, particularly notable with the RRSSS mutant, make it impossible to establish whether a small level of E1^E4 proteolysis observed in the PPSLL and RRSLL mutations is sufficient to support the E4 function of viral genome amplification or that whilst the RRSSS mutation is severely impaired in proteolysis it is actually the lack of full-length E1^E4 within these cells that accounts for the loss of viral genome amplification.

A combination of a reduction in both full-length and processed E4 proteins may contribute towards the impaired ability of the RRSSS mutant genomes to undergo amplification. Indeed the level of the full-length E1^E4 protein may be dependent upon the processed E4 protein for stability through complex formation. The role of processed E4 protein in viral genome amplification may be to stabilise the level of the full-length protein within the cell, or may be to act in complex with the full-length protein, or alone, to perform this viral function. Alternatively the RRSSS mutation may disturb another E1^E4 function aside from proteolysis that is necessary for viral genome amplification.

6.14   Note

I presented the work described within this thesis, particularly detailing the work described within Chapters 5 and 6, as an oral presentation at the 25th International Papillomavirus Conference hosted in Sweden in 2009. The presentation was entitled: Modification of HPV18 E1^E4 by proteolysis, CDK and PKA phosphorylation.

CHAPTER 7

OVERALL DISCUSSION AND

FUTURE DIRECTIONS

7.1   Overview of findings

The work within this thesis has sought to address the hypothesis that HPV18 E1^E4 undergoes modification by proteolysis and phosphorylation, and has investigated the functional significance of these E1^E4 post-translational modifications during the HPV life cycle. This thesis has revealed for the first time that HPV18 E1^E4 exists as a phospho-protein within cells and is a target for *in vitr*o phosphorylation by multiple cellular kinases. Phospho-acceptor residues for these kinases including PKA and CDKs have been identified within this viral protein. This study has also shown that phosphorylation by CDKs is dependent upon a cyclin binding motif (^{43}RRLL^{46}) within the E1^E4 sequence and that integrity of this motif is required for association of the protein with active CDK-cyclin complexes. This association may influence CDK activity since the activity of CDK2-cyclin A was shown to be reduced in the presence of HPV18 E1^E4. Furthermore, it has now been shown that HPV18 E1^E4 is a target for N-terminal proteolysis, and this post-translational modification occurs during the HPV18 replication cycle. An intact leucine-rich motif (LLXLL) in proximity to the N-terminus of E1^E4 is a conserved feature of many E1^E4 proteins and is required for efficient proteolysis of the HPV18 E1^E4 protein. Mutation of this motif in the E4 ORF of HPV18 genomes resulted in reduced viral genome amplification within differentiating HFKs, indicating that this region of the protein is integral to E1^E4 function.

7.2   HPV18 E1^E4 phosphorylation

Whilst the sites of HPV18 E1^E4 phosphorylation are not highly conserved between E1^E4 proteins, the overall positions of phospho-acceptor residues within the E1^E4 protein appear to be similar and are mostly restricted to the central region of the protein (Figure 3.9).
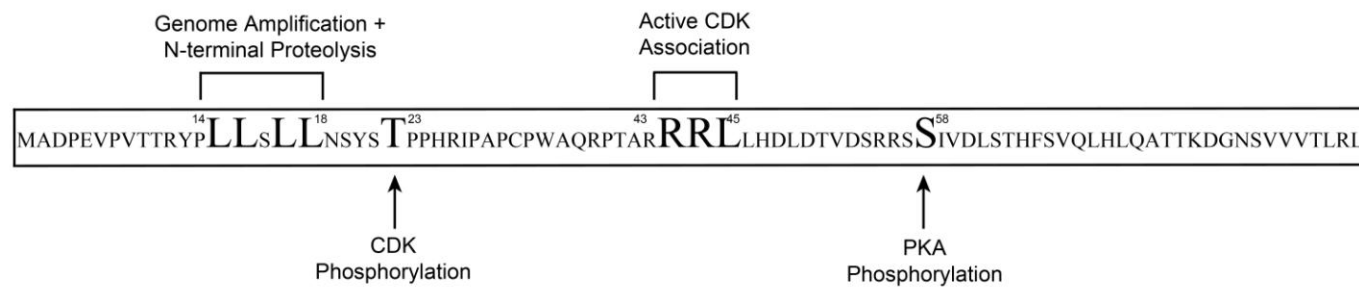
Figure 7.1 A schematic of HPV18 E1^E4 depicting an overview of findings. A leucine-rich motif (LLXLL) at the N-terminus of HPV18 E1^E4 is required for efficient N-terminal proteolysis of the E1^E4 protein and efficient viral genome amplification during the HPV18 life cycle. The positions of CDK and PKA *in vitro* phosphorylation of HPV18 E1^E4 are also indicated, and a cyclin binding motif (RRL) required for association of HPV18 E1^E4 with active CDK-cyclin complexes is highlighted.

Phosphorylation of HPV18 E1^E4 may act to alter the structure of the protein and modulate its biological activity, by enabling or preventing formation of binding sites mediating interactions with partner proteins.

Given that a number of E1^E4 proteins including HPV18 E1^E4, have been shown to be phosphorylated by various cellular kinases (Grand et al., 1989; Bryan et al., 2000; Davy et al., 2006; Wang et al., 2009), it is likely that phosphorylation of E1^E4 may impinge upon the role of this protein during the HPV life cycle. Whilst a phospho-specific antibody has been generated to the CDK phosphorylation site of HPV18 E1^E4, initial attempts to analyse cellular extracts did not succeed in detecting phosphorylated forms of the HPV18 E1^E4 protein. However this antibody may prove to be useful in detection of phosphorylated forms of HPV18 E1^E4 by immunofluorescence, and further work could examine expression of this protein within sections of organotypic rafts cultures of HPV18 genome containing cells, to examine whether the E1^E4 protein is targeted by CDKs within the physiological environment. Similarly, the role of PKA phosphorylation of HPV18 E1^E4 could be examined in this manner. This approach has been previously used to examine MAPK phosphorylation of the HPV16 E1^E4 protein within differentiated cells containing HPV16 genomes (Wang et al., 2009). Interestingly whilst MAPK phosphorylation of HPV16 E1^E4 has been shown to result in increased protein stability and enhanced keratin binding (Wang et al., 2009), within this study there was no evidence from *in vitro* studies that the type 18 protein was a substrate for MAPK.

Another approach to examine the role of E1^E4 CDK phosphorylation is by mutation of the CDK phospho-acceptor site within the E4 ORF of intact HPV18 genomes. This approach was initiated by Dr S. Roberts and Dr G.L. Knight in parallel to the study described within

this thesis. The threonine 23 CDK phospho-acceptor site was mutated to an isoleucine and whilst this mutation did not affect differentiation-dependent viral genome amplification or production of late viral transcripts within differentiating HFKs, it is possible that E1^E4 CDK phosphorylation may affect later stages of the virus life cycle. Further work examining the role of E1^E4 phosphorylation during the HPV replication cycle could also be performed by introduction of an E4 PKA phospho-acceptor mutation into HPV18 genomes and assessing the affect of this within HFKs. HPV18 E1^E4 may also be targeted for phosphorylated by other kinase and contains putative consensus sites for PKC and CKII (Figure 3.1). Further work could examine whether HPV18 E1^E4 is indeed a substrate for these kinases and could examine the implications of phosphorylation at these sites.

7.3   Deciphering the role of HPV18 E1^E4 association with CDKs

The ability of HPV18 E1^E4 to associate with active CDKs is likely to have functional implications, not only in enabling HPV18 E1^E4 to undergo phosphorylation but may also contribute towards wider biological activities of the virus. Loss of integrity of the ^{43}RRLL^{46} sequence of HPV18 E1^E4 has been shown to abrogate cytoplasmic sequestration of cyclin A and cyclin B and their CDK2 and CDK1 binding partners, and the ability of the E1^E4 protein to induce a G2 arrest is subsequently lost (G.L. Knight and S. Roberts, unpublished data). It has previously been hypothesised that the ability of E1^E4 proteins to induce a G2 arrest drives cells into a pseudo S phase in which conditions are conducive to viral genome replication (Davy et al., 2002). Furthermore a role for the HPV18 E1^E4 protein has been demonstrated in viral genome amplification, whereby loss of expression of the full length E1^E4 protein has been shown to disrupt differentiation-dependent HPV18 genome amplification (Wilson et al., 2007). However, introduction of E4 mutations which are

defective for G2 arrest into HPV18 genomes did not alter the extent of amplification of these genomes within differentiating HFKs in comparison with wild type HPV18 genomes (G.L. Knight and S. Roberts, unpublished data). The ability of the HPV18 E1^E4 protein to induce a G2 arrest within ectopic expression systems is therefore not relevant during this stage of the virus life cycle and induction of a G2 arrest is thereby not the mechanism by which HPV18 E1^E4 protein contributes towards viral genome amplification. This is a particularly interesting finding as a recent paper has suggested that generation of a G2 arrested cellular environment is integral to viral genome amplification (Chow et al., 2009). Amplification of HPV18 genomes within organotypic raft cultures followed host DNA replication within S phase cells, occurring later within cells in which cytoplasmic cyclin B accumulated, indicative of a G2 arrest (Chow et al., 2009). Cyclin B accumulated within the cytoplasm of these differentiating epithelia prior to detectable expression of E1^E4 suggesting that this viral protein may not be required for generation of a G2 arrested cellular environment within cells supporting viral genome amplification.

Whilst the ability of E1^E4 to induce a G2 arrest does not appear to enable viral genome amplification (G.L. Knight and S. Roberts, unpublished data), the conserved nature of the G2 arrest function between diverse E1^E4 proteins suggests however that this is a relevant biological action for the virus (Nakahara et al., 2002; Davy et al., 2002; Knight et al., 2004), and the ability of E1^E4 to induce a G2 arrest may serve other functions during the virus life cycle. Induction of a G2 arrest is a common feature of a number of viruses including the human immunodeficiency virus (HIV) and the avian coronavirus, infectious bronchitis virus (IBV) and has been shown to have a number of benefits for viruses (Davy & Doorbar, 2007). It has been reported that the G2 arrest function of IBV enables enhanced expression of viral proteins (Dove et al., 2006), whilst for HIV, the G2 environment increase viral transcription

leading to increased virus production (Goh et al., 1998). It is conceivable that the G2 arrest induced by the HPV18 E1^E4 protein may serve to enhance late functions of the virus life cycle such as capsid protein expression or virion morphogenesis.

HPV18 E1^E4 may act to modulate the activity of associated CDKs, and the presence of this viral protein has been shown to inhibit the activity of CDK2-cyclin A *in vitro*. Further work will be needed to address whether this inhibition is dependent upon the cyclin binding motif of HPV18 E1^E4. The ability of E1^E4 to reduce CDK2 activity may be an additional mechanism by which the protein may induce a G2 arrest within cells. Alternatively inhibition of CDK activity may serve to prevent CDK2-cyclin A directed cellular replication, in a mechanism to enable enhanced viral replication. Further work is needed to establish whether this reduction in activity occurs within a physiological environment, and the level of CDK activity could be examined within keratinocytes containing wild type HPV18 genomes. The ability of the protein to associate with a number of CDK-cyclin complexes may mediate an as yet undefined function of the E1^E4 protein in the virus life cycle.

Association of E1^E4 with CDKs, in addition to resulting in phosphorylation of this viral protein, may also act to direct CDK activity towards other proteins. A recent study has demonstrated an interaction between HPV16 E1^E4 and E2 proteins, with this association resulting in relocalisation of E2 to the cytoplasm and stabilisation of both proteins (Davy et al., 2009). It is proposed that at different times E1^E4 may have differential effects on the stability and localisation of the E2 protein, thereby acting to increase or decrease E2 activity at different stages of the virus life cycle (Davy et al., 2009). Levels of the E2 protein have been shown to be stabilised during S phase, coincident with increased phosphorylation of the protein, with a putative kinase proposed to be CDK2, which has been shown to target E2 for

in vitro phosphorylation (Johansson et al., 2009). Perhaps as a result of the interaction with E2, E1^E4 may act to direct CDK2 towards E2, resulting in enhanced phosphorylation and subsequent stabilisation of the E2 protein. Given the diverse range of functions that E2 performs during the life cycle, the association with E1^E4 could act to regulate a number of E2 activities. Whilst this interaction may contribute towards the ability of E2 to direct viral genome amplification, this would be unlikely to occur as a result of the ability of E1^E4 to interact with CDKS, as association of E1^E4 with CDKs does not appear to contribute towards viral DNA replication (G.L. Knight and S. Roberts, unpublished data). E2 also plays a role in viral transcription, and a role in virion morphogenesis has also been proposed for the E2 protein, demonstrated by enhanced packaging of viral genomes into virions in the presence of BVP1 E2 (Zhao et al., 2000). The ability of E1^E4 to associate with CDKs could act to stabilise the E2 protein to enable key roles during the virus replication cycle to be performed.

