A STUDY OF PDZ PROTEIN EXPRESSION DURING
THE HUMAN PAPILLOMAVIRUS LIFE CYCLE

by

CLAIRe DeBORAh JAmES

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
University of Birmingham
April 2015
This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.
ABSTRACT

The E6 protein of oncogenic α-human papillomaviruses encodes a conserved C-terminal PDZ binding motif (PBM). Conservation of an overlapping kinase recognition site has shown that phosphorylation negatively regulates the E6:PDZ interaction. Many E6 PDZ targets are associated with signalling complexes that regulate cell proliferation and polarity. This study used primary keratinocyte-based models of HPV16 and 18 life cycles to investigate the expression of PDZ targets and the functional role of PBM phosphorylation.

The data indicate that changes in total levels of major E6 PDZ targets are not associated with E6-PBM activity. Interestingly, transcription profiles of E6-PBM targets DLG1, PATJ and PTPN13 are dramatically changed in the presence of HPV16 and 18, whilst others remain unaffected. Further analysis of transcriptional changes of DLG1 revealed upregulation of specific alternatively spliced isoforms, including a novel isoform containing an exon previously thought to be intronic.

This investigation revealed that phosphorylation of the PBM is linked to oncoprotein stability, presenting a potential regulatory mechanism of E6 PDZ interactions during the virus life cycle. Together, these data offer interesting new perspectives on interactions between oncogenic HPV types and PDZ domain-containing targets and indicate that deregulation of their function by the virus may occur through multiple mechanisms.
ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr Sally Roberts for her supervision during this project, and to the members of the Roberts lab, Margaret, Claire, Pete and Elizabeth, for their advice and support. I am especially grateful to Dr Sarah Leonard for her guidance throughout this project. Furthermore, this studentship was made possible by funding from the MRC and Wellcome Trust.

Special thanks are extended to the School of Cancers Sciences, to my colleagues in the lab and office for their encouragement and camaraderie. Thanks also to those in Suffolk and Birmingham, whose friendship has been and continues to be invaluable. I am especially thankful to Corisande Ladies, Alesha, Annika and Kate, who provided much inspiration and entertainment over the last four years.

Lastly, but by no means least, I am hugely grateful for my family, who have always been unquestioningly encouraging and loving.
3.2 Hypothesis and Aims.............................................................................................................. 100
3.3 Results................................................................................................................................. 101
  3.3.1 Generation of cell-based models of the HPV16 and HPV18 life cycle........... 101
  3.3.2 Transcription of E6 PDZ targets in keratinocytes containing high-risk a-
  papillomavirus genomes........................................................................................................ 102
  3.3.3 Analysis of PDZ protein expression during the HPV life cycle ......................... 109
  3.3.4 Proteasomal inhibition of primary keratinocytes.................................................. 119
3.4 Discussion.......................................................................................................................... 122
  3.4.1 Transcriptional upregulation in the presence of HPV genomes......................... 122
  3.4.2 Protein distribution in HPV genome containing cells....................................... 124
  3.4.3 Future directions.......................................................................................................... 130

CHAPTER 4: IDENTIFICATION AND CHARACTERISATION OF A NOVEL
ISOFORM OF DLG1 ......................................................................................................................... 132
4.1 Introduction.......................................................................................................................... 132
  4.1.1 DLG1 protein structure ............................................................................................. 132
  4.1.2 DLG1 expression ........................................................................................................ 139
4.3 Hypothesis and objectives ............................................................................................... 141
4.4 Results................................................................................................................................ 142
  4.4.1 Identification of a novel insertion in DLG1 in mRNA........................................... 142
  4.4.2 Investigation into protein expression of the novel DLG1 isoform....................... 146
  4.4.3 Investigation of DLG1 isoform expression............................................................... 150
4.5 Discussion............................................................................................................................ 157
  4.5.1 Identification of a novel insertion............................................................................. 157
  4.5.2 Consequences of I6 inclusion on DLG1 function.................................................... 159
  4.5.3 Differential expression of DLG1 isoforms............................................................... 163
  4.5.4 Future directions........................................................................................................ 165