7.4 The role of HPV18 E1^E4 N-terminal proteolysis during the HPV18 life cycle

Efficient N-terminal proteolysis of HPV18 E1^E4 within differentiating keratinocytes has been shown to require an intact leucine-rich motif and both di-leucine pairs of this motif appear to contribute towards E1^E4 proteolysis. Mutation of the entire E4 leucine motif within the HPV18 genome results in a severely impaired ability to undergo viral genome amplification within differentiating keratinocytes although plasmid replication within undifferentiated cells is unaffected. This work was performed within two donor HFK backgrounds with introduction of an RRSSS mutation, consistently resulting in impaired amplification of HPV18 genomes. A third HFK cell line has since been generated in which

RRSSS mutant genomes also displays significantly reduced viral genome amplification in comparison with the wild type HPV18 genome (G.L. Knight).

Identification of an E1^E4 sequence ([14]LLSLL[18]), which upon mutation, both prevents efficient proteolysis of the protein and severely reduces amplification of HPV18 genomes is an important finding. It is tempting to speculate that the dual impact of this mutation on both post-translational modification and function of E1^E4 is more than a coincidence and that the ability of the E1^E4 protein to undergo proteolysis is crucial for the HPV life cycle. However further work is needed to investigate the discrepancies observed between protein levels of this mutant protein upon cellular differentiation by organotypic raft culture, and with semi-solid media, to establish whether this mutation acts to impair viral genome amplification through instability of the E1^E4 protein. It will be important to examine viral genome amplification within sections of organotypic raft cultures of HFKs containing the E4 LLXLL mutant HPV18 genome, where levels of the mutant E1^E4 protein are comparable with wild type E1^E4. Mutation of the E4 [17]LL[18] di-peptide in isolation, may also enable a greater understanding of the contribution of this region of the E1^E4 protein to viral genome amplification. Identification of a mutant protein which is impaired in the ability to undergo proteolysis but does not destabilise the E1^E4 protein could clarify whether viral genome amplification requires N-terminal proteolysis of E1^E4.

Whilst efforts to identify the protease responsible for proteolytic cleavage of HPV18 E1^E4 have been undertaken, further work is needed to establish the target protease. This could lead to identification of the exact residues of cleavage of the HPV18 E1^E4 and could also enable inhibition of this proteolytic event within differentiating keratinocytes for performing further functional analyses. Future work could also be performed to establish whether the smaller E4

polypeptide displays any unique functions from that of the full length protein, as has been shown to be the case for the HPV1 truncated E4 protein (Knight et al., 2004; Knight et al., 2006). The cleaved product of the HPV18 protein may associate with novel binding partners and a proteomic analysis could be undertaken using lysates from differentiated cells to identify interactors specific to the truncated E4 protein. Since the N-terminally cleaved HPV18 E4 polypeptide and full length E1^E4 species are able to interact this may extent the functional repertoire of the protein.

The abundance of proteolytically cleaved HPV18 E4 protein increases within differentiated cells, and perhaps cleaved products perform functions within the uppermost cells of the lesion. Further work could examine the expression of the L1 capsid protein within differentiated cells harbouring the leucine-rich E4 mutant genome to establish whether severely reduced expression of the truncated E4 protein impairs production of the late capsid protein. Infectivity assays could also be performed to assess the effect of impaired E1^E4 proteolysis on the production of infectious virus.

7.5   Limitations of the research

Whilst growth of HFKs harbouring HPV genomes is a widely used and highly regarded system in which to study the HPV life cycle, limitations in the ability to manipulate this system must be acknowledged. Introduction of mutations to the E4 ORF, whilst resulting in silent mutations within the E2 ORF, may have adverse effects on previously unidentified, cis-acting elements within the genomic DNA which influence replication of the DNA. Indeed within the HPV16 E4 ORF a splicing enhancer element has been identified which is required for mRNA splicing and polyadenylation of early transcripts, and inhibition of late gene expression (Rush et al., 2005). Whilst disruption of potential cis-acting elements by mutation

of the leucine-rich motif of E4 is an unavoidable possibility, introduction of exogenous wild type E1^E4 would address this issue. E1^E4 could be delivered into HFKs containing the leucine-rich E4 mutant genome by transfection of cells with wild type E1^E4, or by the use of recombinant viruses such as E4 expressing adenoviruses, to establish whether viral genome amplification can be restored within these cells.

7.6 Final statement

The research detailed within this thesis is the first study undertaken to examine the role of post-translational modifications of the HPV18 E1^E4 protein. The findings of this study have supported the hypothesis that the HPV18 E1^E4 protein undergoes proteolysis and phosphorylation, and further supports the proposal that these post-translation modifications may be a conserved mechanism for regulation of E1^E4 function. Furthermore they have indicated that proteolysis of the HPV18 E1^E4 protein may be required for amplification of viral genomes, one of the most tangible of E1^E4 functions, in terms of our understanding of the role of this protein during virus life cycle. Whilst the role of E1^E4 phosphorylation and association with CDKs is yet to be determined, the finding that E1^E4 is able inhibit the activity of CDK2-cyclin A, at least *in vitro,* is a novel finding and may in part explain this interaction and may delineate an as yet undefined function of the E1^E4 protein.

This work has furthered our knowledge of E1^E4 proteins and whilst full comprehension of this protein is far from complete, identification and characterisation of modified forms of HPV18 E1^E4 provides a greater understanding of this complex viral protein. Future research into HPV18 E1^E4, whether resulting in identification of novel binding partners or indeed novel functions, will now be better positioned, with a greater knowledge of this protein, to be able to examine the contribution of this protein and the various modified forms

it may constitute, to the virus life cycle. Furthermore, the findings of this study may indeed help to explain the outcome of future studies.

APPENDIX

S57N

WT E4: R56 **S57** S58
GAGAAGCAGC
WT E2: E256 K257 Q258

E4 S57N: R56 **N57** S58
GAGAAACAGC
WT E2: E256 K257 Q258

S58N

WT E4: S57 **S58** I59
AAGCAGCATT
WT E2: K257 Q258 H259

E4 S58N: S57 **N58** I59
AAGCAACATT
WT E2: K257 Q258 H259

S22N

WT E4: Y21 **S22** T23
CTACAGCACA
WT E2: L221 Q222 H223

E4 S22N: Y21 **N22** T23
CTACAACACA
WT E2: L221 Q222 H223

T23I

WT E4: S22 **T23** P24
CAGCACACCC
WT E2: Q222 H223 T224

E4 T23I: S22 **I23** P24
CAGCATACCC
WT E2: Q222 H223 T224

RXL1

WT E4: R42 **R43** R44 **L45** L46
CAGACGTCGGCTGCTA
WT E2: Q242 T243 S244 A245 A246

E4 RXL1: R42 **H43** R44 **R45** L46
CAGACATCGGCGGCTA
WT E2: Q242 T243 S244 A245 A246

RXL2

WT E4: R43 **R44** L45 **L46** H47
ACGTCGGCTGCTACAC
WT E2: T243 S244 A245 A246 T247

E4 RXL2: R43 **P44** L45 **R46** H47
ACGTCCGCTGCGACAC
WT E2: T243 S244 A245 A246 T247

PPSLL

WT E4: P13 **L14 L15** S16
TCCGCTACTCAGC
WT E2: S213 A214 T215 Q216

E4 PPSLL: P13 **P14 P15** S16
TCCGCCACCCAGC
WT E2: S213 A214 T215 Q216

RRSLL

WT E4: P13 **L14 L15** S16
TCCGCTACTCAGC
WT E2: S213 A214 T215 Q216

E4 RRSLL: P13 **R14 R15** S16
TCCGCGACGCAGC
WT E2: S213 A214 T215 Q216

LLSSS

WT E4: S16 **L17 L18** N19
CAGCTTGTTAAAC
WT E2: Q216 L217 V218 K219

E4 LLSSS: S16 **S17 S18** N19
CAGCTCGTCAAAC
WT E2: Q216 L217 V218 K219

Figure A1: Substitution mutations of the E4 open reading frame (ORF) used within the study which do not alter the E2 coding sequence within the HPV18 genome. Due to the overlapping nature of the E2 and E4 coding sequences substitution mutations were chosen which do not alter the E2 coding sequence. For each mutation, a schematic depicting the wildtype (WT) nucleotide sequence of the HPV18 genome is shown with the E4 and E2 amino acids the sequence codes for indicated above and below the sequence respectively. Below, the HPV18 nucleotide sequence containing substitution mutations is shown. Whilst these mutations do not alter the amino acid residues coded for in the E2 ORF, the E4 ORF codes for different amino acid residues. Target E4 amino acid residues are indicated in large typeface.

APPENDIX II

Generation of an Simian virus 40 (SV40)-epithelial expression system for

expression of HPV18 E1^E4

The use of recombinant SV40 E1^E4 viruses has proved to be very successful in the study of HPV1 and HPV16 E1^E4 proteins, producing high levels of E1^E4 proteins upon expression in SV40-transformed epithelial cells (Roberts et al., 1993). Generation of a recombinant HPV18 E1^E4 SV40 virus was therefore undertaken with a view to generating an expression system in which HPV18 E1^E4 can be characterised.

Briefly, HPV18 E1^E4 cDNA was cloned into the SV40 early replacement vector (pAP16) (Figure A2, A). The plasmid DNA was then digested and re-ligated to remove the bacterial pBR322 genome, which allows manipulation in bacterial strains, and electroporated into the SV40-transformed monkey epithelial cell line, COS-1 packaging cell line which constitutively expresses the SV40 large T antigen required to enable viral replication. Cytopathic effect (CPE) was observed after 7 days and the virus was harvested.

The cellular localisation of HPV18 E1^E4 was examined by immunofluorescence following infection of COS-1 epithelial cells with the rSV40 E1^E4 virus (Figure A2, B). HPV18 E1^E4 was detected with a polyclonal E1^E4 antibody raised against a GST-HPV18 E1^E4 fusion protein (Wilson et al., 2007) and with an N-terminal specific monoclonal E4 antibody raised against a N-terminal HPV16 peptide with cross specificity for the HPV18 E1^E4 N-terminal residues (Roberts et al., 2003). HPV18 E1^E4 is located primarily within the cytoplasm where it displays a predominantly filamentous distribution (Figure A2, B, a) or following collapse of this filamentous network, forms juxta-nuclear bundles (Figure A2, B, b), most likely through association with the keratin intermediate filament network. This is

consistent with observations of other HPV E1^E4 proteins which form highly organised filamentous networks within the cytoplasm with subsequent collapse of HPV16 E1^E4 networks (Doorbar et al., 1991; Roberts et al., 1993).

However the proportion of rSV40-infected cells expressing detectable HPV18 E1^E4 protein by immunofluorescence was less than 5%, suggesting that the efficiency of infection by the virus was extremely low or that the viral protein was expressed at undetectable levels in the majority of cells. Further work with rSV40 viruses was forestalled as the E1^E4 protein was expressed at levels insufficient to be detected by Western blotting and this was therefore not an appropriate system for the study of HPV18 E1^E4 proteolysis.
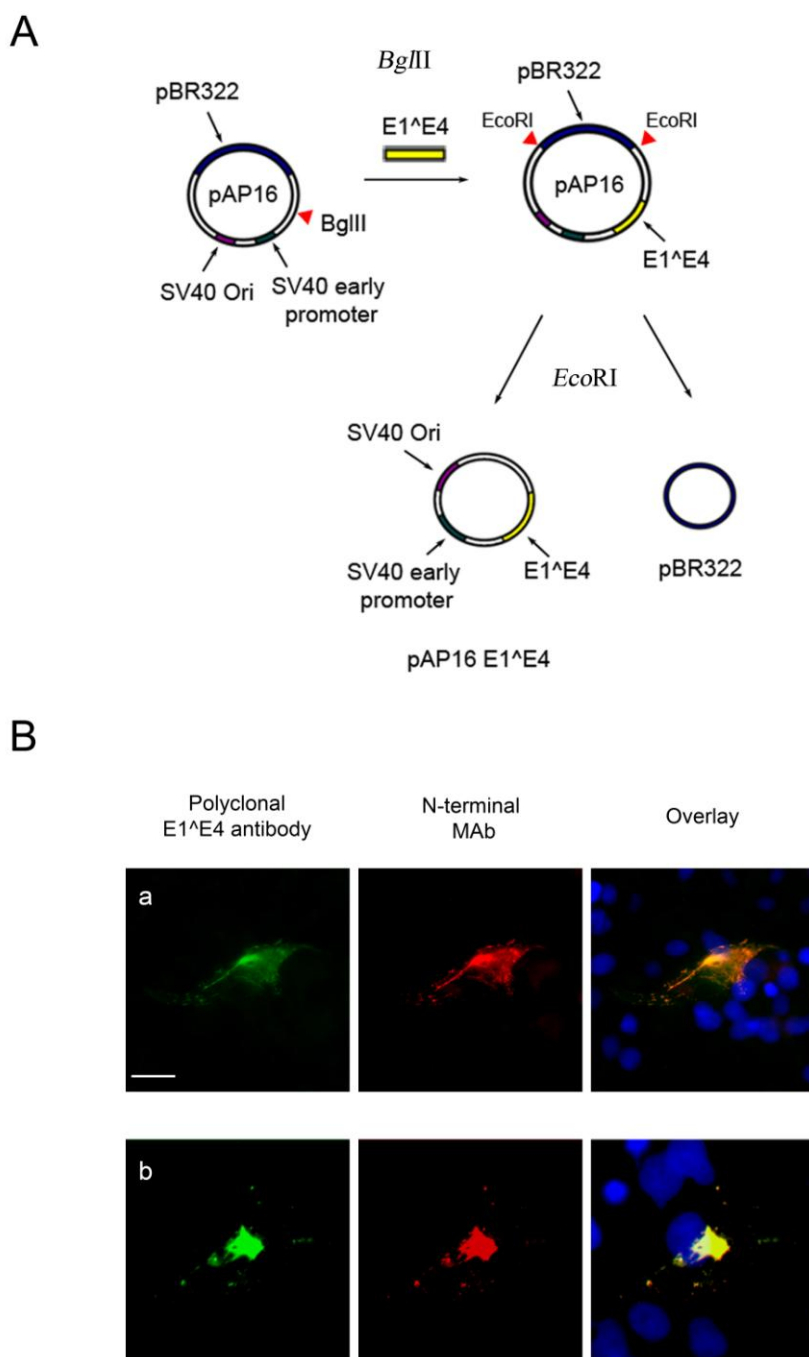
Figure A2: Generation and expression of a recombinant HPV18 E1^E4 SV40 virus. (A) HPV18 E1^E4 cDNA sequence was inserted into the pAP16 vector at the *Bgl*II site. The pBR322 bacterial sequence was then removed by *Eco*RI digestion and the plasmids were recircularised prior to electroporation into COS-1 cells. The recombinant virus was harvested upon observation of cytopathic effect. (B) Cellular distribution of HPV18 E1^E4 in COS-1 cells after infection with recombinant virus. Cells were fixed in 4% paraformaldehyde and permeabilised in acetone, then dual-stained with a rabbit polyclonal E1^E4 antibody and an N-terminal specific E1^E4 monoclonal mouse antibody. Immune complexes were visualised with an anti-rabbit Alexa Fluor 488 antibody (green) and an anti-mouse Alexa Fluor 596 antibody (red) and counter-stained with DAPI for detection of nuclei (blue). E1^E4 is predominantly cytoplasmic and forms cytoplasmic filamentous networks (a) which subsequently collapse to form juxta-nuclear bundles (b). Bar, 10μM.

APPENDIX III

```
AMINO ACID NUMBER            1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18
AMINO ACID SEQUENCE          M    A    D    P    E    V    P    V    T    T    R    Y    P    L    L    S    L    L
NUCLEOTIDE SEQUENCE (WT)    ATG  GCT  GAT  CCA  GAA  GTA  CCA  GTG  ACG  ACA  CGG  TAT  CCG  CTA  CTC  AGC  TTG  TTA
NUCLEOTIDE SEQUENCE (CO)    ATG  GCC  GAC  CCC  GAG  GTG  CCC  GTG  ACA  ACC  AGG  TAC  CCT  CTG  CTC  AGC  CTG  CTC


AMINO ACID NUMBER           19   20   21   22   23   24   25   26   27   28   29   30   31   32   33   34   35   36
AMINO ACID SEQUENCE          N    S    Y    S    T    P    P    H    R    I    P    A    P    C    P    W    A    P
NUCLEOTIDE SEQUENCE (WT)    AAC  AGC  TAC  AGC  ACA  CCC  CCT  CAC  CGT  ATT  CCA  GCA  CCG  TGT  CCG  TGG  GCA  CCG
NUCLEOTIDE SEQUENCE (CO)    AAC  AGC  TAC  AGC  ACC  CCA  CCT  CAC  AGG  ATC  CCC  GCC  CCC  TGT  CCC  TGG  GCC  CCC


AMINO ACID NUMBER           37   38   39   40   41   42   43   44   45   46   47   48   49   50   51   52   53   54
AMINO ACID SEQUENCE          Q    R    P    T    A    R    R    R    L    L    H    D    L    D    T    V    D    S
NUCLEOTIDE SEQUENCE (WT)    CAA  AGA  CCT  ACG  GCC  AGA  CGT  CGG  CTG  CTA  CAC  GAC  CTG  GAC  ACT  GTG  GAC  TCG
NUCLEOTIDE SEQUENCE (CO)    CAG  AGG  CCC  ACC  GCC  AGG  AGG  AGG  CTG  CTG  CAC  GAC  CTG  GAC  ACC  GTG  GAC  AGC


AMINO ACID NUMBER           55   56   57   58   59   60   61   62   63   64   65   66   67   68   69   70   71   72
AMINO ACID SEQUENCE          R    R    S    S    I    V    D    L    S    T    H    F    S    V    Q    L    H    L
NUCLEOTIDE SEQUENCE (WT)    CGG  AGA  AGC  AGC  ATT  GTG  GAC  CTG  TCA  ACC  CAC  TTC  TCG  GTG  CAG  CTA  CAC  CTA
NUCLEOTIDE SEQUENCE (CO)    AGG  AGG  AGC  AGC  ATC  GTG  GAC  CTG  AGC  ACC  CAC  TTC  AGC  GTG  CAG  CTG  CAC  CTG


AMINO ACID NUMBER           73   74   75   76   77   78   79   80   81   82   83   84   85   86   87   88  STOP
AMINO ACID SEQUENCE          Q    A    T    T    K    D    G    N    S    V    V    V    T    L    R    L
NUCLEOTIDE SEQUENCE (WT)    CAG  GCA  ACA  ACA  AAA  GAC  GGA  AAC  TCT  GTA  GTG  GTA  ACA  CTA  CGC  CTA  TAA
NUCLEOTIDE SEQUENCE (CO)    CAG  GCC  ACC  ACC  AAG  GAC  GGG  AAC  AGC  GTG  GTG  GTG  ACC  CTG  AGG  CTG  TAG
```

Figure A2: Codon-optimisation of the HPV18 E1^E4 nucleotide sequence. A number of codons were converted within HPV18 E1^E4 wildtype (WT) nucleotide sequence to those more commonly used in genes of the human genome to generate a codon-optimised E1^E4 sequence (Dr I. Bell). Amino acid sequences and residue numbers are indicated above the HPV18 E1^E4 wildtype (WT) nucleotide sequence with the codon-optimised (CO) sequence shown below. Mutated residues are indicated in bold typeface

APPENDIX IV

Investigation into the identification of the protease responsible for HPV18 E1^E4

proteolysis with the use of protease inhibitors

In an attempt to identify the protease responsible for proteolysis of the HPV18 E1^E4 protein a number of commercially available protease inhibitors were assayed for their ability to abrogate formation of the smaller E4 species. Serine and cysteine proteases constitute two large families of proteases which are categorised by the residues within their active sites. Two broad range inhibitors of these proteases were assayed; E64, an inhibitor of cysteine proteases and AEBSF, an inhibitor of serine proteases, which specifically inhibits: chymotrypsin, kallikrein, plasmin, thrombin, trypsin and related thrombolytic enzymes. The protease calpain is utilised by the HPV E7 protein to promote cleavage of the Rb protein, (Darnell et al., 2007), therefore a range of calpain inhibitors were also assayed to examine whether it is possible that the same protease is utilised by the virus to induce N-terminal proteolytic cleavage of the HPV18 E1^E4 protein. These included: Calpeptin, which targets calpains I and II and papain, and ALLN and ALLM inhibitors which specifically target calpains I and II, cathepsin B and cathepsin L.

Inhibitors were added to the media in which the cells were grown 2 h prior to transfection at concentrations and durations previously used (Richards et al., 2006), or recommended by the manufacturer. Following transfection, cells were grown in media containing the protease inhibitor until harvested, and the ability of the E1^E4 protein to undergo proteolysis was assessed by Western blot analysis with a polyclonal E1^E4 antibody (Figure A4 A). The formation of the full-length E1^E4 protein and the smaller processed E4 species following treatment with the aforementioned inhibitors was however comparable to that of untreated cells, suggesting that the proteases which are targeted by these panel of inhibitors are not

responsible for cleavage of the E1^E4 protein. The effectiveness of the inhibitors used at these concentrations within this study, was not be confirmed, therefore a definitive conclusion cannot be made regarding ability of the corresponding proteases to target HPV18 E1^E4 for N-terminal proteolysis. Although knowledge of the protease required for E1^E4 proteolysis would have been desirable, this was not pursued further due to time constraints and partly due to the expensive nature of utilising commercial inhibitors.

A
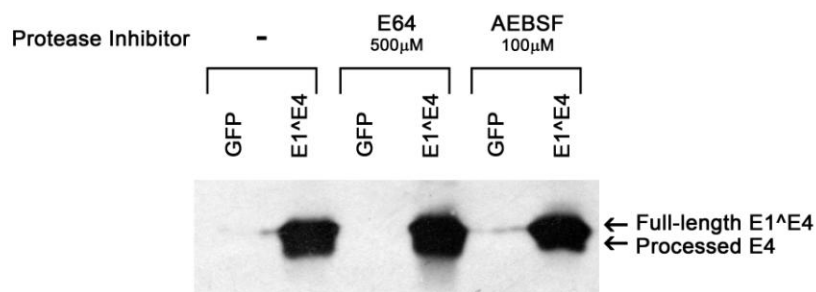
B
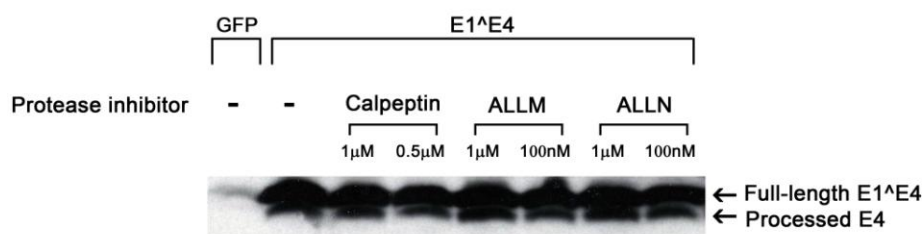
Figure A4: The use of protease inhibitors to examine N-terminal proteolysis of HPV18 E1^E4. Prior to transfection of 293T cells with codon-optimised E1^E4, cells were incubated in media containing commercially available protease inhibitors at stated concentrations. Following a 6 hour transfection, media containing the protease inhibitor was replaced and cells were harvest 48 h post-transfection and analysed by Western blot analysis with a polyclonal E1^E4 antibody. The inhibitors used were as follows: **(A)** E64 (500 μM) and AEBSF (100 μM), **(B)** Calpeptin (1 μM and 0.5 μM), ALLM (1 μM and 100 nM) and ALLM (1 μM and 100 nM). The use of inhibitors at the concentrations stated did not impair the production of smaller E4 species which in all cases was comparable to untreated cells.

PUBLICATIONS AND PRESENTATIONS OF RESEARCH

Journal papers

Manuscript in preparation:

Knight, G.L.,[†] Pugh, A.G.,[†] Yates, E., Bell, I., Wilson, W., Laimins, L.L., Roberts, S.,
[†]These authors contributed equally to this work.