CHAPTER 5: REGULATION OF THE HUMAN PAPILLOMAVIRUS E6 PDZ
BINDING MOTIF DURING THE VIRUS LIFE CYCLE................................................................. 167
5.1 Introduction.......................................................................................................................... 167
  5.1.1 Cellular PDZ interactions ....................................................................................... 167
  5.1.2 Viral PBM regulation ............................................................................................. 169
5.3 Hypothesis and objectives ............................................................................................... 174
5.4 Results................................................................................................................................ 175
  5.4.1 Computational screening predicts phosphorylation of a-HPV E6 PBMs.............. 175
  5.4.2 Investigation into the effect of PBM phosphorylation upon E6......................... 176
  5.4.3 14-3-3z...................................................................................................................... 181
5.5 Discussion............................................................................................................................ 182

CHAPTER 6: FINAL DISCUSSION ............................................................................................... 187
6.1 Overview of findings ........................................................................................................... 187
6.2 Future directions ................................................................................................................ 189
6.3 Final statement.................................................................................................................... 190

REFERENCES ............................................................................................................................ 192

APPENDIX I: PUBLICATIONS

APPENDIX II: PRIMER OPTIMISATION AND SEQUENCE ALIGNMENTS

APPENDIX III: CELL CULTURE REAGENTS AND STOCK SOLUTIONS
LIST OF FIGURES

Page numbers represent the preceding page

Chapter 1

Figure 1.1: Human papillomavirus classification 1
Figure 1.2: Stages of cervical intraepithelial neoplasia 5
Figure 1.3: HPV genome organisation 11
Figure 1.4: Epithelial differentiation and the HPV life cycle 17
Figure 1.5: E7 acting upon the cell cycle 18
Figure 1.6: Structure of E6 19
Figure 1.7: E6 interacts with a number of cellular proteins 19
Figure 1.8: HPV oncoprotein cooperation in cervical carcinogenesis 29
Figure 1.9: E6 and E7 affect a number of cellular targets in carcinogenesis 29
Figure 1.10: Epithelial polarity and junction complexes 29
Figure 1.11: The PDZ domain structure 31
Figure 1.12: PDZ domain containing protein domain structure 33
Figure 1.13: Patterns of alternative splicing 33
Figure 1.14: Differential expression of PDZ proteins in tumour progression 35
Figure 1.15: Viral proteins target common cellular PDZ proteins 41
Figure 1.16: C-terminal alignment of high- and low-risk HPV E6 proteins 53

Chapter 2

Figure 2.1: Transfection of primary keratinocytes with HPV genomes 59
Figure 2.2: Organotypic raft culture 63

Chapter 3

Figure 3.1: C-terminal alignment of high- and low-risk a-HPV E6 proteins 97
Figure 3.2: Generation of HPV16 and HPV18 genome containing primary keratinocyte cell lines 103
Figure 3.3: Confirmation of episomal genomes in HPV18 genomes with modified PDZ binding activity 103
Figure 3.4: Quantification of selected E6-PDZ target transcription in HPV16 and HPV18 genome containing keratinocytes 108
Figure 3.5: Quantification of MAGI-1, MAGI-2 and MAGI-3 transcription in HPV16 and HPV18 genome-containing keratinocytes 109
Figure 3.6: Cumulative qRTPCR analysis PDZ targets of E6 expression in HPV genome containing HFKs 109
Figure 3.7: Quantification of E6-PDZ target transcription in HPV16 and HPV18 genome containing keratinocytes following cell propagation 111
Figure 3.8: Quantification of E6-PDZ target transcription in HPV18 genome containing keratinocytes 113
Figure 3.9: Confirmation of successful enrichment of protein from subcellular fractions 113
Figure 3.10: DLG1 protein expression in keratinocytes replicating HPV genomes

Figure 3.11: DLG1 protein distribution in keratinocytes

Figure 3.12: DLG1 protein in differentiated keratinocytes containing HPV18 genomes

Figure 3.13: Immunofluorescence of DLG1 in differentiated keratinocytes

Figure 3.14: SCRIB protein levels in HPV genome containing keratinocytes

Figure 3.15: Distribution of SCRIB in monolayer keratinocytes

Figure 3.16: SCRIB protein in differentiated keratinocytes containing HPV18 genomes