A cyclin binding motif in human papillomavirus type 18 (HPV18) E1^E4 is necessary for cytoplasmic sequestration of CDK-cyclin complexes and G2/M cell cycle arrest but is not required for differentiation-dependent HPV18 genome amplification.

Conference Abstracts

Oral Presentation:

A Pugh, G Knight, I Bell, S, Roberts (2009) Modification of HPV18 E1^E4 by proteolysis, CDK and PKA phosphorylation. 25th International Papillomavirus Conference, Malmo, Sweden.

Poster Presentation:

A G Pugh, G L Knight, I Bell, S Roberts (2006) Post-translational modification of E1^E4 in the HPV18 replication cycle. 23rd International Papillomavirus Conference, Prague, Czech Republic.

REFERENCES

References

Agosti, J. M. & Goldie, S. J. 2007, "Introducing HPV vaccine in developing countries--key challenges and issues", *N.Engl.J.Med.*, vol. 356, no. 19, pp. 1908-1910.

Ashmole, I., Gallimore, P. H., & Roberts, S. 1998, "Identification of conserved hydrophobic C-terminal residues of the human papillomavirus type 1 E1E4 protein necessary for E4 oligomerisation in vivo", *Virology*, vol. 240, no. 2, pp. 221-231.

Ashrafi, G. H., Brown, D. R., Fife, K. H., & Campo, M. S. 2006, "Down-regulation of MHC class I is a property common to papillomavirus E5 proteins", *Virus Res.*, vol. 120, no. 1-2, pp. 208-211.

Avvakumov, N., Torchia, J., & Mymryk, J. S. 2003, "Interaction of the HPV E7 proteins with the pCAF acetyltransferase", *Oncogene*, vol. 22, no. 25, pp. 3833-3841.

Balsitis, S., Dick, F., Dyson, N., & Lambert, P. F. 2006, "Critical roles for non-pRb targets of human papillomavirus type 16 E7 in cervical carcinogenesis", *Cancer Res.*, vol. 66, no. 19, pp. 9393-9400.

Banks, L., Edmonds, C., & Vousden, K. H. 1990, "Ability of the HPV16 E7 protein to bind RB and induce DNA synthesis is not sufficient for efficient transforming activity in NIH3T3 cells", *Oncogene*, vol. 5, no. 9, pp. 1383-1389.

Barbosa, M. S., Edmonds, C., Fisher, C., Schiller, J. T., Lowy, D. R., & Vousden, K. H. 1990, "The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation", *EMBO J.*, vol. 9, no. 1, pp. 153-160.

Bardwell, L. 2006, "Mechanisms of MAPK signalling specificity", *Biochem.Soc.Trans.*, vol. 34, no. Pt 5, pp. 837-841.

Baseman, J. G. & Koutsky, L. A. 2005, "The epidemiology of human papillomavirus infections", *J.Clin.Virol.*, vol. 32 Suppl 1, p. S16-S24.

Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., & Gygi, S. P. 2006, "A probability-based approach for high-throughput protein phosphorylation analysis and site localization", *Nat.Biotechnol.*, vol. 24, no. 10, pp. 1285-1292.

Bedell, M. A., Hudson, J. B., Golub, T. R., Turyk, M. E., Hosken, M., Wilbanks, G. D., & Laimins, L. A. 1991, "Amplification of human papillomavirus genomes in vitro is dependent on epithelial differentiation", *J.Virol.*, vol. 65, no. 5, pp. 2254-2260.

Bekkers, R. L., Massuger, L. F., Bulten, J., & Melchers, W. J. 2004, "Epidemiological and clinical aspects of human papillomavirus detection in the prevention of cervical cancer", *Rev.Med.Virol.*, vol. 14, no. 2, pp. 95-105.

Bell, I., Martin, A., & Roberts, S. 2007, "The E1circumflexE4 protein of human papillomavirus interacts with the serine-arginine-specific protein kinase SRPK1", *J.Virol.*, vol. 81, no. 11, pp. 5437-5448.

Bernat, A., Avvakumov, N., Mymryk, J. S., & Banks, L. 2003, "Interaction between the HPV E7 oncoprotein and the transcriptional coactivator p300", *Oncogene*, vol. 22, no. 39, pp. 7871-7881.

Blom, N., Gammeltoft, S., & Brunak, S. 1999, "Sequence and structure-based prediction of eukaryotic protein phosphorylation sites", *J.Mol.Biol.*, vol. 294, no. 5, pp. 1351-1362.

Bousarghin, L., Touze, A., Sizaret, P. Y., & Coursaget, P. 2003, "Human papillomavirus types 16, 31, and 58 use different endocytosis pathways to enter cells", *J.Virol.*, vol. 77, no. 6, pp. 3846-3850.

Bouvard, V., Storey, A., Pim, D., & Banks, L. 1994, "Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes", *EMBO J.*, vol. 13, no. 22, pp. 5451-5459.

Boyer, S. N., Wazer, D. E., & Band, V. 1996, "E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway", *Cancer Res.*, vol. 56, no. 20, pp. 4620-4624.

Brandsma, J. L., Yang, Z. H., Barthold, S. W., & Johnson, E. A. 1991, "Use of a rapid, efficient inoculation method to induce papillomas by cottontail rabbit papillomavirus DNA shows that the E7 gene is required", *Proc.Natl.Acad.Sci.U.S.A*, vol. 88, no. 11, pp. 4816-4820.

Bray, F., Carstensen, B., Moller, H., Zappa, M., Zakelj, M. P., Lawrence, G., Hakama, M., & Weiderpass, E. 2005, "Incidence trends of adenocarcinoma of the cervix in 13 European countries", *Cancer Epidemiol.Biomarkers Prev.*, vol. 14, no. 9, pp. 2191-2199.

Breitburd, F., Croissant, O., & Orth, G. 1987, "Expression of human papillomavirus type 1 E4 gene products in warts", *In: Steinberg BM, Brandsma JL, Taichman LB (eds.), Cancer Cells, Vol.5.Papillomaviruses Cold Spring Harbor Laboratory, Cold Spring Harbor, New York* pp. 115-122.

Breitburd, F., Kirnbauer, R., Hubbert, N. L., Nonnenmacher, B., Trin-Dinh-Desmarquet, C., Orth, G., Schiller, J. T., & Lowy, D. R. 1995, "Immunization with viruslike particles from cottontail rabbit papillomavirus (CRPV) can protect against experimental CRPV infection", *J.Virol.*, vol. 69, no. 6, pp. 3959-3963.

Brown, D. R., Brown, C. R., & Lehr, E. E. 2004, "Intracellular expression patterns of the human papillomavirus type 59 E1/E4 protein in COS cells, keratinocytes, and genital epithelium", *Intervirology*, vol. 47, no. 6, pp. 321-327.

Brown, N. R., Noble, M. E., Endicott, J. A., & Johnson, L. N. 1999, "The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases", *Nat.Cell Biol.*, vol. 1, no. 7, pp. 438-443.

Bryan, J. T. & Brown, D. R. 2000, "Association of the human papillomavirus type 11 E1()E4 protein with cornified cell envelopes derived from infected genital epithelium", *Virology*, vol. 277, no. 2, pp. 262-269.

Bryan, J. T., Han, A., Fife, K. H., & Brown, D. R. 2000, "The human papillomavirus type 11 E1E4 protein is phosphorylated in genital epithelium", *Virology*, vol. 268, no. 2, pp. 430-439.

Buck, C. B., Thompson, C. D., Pang, Y. Y., Lowy, D. R., & Schiller, J. T. 2005, "Maturation of papillomavirus capsids", *J.Virol.*, vol. 79, no. 5, pp. 2839-2846.

Bulk, S., Berkhof, J., Bulkmans, N. W., Zielinski, G. D., Rozendaal, L., van Kemenade, F. J., Snijders, P. J., & Meijer, C. J. 2006, "Preferential risk of HPV16 for squamous cell carcinoma and of HPV18 for adenocarcinoma of the cervix compared to women with normal cytology in The Netherlands", *Br.J.Cancer*, vol. 94, no. 1, pp. 171-175.

Carrillo-Infante, C., Abbadessa, G., Bagella, L., & Giordano, A. 2007, "Viral infections as a cause of cancer (review)", *Int.J.Oncol.*, vol. 30, no. 6, pp. 1521-1528.

Castellsague, X. & Munoz, N. 2003, "Chapter 3: Cofactors in human papillomavirus carcinogenesis--role of parity, oral contraceptives, and tobacco smoking", *J.Natl.Cancer Inst.Monogr* no. 31, pp. 20-28.

Chemin, I. & Zoulim, F. 2009, "Hepatitis B virus induced hepatocellular carcinoma", *Cancer Lett.*, vol. 286, no. 1, pp. 52-59.

Chen, G. & Stenlund, A. 2002, "Sequential and ordered assembly of E1 initiator complexes on the papillomavirus origin of DNA replication generates progressive structural changes related to melting", *Mol.Cell Biol.*, vol. 22, no. 21, pp. 7712-7720.

Chen, G. & Stenlund, A. 2001, "The E1 initiator recognizes multiple overlapping sites in the papillomavirus origin of DNA replication", *J.Virol.*, vol. 75, no. 1, pp. 292-302.

Cheng, S., Schmidt-Grimminger, D. C., Murant, T., Broker, T. R., & Chow, L. T. 1995, "Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes", *Genes Dev.*, vol. 9, no. 19, pp. 2335-2349.

Chesters, P. M., Vousden, K. H., Edmonds, C., & McCance, D. J. 1990, "Analysis of human papillomavirus type 16 open reading frame E7 immortalizing function in rat embryo fibroblast cells", *J.Gen.Virol.*, vol. 71 (Pt 2), pp. 449-453.

Chow, L. T., Duffy, A. A., Wang, H. K., & Broker, T. R. 2009, "A highly efficient system to produce infectious human papillomavirus: Elucidation of natural virus-host interactions", *Cell Cycle*, vol. 8, no. 9, pp. 1319-1323.

Chow, L. T. T. R. & Broker 2006, "Mechanisms and regulation of papillomavirus DNA replication. *In* M. S. Campo (ed.), Papillomavirus research: from natural history to vaccines and beyond. Caister Academic Press, Norwich, United Kingdom.," pp. 53-71.

Ciccolini, F., Di, P. G., Carlotti, F., Crawford, L., & Tommasino, M. 1994, "Functional studies of E7 proteins from different HPV types", *Oncogene*, vol. 9, no. 9, pp. 2633-2638.

Clifford, G., Franceschi, S., Diaz, M., Munoz, N., & Villa, L. L. 2006, "Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases", *Vaccine*, vol. 24 Suppl 3, pp. S3-26-S3/34.

Clower, R. V., Fisk, J. C., & Melendy, T. 2006a, "Papillomavirus E1 protein binds to and stimulates human topoisomerase I", *J.Virol.*, vol. 80, no. 3, pp. 1584-1587.

Clower, R. V., Hu, Y., & Melendy, T. 2006b, "Papillomavirus E2 protein interacts with and stimulates human topoisomerase I", *Virology*, vol. 348, no. 1, pp. 13-18.

Cohen, P. 2002a, "The origins of protein phosphorylation", *Nat.Cell Biol.*, vol. 4, no. 5, p. E127-E130.

Cohen, P. T. 2002b, "Protein phosphatase 1--targeted in many directions", *J.Cell Sci.*, vol. 115, no. Pt 2, pp. 241-256.

Collins, S. I., Constandinou-Williams, C., Wen, K., Young, L. S., Roberts, S., Murray, P. G., & Woodman, C. B. 2009, "Disruption of the E2 gene is a common and early event in the natural history of cervical human papillomavirus infection: a longitudinal cohort study", *Cancer Res.*, vol. 69, no. 9, pp. 3828-3832.

Conrad, M., Bubb, V. J., & Schlegel, R. 1993, "The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein", *J.Virol.*, vol. 67, no. 10, pp. 6170-6178.

Conway, M. J. & Meyers, C. 2009, "Replication and assembly of human papillomaviruses", *J.Dent.Res.*, vol. 88, no. 4, pp. 307-317.

Correa, P. & Houghton, J. 2007, "Carcinogenesis of Helicobacter pylori", *Gastroenterology*, vol. 133, no. 2, pp. 659-672.

Coverley, D., Laman, H., & Laskey, R. A. 2002, "Distinct roles for cyclins E and A during DNA replication complex assembly and activation", *Nat.Cell Biol.*, vol. 4, no. 7, pp. 523-528.

Crook, T., Fisher, C., Masterson, P. J., & Vousden, K. H. 1994, "Modulation of transcriptional regulatory properties of p53 by HPV E6", *Oncogene*, vol. 9, no. 4, pp. 1225-1230.

Crusius, K., Auvinen, E., Steuer, B., Gaissert, H., & Alonso, A. 1998, "The human papillomavirus type 16 E5-protein modulates ligand-dependent activation of the EGF receptor family in the human epithelial cell line HaCaT", *Exp.Cell Res.*, vol. 241, no. 1, pp. 76-83.

Crusius, K., Rodriguez, I., & Alonso, A. 2000, "The human papillomavirus type 16 E5 protein modulates ERK1/2 and p38 MAP kinase activation by an EGFR-independent process in stressed human keratinocytes", *Virus Genes*, vol. 20, no. 1, pp. 65-69.

Cumming, S. A., Cheun-Im, T., Milligan, S. G., & Graham, S. V. 2008, "Human papillomavirus type 16 late gene expression is regulated by cellular RNA processing factors in response to epithelial differentiation", *Biochem.Soc.Trans.*, vol. 36, no. Pt 3, pp. 522-524.

Cumming, S. A., Chuen-Im, T., Zhang, J., & Graham, S. V. 2009, "The RNA stability regulator HuR regulates L1 protein expression in vivo in differentiating cervical epithelial cells", *Virology*, vol. 383, no. 1, pp. 142-149.

Darnell, G. A., Schroder, W. A., Antalis, T. M., Lambley, E., Major, L., Gardner, J., Birrell, G., Cid-Arregui, A., & Suhrbier, A. 2007, "Human papillomavirus E7 requires the protease calpain to degrade the retinoblastoma protein", *J.Biol.Chem.*, vol. 282, no. 52, pp. 37492-37500.

Davies, R., Hicks, R., Crook, T., Morris, J., & Vousden, K. 1993, "Human papillomavirus type 16 E7 associates with a histone H1 kinase and with p107 through sequences necessary for transformation", *J.Virol.*, vol. 67, no. 5, pp. 2521-2528.

Davy, C. & Doorbar, J. 2007, "G2/M cell cycle arrest in the life cycle of viruses", *Virology*, vol. 368, no. 2, pp. 219-226.

Davy, C., McIntosh, P., Jackson, D. J., Sorathia, R., Miell, M., Wang, Q., Khan, J., Soneji, Y., & Doorbar, J. 2009, "A novel interaction between the human papillomavirus type 16 E2 and E1;E4 proteins leads to stabilization of E2", *Virology*.

Davy, C. E., Ayub, M., Jackson, D. J., Das, P., McIntosh, P., & Doorbar, J. 2006, "HPV16 E1--E4 protein is phosphorylated by Cdk2/cyclin A and relocalizes this complex to the cytoplasm", *Virology*, vol. 349, no. 1, pp. 230-244.

Davy, C. E., Jackson, D. J., Raj, K., Peh, W. L., Southern, S. A., Das, P., Sorathia, R., Laskey, P., Middleton, K., Nakahara, T., Wang, Q., Masterson, P. J., Lambert, P. F., Cuthill, S., Millar, J. B., & Doorbar, J. 2005, "Human papillomavirus type 16 E1 E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes", *J.Virol.*, vol. 79, no. 7, pp. 3998-4011.

Davy, C. E., Jackson, D. J., Wang, Q., Raj, K., Masterson, P. J., Fenner, N. F., Southern, S., Cuthill, S., Millar, J. B., & Doorbar, J. 2002, "Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16", *J.Virol.*, vol. 76, no. 19, pp. 9806-9818.

Day, P. M., Baker, C. C., Lowy, D. R., & Schiller, J. T. 2004, "Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression", *Proc.Natl.Acad.Sci.U.S.A*, vol. 101, no. 39, pp. 14252-14257.

Day, P. M., Lowy, D. R., & Schiller, J. T. 2003, "Papillomaviruses infect cells via a clathrin-dependent pathway", *Virology*, vol. 307, no. 1, pp. 1-11.

Day, P. M., Roden, R. B., Lowy, D. R., & Schiller, J. T. 1998, "The papillomavirus minor capsid protein, L2, induces localization of the major capsid protein, L1, and the viral transcription/replication protein, E2, to PML oncogenic domains", *J.Virol.*, vol. 72, no. 1, pp. 142-150.

Dayaram, T. & Marriott, S. J. 2008, "Effect of transforming viruses on molecular mechanisms associated with cancer", *J.Cell Physiol*, vol. 216, no. 2, pp. 309-314.

de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U., & zur, H. H. 2004, "Classification of papillomaviruses", *Virology*, vol. 324, no. 1, pp. 17-27.

Deng, W., Lin, B. Y., Jin, G., Wheeler, C. G., Ma, T., Harper, J. W., Broker, T. R., & Chow, L. T. 2004, "Cyclin/CDK regulates the nucleocytoplasmic localization of the human papillomavirus E1 DNA helicase", *J.Virol.*, vol. 78, no. 24, pp. 13954-13965.

Desaintes, C. & Demeret, C. 1996, "Control of papillomavirus DNA replication and transcription", *Semin.Cancer Biol.*, vol. 7, no. 6, pp. 339-347.

DiMaio, D. & Mattoon, D. 2001, "Mechanisms of cell transformation by papillomavirus E5 proteins", *Oncogene*, vol. 20, no. 54, pp. 7866-7873.

Doorbar, J., Campbell, D., Grand, R. J., & Gallimore, P. H. 1986, "Identification of the human papilloma virus-1a E4 gene products", *EMBO J.*, vol. 5, no. 2, pp. 355-362.

Doorbar, J., Coneron, I., & Gallimore, P. H. 1989, "Sequence divergence yet conserved physical characteristics among the E4 proteins of cutaneous human papillomaviruses", *Virology*, vol. 172, no. 1, pp. 51-62.

Doorbar, J., Elston, R. C., Napthine, S., Raj, K., Medcalf, E., Jackson, D., Coleman, N., Griffin, H. M., Masterson, P., Stacey, S., Mengistu, Y., & Dunlop, J. 2000, "The E1E4 protein of human papillomavirus type 16 associates with a putative RNA helicase through sequences in its C terminus", *J.Virol.*, vol. 74, no. 21, pp. 10081-10095.

Doorbar, J., Ely, S., Sterling, J., McLean, C., & Crawford, L. 1991, "Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network", *Nature*, vol. 352, no. 6338, pp. 824-827.

Doorbar, J., Evans, H. S., Coneron, I., Crawford, L. V., & Gallimore, P. H. 1988, "Analysis of HPV-1 E4 gene expression using epitope-defined antibodies", *EMBO J.*, vol. 7, no. 3, pp. 825-833.

Doorbar, J., Medcalf, E., & Napthine, S. 1996, "Analysis of HPV1 E4 complexes and their association with keratins in vivo", *Virology*, vol. 218, no. 1, pp. 114-126.

Dostatni, N., Lambert, P. F., Sousa, R., Ham, J., Howley, P. M., & Yaniv, M. 1991, "The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex", *Genes Dev.*, vol. 5, no. 9, pp. 1657-1671.

Dove, B., Brooks, G., Bicknell, K., Wurm, T., & Hiscox, J. A. 2006, "Cell cycle perturbations induced by infection with the coronavirus infectious bronchitis virus and their effect on virus replication", *J.Virol.*, vol. 80, no. 8, pp. 4147-4156.

Duensing, S., Duensing, A., Crum, C. P., & Munger, K. 2001, "Human papillomavirus type 16 E7 oncoprotein-induced abnormal centrosome synthesis is an early event in the evolving malignant phenotype", *Cancer Res.*, vol. 61, no. 6, pp. 2356-2360.

Dyson, N., Howley, P. M., Munger, K., & Harlow, E. 1989, "The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product", *Science*, vol. 243, no. 4893, pp. 934-937.

Egawa, K., Iftner, A., Doorbar, J., Honda, Y., & Iftner, T. 2000, "Synthesis of viral DNA and late capsid protein L1 in parabasal spinous cell layers of naturally occurring benign warts infected with human papillomavirus type 1", *Virology*, vol. 268, no. 2, pp. 281-293.

Everett, R. D. 2001, "DNA viruses and viral proteins that interact with PML nuclear bodies", *Oncogene*, vol. 20, no. 49, pp. 7266-7273.

Fang, L., Budgeon, L. R., Doorbar, J., Briggs, E. R., & Howett, M. K. 2006a, "The human papillomavirus type 11 E1/\E4 protein is not essential for viral genome amplification", *Virology*, vol. 351, no. 2, pp. 271-279.

Fang, L., Meyers, C., Budgeon, L. R., & Howett, M. K. 2006b, "Induction of productive human papillomavirus type 11 life cycle in epithelial cells grown in organotypic raft cultures", *Virology*, vol. 347, no. 1, pp. 28-35.

Fehrmann, F., Klumpp, D. J., & Laimins, L. A. 2003, "Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation", *J.Virol.*, vol. 77, no. 5, pp. 2819-2831.

Feller, L., Wood, N. H., & Lemmer, J. 2007, "HIV-associated Kaposi sarcoma: pathogenic mechanisms", *Oral Surg.Oral Med.Oral Pathol.Oral Radiol.Endod.*, vol. 104, no. 4, pp. 521-529.

Filippova, M., Parkhurst, L., & Duerksen-Hughes, P. J. 2004, "The human papillomavirus 16 E6 protein binds to Fas-associated death domain and protects cells from Fas-triggered apoptosis", *J.Biol.Chem.*, vol. 279, no. 24, pp. 25729-25744.

Filippova, M., Song, H., Connolly, J. L., Dermody, T. S., & Duerksen-Hughes, P. J. 2002, "The human papillomavirus 16 E6 protein binds to tumor necrosis factor (TNF) R1 and protects cells from TNF-induced apoptosis", *J.Biol.Chem.*, vol. 277, no. 24, pp. 21730-21739.

Flores, E. R. & Lambert, P. F. 1997, "Evidence for a switch in the mode of human papillomavirus type 16 DNA replication during the viral life cycle", *J.Virol.*, vol. 71, no. 10, pp. 7167-7179.

Flores, E. R., len-Hoffmann, B. L., Lee, D., Sattler, C. A., & Lambert, P. F. 1999, "Establishment of the human papillomavirus type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line", *Virology*, vol. 262, no. 2, pp. 344-354.

Florin, L., Sapp, C., Streeck, R. E., & Sapp, M. 2002a, "Assembly and translocation of papillomavirus capsid proteins", *J.Virol.*, vol. 76, no. 19, pp. 10009-10014.

Florin, L., Schafer, F., Sotlar, K., Streeck, R. E., & Sapp, M. 2002b, "Reorganization of nuclear domain 10 induced by papillomavirus capsid protein l2", *Virology*, vol. 295, no. 1, pp. 97-107.

Frattini, M. G. & Laimins, L. A. 1994, "Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 26, pp. 12398-12402.

Frattini, M. G., Lim, H. B., & Laimins, L. A. 1996, "In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression", *Proc.Natl.Acad.Sci.U.S.A*, vol. 93, no. 7, pp. 3062-3067.

Fuller-Pace, F. V. 2006, "DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation", *Nucleic Acids Res.*, vol. 34, no. 15, pp. 4206-4215.

Fuller-Pace, F. V. 1994, "RNA helicases: modulators of RNA structure", *Trends Cell Biol.*, vol. 4, no. 8, pp. 271-274.

Funk, J. O., Waga, S., Harry, J. B., Espling, E., Stillman, B., & Galloway, D. A. 1997, "Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein", *Genes Dev.*, vol. 11, no. 16, pp. 2090-2100.

Furuno, N., den, E. N., & Pines, J. 1999, "Human cyclin A is required for mitosis until mid prophase", *J.Cell Biol.*, vol. 147, no. 2, pp. 295-306.

Garnett, T. O., Filippova, M., & Duerksen-Hughes, P. J. 2006, "Accelerated degradation of FADD and procaspase 8 in cells expressing human papilloma virus 16 E6 impairs TRAIL-mediated apoptosis", *Cell Death.Differ.*, vol. 13, no. 11, pp. 1915-1926.

Genovese, N. J., Banerjee, N. S., Broker, T. R., & Chow, L. T. 2008, "Casein kinase II motif-dependent phosphorylation of human papillomavirus E7 protein promotes p130 degradation and S-phase induction in differentiated human keratinocytes", *J.Virol.*, vol. 82, no. 10, pp. 4862-4873.

Gillison, M. L. & Shah, K. V. 2003, "Chapter 9: Role of mucosal human papillomavirus in nongenital cancers", *J.Natl.Cancer Inst.Monogr* no. 31, pp. 57-65.

Giroglou, T., Florin, L., Schafer, F., Streeck, R. E., & Sapp, M. 2001, "Human papillomavirus infection requires cell surface heparan sulfate", *J.Virol.*, vol. 75, no. 3, pp. 1565-1570.

Glaunsinger, B. A., Lee, S. S., Thomas, M., Banks, L., & Javier, R. 2000, "Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins", *Oncogene*, vol. 19, no. 46, pp. 5270-5280.

Goh, W. C., Rogel, M. E., Kinsey, C. M., Michael, S. F., Fultz, P. N., Nowak, M. A., Hahn, B. H., & Emerman, M. 1998, "HIV-1 Vpr increases viral expression by manipulation of the cell cycle: a mechanism for selection of Vpr in vivo", *Nat.Med.*, vol. 4, no. 1, pp. 65-71.

Graham, S. V. 2008, "Papillomavirus 3' UTR regulatory elements", *Front Biosci.*, vol. 13, pp. 5646-5663.

Graham, S. V. 2006, "Late Events in the Life Cycle of Human papillomaviruses. *In* M. S. Campo (ed.), Papillomavirus research: from natural history to vaccines and beyond. Caister Academic Press, Norwich, United Kingdom.," pp. 193-212.

Grand, R. J., Doorbar, J., Smith, K. J., Coneron, I., & Gallimore, P. H. 1989, "Phosphorylation of the human papillomavirus type 1 E4 proteins in vivo and in vitro", *Virology*, vol. 170, no. 1, pp. 201-213.

Graveley, B. R. 2000, "Sorting out the complexity of SR protein functions", *RNA.*, vol. 6, no. 9, pp. 1197-1211.

Green, J., Berrington de, G. A., Sweetland, S., Beral, V., Chilvers, C., Crossley, B., Deacon, J., Hermon, C., Jha, P., Mant, D., Peto, J., Pike, M., & Vessey, M. P. 2003, "Risk factors for adenocarcinoma and squamous cell carcinoma of the cervix in women aged 20-44 years: the UK National Case-Control Study of Cervical Cancer", *Br.J.Cancer*, vol. 89, no. 11, pp. 2078-2086.

Hanks, S. K. & Hunter, T. 1995, "Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification", *FASEB J.*, vol. 9, no. 8, pp. 576-596.

Hawley-Nelson, P., Vousden, K. H., Hubbert, N. L., Lowy, D. R., & Schiller, J. T. 1989, "HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes", *EMBO J.*, vol. 8, no. 12, pp. 3905-3910.

He, W., Staples, D., Smith, C., & Fisher, C. 2003, "Direct activation of cyclin-dependent kinase 2 by human papillomavirus E7", *J.Virol.*, vol. 77, no. 19, pp. 10566-10574.

Heino, P., Zhou, J., & Lambert, P. F. 2000, "Interaction of the papillomavirus transcription/replication factor, E2, and the viral capsid protein, L2", *Virology*, vol. 276, no. 2, pp. 304-314.

Hendrix, S. L. 2008, "Assessing human papillomavirus vaccine efficacy and safety", *J.Am.Osteopath.Assoc.*, vol. 108, no. 4 Suppl 2, p. S8-S12.

Ho, G. Y., Burk, R. D., Klein, S., Kadish, A. S., Chang, C. J., Palan, P., Basu, J., Tachezy, R., Lewis, R., & Romney, S. 1995, "Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia", *J.Natl.Cancer Inst.*, vol. 87, no. 18, pp. 1365-1371.

Holmgren, S. C., Patterson, N. A., Ozbun, M. A., & Lambert, P. F. 2005, "The minor capsid protein L2 contributes to two steps in the human papillomavirus type 31 life cycle", *J.Virol.*, vol. 79, no. 7, pp. 3938-3948.

Houtgraaf, J. H., Versmissen, J., & van der Giessen, W. J. 2006, "A concise review of DNA damage checkpoints and repair in mammalian cells", *Cardiovasc.Revasc.Med.*, vol. 7, no. 3, pp. 165-172.

Howie, H. L., Katzenellenbogen, R. A., & Galloway, D. A. 2009, "Papillomavirus E6 proteins", *Virology*, vol. 384, no. 2, pp. 324-334.

Howley, P. M. & Lowy, D. R. 2009, "Papillomaviruses and their replication. In: D.M. a. H. and P.M. Knipe, Editors, *Fields' Virology*, Lippincott Williams & Wilkins, Philadelphia (2001),".

Hu, T. H., S.C.Ferril, A.M.Snider, & M.S.Barbosa 1995, "In vivo analysis of HPV E7 protein association with pRb, p107 and p130.", *Int.J.Oncol.*, vol. 6, pp. 167-174.

Hughes, F. J. & Romanos, M. A. 1993, "E1 protein of human papillomavirus is a DNA helicase/ATPase", *Nucleic Acids Res.*, vol. 21, no. 25, pp. 5817-5823.

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De, C. E., Langendijk-Genevaux, P. S., Pagni, M., & Sigrist, C. J. 2006, "The PROSITE database", *Nucleic Acids Res.*, vol. 34, no. Database issue, p. D227-D230.

Iost, I. & Dreyfus, M. 1994, "mRNAs can be stabilized by DEAD-box proteins", *Nature*, vol. 372, no. 6502, pp. 193-196.

Johansson, C., Graham, S. V., Dornan, E. S., & Morgan, I. M. 2009, "The human papillomavirus 16 E2 protein is stabilised in S phase", *Virology*.

Jones, D. L., Alani, R. M., & Munger, K. 1997, "The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1-mediated inhibition of cdk2", *Genes Dev.*, vol. 11, no. 16, pp. 2101-2111.

Kadaja, M., Silla, T., Ustav, E., & Ustav, M. 2009, "Papillomavirus DNA replication - from initiation to genomic instability", *Virology*, vol. 384, no. 2, pp. 360-368.

Kahn, J. A. 2009, "HPV vaccination for the prevention of cervical intraepithelial neoplasia", *N.Engl.J.Med.*, vol. 361, no. 3, pp. 271-278.

Kalantari, M. & Bernard, H. U. 2006, "Gene Expression of Papillomaviruses. *In* M. S. Campo (ed.), Papillomavirus research: from natural history to vaccines and beyond. Caister Academic Press, Norwich, United Kingdom," pp. 41-52.

Kim, K. Y. & Taylor, M. W. 2003, "Identification of a novel promoter in the E2 open reading frame of the human papillomavirus type 18 genome", *J.Med.Virol.*, vol. 69, no. 1, pp. 122-131.

Kiyono, T., Hiraiwa, A., Fujita, M., Hayashi, Y., Akiyama, T., & Ishibashi, M. 1997, "Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein", *Proc.Natl.Acad.Sci.U.S.A*, vol. 94, no. 21, pp. 11612-11616.

Klingelhutz, A. J., Foster, S. A., & McDougall, J. K. 1996, "Telomerase activation by the E6 gene product of human papillomavirus type 16", *Nature*, vol. 380, no. 6569, pp. 79-82.

Klumpp, D. J. & Laimins, L. A. 1999, "Differentiation-induced changes in promoter usage for transcripts encoding the human papillomavirus type 31 replication protein E1", *Virology*, vol. 257, no. 1, pp. 239-246.

Knapp, A. A., McManus, P. M., Bockstall, K., & Moroianu, J. 2009, "Identification of the nuclear localization and export signals of high risk HPV16 E7 oncoprotein", *Virology*, vol. 383, no. 1, pp. 60-68.

Knight, G. L., Grainger, J. R., Gallimore, P. H., & Roberts, S. 2004, "Cooperation between different forms of the human papillomavirus type 1 E4 protein to block cell cycle progression and cellular DNA synthesis", *J.Virol.*, vol. 78, no. 24, pp. 13920-13933.

Knight, G. L., Turnell, A. S., & Roberts, S. 2006, "Role for Wee1 in inhibition of G2-to-M transition through the cooperation of distinct human papillomavirus type 1 E4 proteins", *J.Virol.*, vol. 80, no. 15, pp. 7416-7426.

Kreider, J. W., Howett, M. K., Leure-Dupree, A. E., Zaino, R. J., & Weber, J. A. 1987, "Laboratory production in vivo of infectious human papillomavirus type 11", *J.Virol.*, vol. 61, no. 2, pp. 590-593.

Kuhne, C., Gardiol, D., Guarnaccia, C., Amenitsch, H., & Banks, L. 2000, "Differential regulation of human papillomavirus E6 by protein kinase A: conditional degradation of human discs large protein by oncogenic E6", *Oncogene*, vol. 19, no. 51, pp. 5884-5891.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. 2007, "Clustal W and Clustal X version 2.0", *Bioinformatics.*, vol. 23, no. 21, pp. 2947-2948.

Laughlin-Drubin, M. E. & Munger, K. 2009, "The human papillomavirus E7 oncoprotein", *Virology*, vol. 384, no. 2, pp. 335-344.

Lechner, M. S. & Laimins, L. A. 1994, "Inhibition of p53 DNA binding by human papillomavirus E6 proteins", *J.Virol.*, vol. 68, no. 7, pp. 4262-4273.

Lee, S. S., Glaunsinger, B., Mantovani, F., Banks, L., & Javier, R. T. 2000, "Multi-PDZ domain protein MUPP1 is a cellular target for both adenovirus E4-ORF1 and high-risk papillomavirus type 18 E6 oncoproteins", *J.Virol.*, vol. 74, no. 20, pp. 9680-9693.

Liao, J. B. 2006, "Viruses and human cancer", *Yale J.Biol.Med.*, vol. 79, no. 3-4, pp. 115-122.

Liu, X., Dakic, A., Zhang, Y., Dai, Y., Chen, R., & Schlegel, R. 2009, "HPV E6 protein interacts physically and functionally with the cellular telomerase complex", *Proc.Natl.Acad.Sci.U.S.A*, vol. 106, no. 44, pp. 18780-18785.

Longworth, M. S. & Laimins, L. A. 2004, "Pathogenesis of human papillomaviruses in differentiating epithelia", *Microbiol.Mol.Biol.Rev.*, vol. 68, no. 2, pp. 362-372.

Longworth, M. S., Wilson, R., & Laimins, L. A. 2005, "HPV31 E7 facilitates replication by activating E2F2 transcription through its interaction with HDACs", *EMBO J.*, vol. 24, no. 10, pp. 1821-1830.

Loo, Y. M. & Melendy, T. 2004, "Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA", *J.Virol.*, vol. 78, no. 4, pp. 1605-1615.

Ma, T., Zou, N., Lin, B. Y., Chow, L. T., & Harper, J. W. 1999, "Interaction between cyclin-dependent kinases and human papillomavirus replication-initiation protein E1 is required for efficient viral replication", *Proc.Natl.Acad.Sci.U.S.A*, vol. 96, no. 2, pp. 382-387.

Maldonado, E., Cabrejos, M. E., Banks, L., & Allende, J. E. 2002, "Human papillomavirus-16 E7 protein inhibits the DNA interaction of the TATA binding transcription factor", *J.Cell Biochem.*, vol. 85, no. 4, pp. 663-669.

Mammone, T., Marenus, K., Maes, D., & Lockshin, R. A. 1998, "The induction of terminal differentiation markers by the cAMP pathway in human HaCaT keratinocytes", *Skin Pharmacol.Appl.Skin Physiol*, vol. 11, no. 3, pp. 152-160.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. 2002, "The protein kinase complement of the human genome", *Science*, vol. 298, no. 5600, pp. 1912-1934.

Mantovani, F. & Banks, L. 2001, "The human papillomavirus E6 protein and its contribution to malignant progression", *Oncogene*, vol. 20, no. 54, pp. 7874-7887.

Massimi, P., Pim, D., Storey, A., & Banks, L. 1996, "HPV-16 E7 and adenovirus E1a complex formation with TATA box binding protein is enhanced by casein kinase II phosphorylation", *Oncogene*, vol. 12, no. 11, pp. 2325-2330.

Masterson, P. J., Stanley, M. A., Lewis, A. P., & Romanos, M. A. 1998, "A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha-primase p68 subunit", *J.Virol.*, vol. 72, no. 9, pp. 7407-7419.

Maul, G. G., Negorev, D., Bell, P., & Ishov, A. M. 2000, "Review: properties and assembly mechanisms of ND10, PML bodies, or PODs", *J.Struct.Biol.*, vol. 129, no. 2-3, pp. 278-287.

McCance, D. J., Kopan, R., Fuchs, E., & Laimins, L. A. 1988, "Human papillomavirus type 16 alters human epithelial cell differentiation in vitro", *Proc.Natl.Acad.Sci.U.S.A*, vol. 85, no. 19, pp. 7169-7173.

McIntosh, P. B., Martin, S. R., Jackson, D. J., Khan, J., Isaacson, E. R., Calder, L., Raj, K., Griffin, H. M., Wang, Q., Laskey, P., Eccleston, J. F., & Doorbar, J. 2008, "Structural analysis reveals an amyloid form of the human papillomavirus type 16 E1--E4 protein and provides a molecular basis for its accumulation", *J.Virol.*, vol. 82, no. 16, pp. 8196-8203.

McIntyre, M. C., Ruesch, M. N., & Laimins, L. A. 1996, "Human papillomavirus E7 oncoproteins bind a single form of cyclin E in a complex with cdk2 and p107", *Virology*, vol. 215, no. 1, pp. 73-82.

Meyers, C., Frattini, M. G., Hudson, J. B., & Laimins, L. A. 1992, "Biosynthesis of human papillomavirus from a continuous cell line upon epithelial differentiation", *Science*, vol. 257, no. 5072, pp. 971-973.

Meyers, C. & Laimins, L. A. 1994, "In vitro systems for the study and propagation of human papillomaviruses", *Curr.Top.Microbiol.Immunol.*, vol. 186, pp. 199-215.

Meyers, C., Mayer, T. J., & Ozbun, M. A. 1997, "Synthesis of infectious human papillomavirus type 18 in differentiating epithelium transfected with viral DNA", *J.Virol.*, vol. 71, no. 10, pp. 7381-7386.

Milligan, S. G., Veerapraditsin, T., Ahamet, B., Mole, S., & Graham, S. V. 2007, "Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells", *Virology*, vol. 360, no. 1, pp. 172-181.

Morgan, I. M. & Donaldson, M. M. 2006, "The Papillomavirus Transcription/Replication Factor E2: Structure, Function, Cancer and Therapy *In* M. S. Campo (ed.), Papillomavirus research: from natural history to vaccines and beyond. Caister Academic Press, Norwich, United Kingdom.," pp. 73-82.

Munger, K. & Howley, P. M. 2002, "Human papillomavirus immortalization and transformation functions", *Virus Res.*, vol. 89, no. 2, pp. 213-228.

Munger, K., Werness, B. A., Dyson, N., Phelps, W. C., Harlow, E., & Howley, P. M. 1989, "Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product", *EMBO J.*, vol. 8, no. 13, pp. 4099-4105.

Munoz, N., Bosch, F. X., de, S. S., Herrero, R., Castellsague, X., Shah, K. V., Snijders, P. J., & Meijer, C. J. 2003, "Epidemiologic classification of human papillomavirus types associated with cervical cancer", *N.Engl.J.Med.*, vol. 348, no. 6, pp. 518-527.

Nakagawa, S. & Huibregtse, J. M. 2000, "Human scribble (Vartul) is targeted for ubiquitin-mediated degradation by the high-risk papillomavirus E6 proteins and the E6AP ubiquitin-protein ligase", *Mol.Cell Biol.*, vol. 20, no. 21, pp. 8244-8253.

Nakahara, T., Nishimura, A., Tanaka, M., Ueno, T., Ishimoto, A., & Sakai, H. 2002, "Modulation of the cell division cycle by human papillomavirus type 18 E4", *J.Virol.*, vol. 76, no. 21, pp. 10914-10920.

Nakahara, T., Peh, W. L., Doorbar, J., Lee, D., & Lambert, P. F. 2005, "Human papillomavirus type 16 E1circumflexE4 contributes to multiple facets of the papillomavirus life cycle", *J.Virol.*, vol. 79, no. 20, pp. 13150-13165.

Nasseri, M., Hirochika, R., Broker, T. R., & Chow, L. T. 1987, "A human papilloma virus type 11 transcript encoding an E1--E4 protein", *Virology*, vol. 159, no. 2, pp. 433-439.

Nguyen, C. L. & Munger, K. 2008, "Direct association of the HPV16 E7 oncoprotein with cyclin A/CDK2 and cyclin E/CDK2 complexes", *Virology*, vol. 380, no. 1, pp. 21-25.

Oh, S. T., Kyo, S., & Laimins, L. A. 2001, "Telomerase activation by human papillomavirus type 16 E6 protein: induction of human telomerase reverse transcriptase expression through Myc and GC-rich Sp1 binding sites", *J.Virol.*, vol. 75, no. 12, pp. 5559-5566.

Ozbun, M. A. & Meyers, C. 1998, "Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b", *J.Virol.*, vol. 72, no. 4, pp. 2715-2722.

Palefsky, J. 2009, "Human papillomavirus-related disease in people with HIV", *Curr.Opin.HIV.AIDS*, vol. 4, no. 1, pp. 52-56.

Park, P., Copeland, W., Yang, L., Wang, T., Botchan, M. R., & Mohr, I. J. 1994, "The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase", *Proc.Natl.Acad.Sci.U.S.A*, vol. 91, no. 18, pp. 8700-8704.

Parkin, D. M. & Bray, F. 2006, "Chapter 2: The burden of HPV-related cancers", *Vaccine*, vol. 24 Suppl 3, pp. S3-11-S3/25.

Patel, D., Huang, S. M., Baglia, L. A., & McCance, D. J. 1999, "The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300", *EMBO J.*, vol. 18, no. 18, pp. 5061-5072.

Peh, W. L., Brandsma, J. L., Christensen, N. D., Cladel, N. M., Wu, X., & Doorbar, J. 2004, "The viral E4 protein is required for the completion of the cottontail rabbit papillomavirus productive cycle in vivo", *J.Virol.*, vol. 78, no. 4, pp. 2142-2151.

Peh, W. L., Middleton, K., Christensen, N., Nicholls, P., Egawa, K., Sotlar, K., Brandsma, J., Percival, A., Lewis, J., Liu, W. J., & Doorbar, J. 2002, "Life cycle heterogeneity in animal models of human papillomavirus-associated disease", *J.Virol.*, vol. 76, no. 20, pp. 10401-10416.

Pereira, R., Hitzeroth, I. I., & Rybicki, E. P. 2009, "Insights into the role and function of L2, the minor capsid protein of papillomaviruses", *Arch.Virol.*, vol. 154, no. 2, pp. 187-197.

Pett, M. R., Herdman, M. T., Palmer, R. D., Yeo, G. S., Shivji, M. K., Stanley, M. A., & Coleman, N. 2006, "Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response", *Proc.Natl.Acad.Sci.U.S.A*, vol. 103, no. 10, pp. 3822-3827.

Pirog, E. C., Kleter, B., Olgac, S., Bobkiewicz, P., Lindeman, J., Quint, W. G., Richart, R. M., & Isacson, C. 2000, "Prevalence of human papillomavirus DNA in different histological subtypes of cervical adenocarcinoma", *Am.J.Pathol.*, vol. 157, no. 4, pp. 1055-1062.

Poddar, A., Reed, S. C., McPhillips, M. G., Spindler, J. E., & McBride, A. A. 2009, "The human papillomavirus type 8 E2 tethering protein targets the ribosomal DNA loci of host mitotic chromosomes", *J.Virol.*, vol. 83, no. 2, pp. 640-650.

Pray, T. R. & Laimins, L. A. 1995, "Differentiation-dependent expression of E1--E4 proteins in cell lines maintaining episomes of human papillomavirus type 31b", *Virology*, vol. 206, no. 1, pp. 679-685.

Raj, K., Berguerand, S., Southern, S., Doorbar, J., & Beard, P. 2004, "E1 empty set E4 protein of human papillomavirus type 16 associates with mitochondria", *J.Virol.*, vol. 78, no. 13, pp. 7199-7207.

Regad, T. & Chelbi-Alix, M. K. 2001, "Role and fate of PML nuclear bodies in response to interferon and viral infections", *Oncogene*, vol. 20, no. 49, pp. 7274-7286.

Rezk, S. A. & Weiss, L. M. 2007, "Epstein-Barr virus-associated lymphoproliferative disorders", *Hum.Pathol.*, vol. 38, no. 9, pp. 1293-1304.

Richards, R. M., Lowy, D. R., Schiller, J. T., & Day, P. M. 2006, "Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection", *Proc.Natl.Acad.Sci.U.S.A*, vol. 103, no. 5, pp. 1522-1527.

Roberts, S. 2006, "The E4 protein - A late starter. *In* M. S. Campo (ed.), Papillomavirus research: from natural history to vaccines and beyond. Caister Academic Press, Norwich, United Kingdom.," pp. 83-96.

Roberts, S., Ashmole, I., Gibson, L. J., Rookes, S. M., Barton, G. J., & Gallimore, P. H. 1994, "Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells", *J.Virol.*, vol. 68, no. 10, pp. 6432-6445.

Roberts, S., Ashmole, I., Johnson, G. D., Kreider, J. W., & Gallimore, P. H. 1993, "Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells", *Virology*, vol. 197, no. 1, pp. 176-187.

Roberts, S., Ashmole, I., Rookes, S. M., & Gallimore, P. H. 1997, "Mutational analysis of the human papillomavirus type 16 E1--E4 protein shows that the C terminus is dispensable for keratin cytoskeleton association but is involved in inducing disruption of the keratin filaments", *J.Virol.*, vol. 71, no. 5, pp. 3554-3562.

Roberts, S., Hillman, M. L., Knight, G. L., & Gallimore, P. H. 2003, "The ND10 component promyelocytic leukemia protein relocates to human papillomavirus type 1 E4 intranuclear inclusion bodies in cultured keratinocytes and in warts", *J.Virol.*, vol. 77, no. 1, pp. 673-684.

Roberts, S., Kingsbury, S. R., Stoeber, K., Knight, G. L., Gallimore, P. H., & Williams, G. H. 2008, "Identification of an arginine-rich motif in human papillomavirus type 1 E1;E4 protein necessary for E4-mediated inhibition of cellular DNA synthesis in vitro and in cells", *J.Virol.*, vol. 82, no. 18, pp. 9056-9064.

Rogel-Gaillard, C., Breitburd, F., & Orth, G. 1992, "Human papillomavirus type 1 E4 proteins differing by their N-terminal ends have distinct cellular localizations when transiently expressed in vitro", *J.Virol.*, vol. 66, no. 2, pp. 816-823.

Rogel-Gaillard, C., Pehau-Arnaudet, G., Breitburd, F., & Orth, G. 1993, "Cytopathic effect in human papillomavirus type 1-induced inclusion warts: in vitro analysis of the contribution of two forms of the viral E4 protein", *J.Invest Dermatol.*, vol. 101, no. 6, pp. 843-851.

Ruesch, M. N. & Laimins, L. A. 1997, "Initiation of DNA synthesis by human papillomavirus E7 oncoproteins is resistant to p21-mediated inhibition of cyclin E-cdk2 activity", *J.Virol.*, vol. 71, no. 7, pp. 5570-5578.

Ruesch, M. N., Stubenrauch, F., & Laimins, L. A. 1998, "Activation of papillomavirus late gene transcription and genome amplification upon differentiation in semisolid medium is coincident with expression of involucrin and transglutaminase but not keratin-10", *J.Virol.*, vol. 72, no. 6, pp. 5016-5024.

Rush, M., Zhao, X., & Schwartz, S. 2005, "A splicing enhancer in the E4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression", *J.Virol.*, vol. 79, no. 18, pp. 12002-12015.

Scheffner, M., Huibregtse, J. M., Vierstra, R. D., & Howley, P. M. 1993, "The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53", *Cell*, vol. 75, no. 3, pp. 495-505.

Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J., & Howley, P. M. 1990, "The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53", *Cell*, vol. 63, no. 6, pp. 1129-1136.

Schiffman, M., Castle, P. E., Jeronimo, J., Rodriguez, A. C., & Wacholder, S. 2007, "Human papillomavirus and cervical cancer", *Lancet*, vol. 370, no. 9590, pp. 890-907.

Schwartz, D. & Gygi, S. P. 2005, "An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets", *Nat.Biotechnol.*, vol. 23, no. 11, pp. 1391-1398.

Shafti-Keramat, S., Handisurya, A., Kriehuber, E., Meneguzzi, G., Slupetzky, K., & Kirnbauer, R. 2003, "Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses", *J.Virol.*, vol. 77, no. 24, pp. 13125-13135.

Sherr, C. J. & McCormick, F. 2002, "The RB and p53 pathways in cancer", *Cancer Cell*, vol. 2, no. 2, pp. 103-112.

Shin, M. K., Balsitis, S., Brake, T., & Lambert, P. F. 2009, "Human papillomavirus E7 oncoprotein overrides the tumor suppressor activity of p21Cip1 in cervical carcinogenesis", *Cancer Res.*, vol. 69, no. 14, pp. 5656-5663.

Solomon, D., Davey, D., Kurman, R., Moriary, A., O'Connor, D., Prey, M., Raab, S., Sherman, M., Wilbur, D., Wright, T., & Young, N. 2002, "The 2001 Bethesda System - Terminology for reporting results of cervical cytology", *Jama-Journal of the American Medical Association*, vol. 287, no. 16, pp. 2114-2119.

Spink, K. M. & Laimins, L. A. 2005, "Induction of the human papillomavirus type 31 late promoter requires differentiation but not DNA amplification", *J.Virol.*, vol. 79, no. 8, pp. 4918-4926.

Stanley, M. A. 2001, "Human papillomavirus and cervical carcinogenesis", *Best.Pract.Res.Clin.Obstet.Gynaecol.*, vol. 15, no. 5, pp. 663-676.

Stanley, M. A., Browne, H. M., Appleby, M., & Minson, A. C. 1989, "Properties of a non-tumorigenic human cervical keratinocyte cell line", *Int.J.Cancer*, vol. 43, no. 4, pp. 672-676.

Steger, G. & Corbach, S. 1997, "Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein", *J.Virol.*, vol. 71, no. 1, pp. 50-58.

Stenlund, A. 2003, "E1 initiator DNA binding specificity is unmasked by selective inhibition of non-specific DNA binding", *EMBO J.*, vol. 22, no. 4, pp. 954-963.

Storey, A., Pim, D., Murray, A., Osborn, K., Banks, L., & Crawford, L. 1988, "Comparison of the in vitro transforming activities of human papillomavirus types", *EMBO J.*, vol. 7, no. 6, pp. 1815-1820.

Straight, S. W., Herman, B., & McCance, D. J. 1995, "The E5 oncoprotein of human papillomavirus type 16 inhibits the acidification of endosomes in human keratinocytes", *J.Virol.*, vol. 69, no. 5, pp. 3185-3192.

Straight, S. W., Hinkle, P. M., Jewers, R. J., & McCance, D. J. 1993, "The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes", *J.Virol.*, vol. 67, no. 8, pp. 4521-4532.

Stubenrauch, F. & Laimins, L. A. 1999, "Human papillomavirus life cycle: active and latent phases", *Semin.Cancer Biol.*, vol. 9, no. 6, pp. 379-386.

Suzich, J. A., Ghim, S. J., Palmer-Hill, F. J., White, W. I., Tamura, J. K., Bell, J. A., Newsome, J. A., Jenson, A. B., & Schlegel, R. 1995, "Systemic immunization with papillomavirus L1 protein completely prevents the development of viral mucosal papillomas", *Proc.Natl.Acad.Sci.U.S.A*, vol. 92, no. 25, pp. 11553-11557.

Sweet, S. M., Creese, A. J., & Cooper, H. J. 2006, "Strategy for the identification of sites of phosphorylation in proteins: neutral loss triggered electron capture dissociation", *Anal.Chem.*, vol. 78, no. 21, pp. 7563-7569.

Swindle, C. S., Zou, N., Van Tine, B. A., Shaw, G. M., Engler, J. A., & Chow, L. T. 1999, "Human papillomavirus DNA replication compartments in a transient DNA replication system", *J.Virol.*, vol. 73, no. 2, pp. 1001-1009.

Tan, S. H., Leong, L. E., Walker, P. A., & Bernard, H. U. 1994, "The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID", *J.Virol.*, vol. 68, no. 10, pp. 6411-6420.

Thierry, F. 2009, "Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma", *Virology*, vol. 384, no. 2, pp. 375-379.

Thomas, J. T., Oh, S. T., Terhune, S. S., & Laimins, L. A. 2001, "Cellular changes induced by low-risk human papillomavirus type 11 in keratinocytes that stably maintain viral episomes", *J.Virol.*, vol. 75, no. 16, pp. 7564-7571.

Thomas, M. & Banks, L. 1998, "Inhibition of Bak-induced apoptosis by HPV-18 E6", *Oncogene*, vol. 17, no. 23, pp. 2943-2954.

Thomas, M. & Banks, L. 1999, "Human papillomavirus (HPV) E6 interactions with Bak are conserved amongst E6 proteins from high and low risk HPV types", *J.Gen.Virol.*, vol. 80 (Pt 6), pp. 1513-1517.

Thomas, M., Laura, R., Hepner, K., Guccione, E., Sawyers, C., Lasky, L., & Banks, L. 2002, "Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation", *Oncogene*, vol. 21, no. 33, pp. 5088-5096.

Thomas, M., Narayan, N., Pim, D., Tomaic, V., Massimi, P., Nagasaka, K., Kranjec, C., Gammoh, N., & Banks, L. 2008, "Human papillomaviruses, cervical cancer and cell polarity", *Oncogene*, vol. 27, no. 55, pp. 7018-7030.

Thomas, M., Pim, D., & Banks, L. 2006, "The Role of the HPV16 E6 Oncoprotein in Malignant Progression *In* M. S. Campo (ed.), Papillomavirus research: from natural history to vaccines and beyond. Caister Academic Press, Norwich, United Kingdom.,".

Thomsen, P., van, D. B., Norrild, B., & Kayser, L. 2000, "The HPV16 E5 oncogene inhibits endocytic trafficking", *Oncogene*, vol. 19, no. 52, pp. 6023-6032.

Tomaic, V., Pim, D., & Banks, L. 2009, "The stability of the human papillomavirus E6 oncoprotein is E6AP dependent", *Virology*.

Tomita, Y., Fuse, A., Sekine, H., Shirasawa, H., Simizu, B., Sugimoto, M., & Funahashi, S. 1991, "Human papillomavirus type 6 and 11 E4 gene products in condyloma acuminata", *J.Gen.Virol.*, vol. 72 (Pt 3), pp. 731-734.

Tommasino, M., Adamczewski, J. P., Carlotti, F., Barth, C. F., Manetti, R., Contorni, M., Cavalieri, F., Hunt, T., & Crawford, L. 1993, "HPV16 E7 protein associates with the protein kinase p33CDK2 and cyclin A", *Oncogene*, vol. 8, no. 1, pp. 195-202.

Tran, H., Nourse, J., Hall, S., Green, M., Griffiths, L., & Gandhi, M. K. 2008, "Immunodeficiency-associated lymphomas", *Blood Rev.*, vol. 22, no. 5, pp. 261-281.

Tsai, T. C. & Chen, S. L. 2003, "The biochemical and biological functions of human papillomavirus type 16 E5 protein", *Arch.Virol.*, vol. 148, no. 8, pp. 1445-1453.

Van Tine, B. A., Dao, L. D., Wu, S. Y., Sonbuchner, T. M., Lin, B. Y., Zou, N., Chiang, C. M., Broker, T. R., & Chow, L. T. 2004, "Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning", *Proc.Natl.Acad.Sci.U.S.A*, vol. 101, no. 12, pp. 4030-4035.

Wang, Q., Kennedy, A., Das, P., McIntosh, P. B., Howell, S. A., Isaacson, E. R., Hinz, S. A., Davy, C., & Doorbar, J. 2009, "Phosphorylation of the human papillomavirus type 16 E1--E4 protein at T57 by ERK triggers a structural change that enhances keratin binding and protein stability", *J.Virol.*, vol. 83, no. 8, pp. 3668-3683.

Watson, R. A., Rollason, T. P., Reynolds, G. M., Murray, P. G., Banks, L., & Roberts, S. 2002, "Changes in expression of the human homologue of the Drosophila discs large tumour suppressor protein in high-grade premalignant cervical neoplasias", *Carcinogenesis*, vol. 23, no. 11, pp. 1791-1796.

Watson, R. A., Thomas, M., Banks, L., & Roberts, S. 2003, "Activity of the human papillomavirus E6 PDZ-binding motif correlates with an enhanced morphological

transformation of immortalized human keratinocytes", *J.Cell Sci.*, vol. 116, no. Pt 24, pp. 4925-4934.

Werness, B. A., Levine, A. J., & Howley, P. M. 1990, "Association of human papillomavirus types 16 and 18 E6 proteins with p53", *Science*, vol. 248, no. 4951, pp. 76-79.

Wilson, R., Fehrmann, F., & Laimins, L. A. 2005, "Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31", *J.Virol.*, vol. 79, no. 11, pp. 6732-6740.

Wilson, R. & Laimins, L. A. 2005, "Differentiation of HPV-containing cells using organotypic "raft" culture or methylcellulose", *Methods Mol.Med.*, vol. 119, pp. 157-169.

Wilson, R., Ryan, G. B., Knight, G. L., Laimins, L. A., & Roberts, S. 2007, "The full-length E1E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression", *Virology*, vol. 362, no. 2, pp. 453-460.

Wilson, V. G., West, M., Woytek, K., & Rangasamy, D. 2002, "Papillomavirus E1 proteins: form, function, and features", *Virus Genes*, vol. 24, no. 3, pp. 275-290.

World Health Organization 2002, "International Agency for Research on Cancer, GLOBOCAN database 2002.", *CANCERMondial [online], http://www-dep.iarc.fr/ (2002).*

You, J., Croyle, J. L., Nishimura, A., Ozato, K., & Howley, P. M. 2004, "Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes", *Cell*, vol. 117, no. 3, pp. 349-360.

Yu, J. H., Lin, B. Y., Deng, W., Broker, T. R., & Chow, L. T. 2007, "Mitogen-activated protein kinases activate the nuclear localization sequence of human papillomavirus type 11 E1 DNA helicase to promote efficient nuclear import", *J.Virol.*, vol. 81, no. 10, pp. 5066-5078.

Zanardi, T. A., Stanley, C. M., Saville, B. M., Spacek, S. M., & Lentz, M. R. 1997, "Modulation of bovine papillomavirus DNA replication by phosphorylation of the viral E1 protein", *Virology*, vol. 228, no. 1, pp. 1-10.

Zerfass-Thome, K., Zwerschke, W., Mannhardt, B., Tindle, R., Botz, J. W., & Jansen-Durr, P. 1996, "Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein", *Oncogene*, vol. 13, no. 11, pp. 2323-2330.

Zhang, B., Spandau, D. F., & Roman, A. 2002, "E5 protein of human papillomavirus type 16 protects human foreskin keratinocytes from UV B-irradiation-induced apoptosis", *J.Virol.*, vol. 76, no. 1, pp. 220-231.

Zhao, K. N., Hengst, K., Liu, W. J., Liu, Y. H., Liu, X. S., McMillan, N. A., & Frazer, I. H. 2000, "BPV1 E2 protein enhances packaging of full-length plasmid DNA in BPV1 pseudovirions", *Virology*, vol. 272, no. 2, pp. 382-393.

Zhou, J., Liu, W. J., Peng, S. W., Sun, X. Y., & Frazer, I. 1999, "Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability", *J.Virol.*, vol. 73, no. 6, pp. 4972-4982.

Zhu, Q. 1999, "Evolutionary relationships among seventeen human papillomavirus genotypes", *Chin Med.Sci.J.*, vol. 14, no. 2, pp. 85-95.

Zimmermann, H., Degenkolbe, R., Bernard, H. U., & O'Connor, M. J. 1999, "The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300", *J.Virol.*, vol. 73, no. 8, pp. 6209-6219.

Zou, N., Lin, B. Y., Duan, F., Lee, K. Y., Jin, G., Guan, R., Yao, G., Lefkowitz, E. J., Broker, T. R., & Chow, L. T. 2000, "The hinge of the human papillomavirus type 11 E2 protein contains major determinants for nuclear localization and nuclear matrix association", *J.Virol.*, vol. 74, no. 8, pp. 3761-3770.

zur Hausen H. 2009, "The search for infectious causes of human cancers: where and why", *Virology*, vol. 392, no. 1, pp. 1-10.