Figure 3.17: SCRIB distribution in differentiated keratinocytes

Figure 3.18: CASK protein level and distribution in monolayer grown keratinocytes

Figure 3.19: PATJ protein levels in HPV genome containing keratinocytes

Figure 3.20: Distribution of PATJ in monolayer keratinocytes

Figure 3.21: PARD3 protein levels and distribution in primary keratinocytes

Figure 3.22: MAGI-1 protein levels and distribution in primary keratinocytes

Figure 3.23: MUPP1 protein levels and distribution in primary keratinocytes

Figure 3.24: ZO-1 protein levels in HPV genome containing keratinocytes

Figure 3.25: ZO-1 protein in differentiated keratinocyte culture

Figure 3.26: Distribution of ZO-1 in monolayer keratinocytes

Figure 3.27: ZO-1 distribution in differentiated keratinocytes

Figure 3.28: ZO-2 protein levels in HPV genome containing keratinocytes

Figure 3.29: Distribution of ZO-2 protein in monolayer keratinocytes

Figure 3.30: ZO-2 protein in differentiated keratinocytes

Figure 3.31: NHERF1 protein levels in HPV genome containing keratinocytes

Figure 3.32: Distribution of NHERF1 protein in monolayer keratinocytes

Figure 3.33: NHERF1 protein in differentiated keratinocytes

Figure 3.34: NHERF1 distribution in differentiated keratinocytes

Figure 3.35: PTPN13 protein levels in HPV genome containing keratinocytes

Figure 3.36: Distribution of PTPN13 in monolayer keratinocytes

Figure 3.37: PTPN13 protein in differentiated keratinocytes

Figure 3.38: Proteosomal inhibition of primary keratinocytes

Figure 3.39: DLG1 protein levels in keratinocytes under proteosomal inhibition

Figure 3.40: SCRIB protein in keratinocytes under conditions of proteosomal inhibition

Figure 3.41: NHERF1 and PTPN13 protein levels in keratinocytes under proteosomal inhibition

Figure 3.42: Summary of PDZ protein expression in the presence of high-risk α-HPV genomes

Chapter 4

Figure 4.1: Domain structure of DLG1 isoform

Figure 4.2: Regions of alternative splicing of DLG1

Figure 4.3: DLG1 HOOK domain corresponding primers and probes

Figure 4.4: Amplification of the DLG1 HOOK region

Figure 4.5: Sequencing across exon junctions

Figure 4.6: Amplification of the DLG1 N-terminus

Figure 4.7 In silico translation of nucleotide sequence
Figure 4.8: Cross-reaction of subcellular fractions with HRP-conjugated secondary antibody
Figure 4.9: Detection of DLG1 protein in cellular compartments of HeLa and HPV18 genome containing keratinocytes
Figure 4.10: Detection of DLG1α protein
Figure 4.11: DLG1 isoform expression in primary keratinocytes containing α-HPV genomes
Figure 4.12: Expression of DLG1 in differentiating keratinocytes containing HPV18
Figure 4.13: DLG1 isoform expression in keratinocytes containing HPV18 genomes
Figure 4.14: DLG1 isoform expression in keratinocytes containing episomal and integrated forms of the HPV18 genome
Figure 4.15: DLG1 isoform expression in normal tissue
Figure 4.16: DLG1 isoform expression in cancer cell lines
Figure 4.17: DLG1 isoform expression in head and neck cancer derived cell lines
Figure 4.18: Predicted protein structure of I6-containing DLG1

Chapter 5

Figure 5.1: Alignment of E6 C-terminal sequences
Figure 5.2: The HPV18 E6 PDZ binding motif
Figure 5.3: Computational screen of E6 C-terminal sequences
Figure 5.4: Phosphorylation of E6 in HPV18 positive cells
Figure 5.5: Distribution of phosphorylated E6 in monolayer-grown cells
Figure 5.6: The effect of PKA inhibition upon HPV18 E6 protein
Figure 5.7: HPV oncoprotein expression in HeLa expressing mPKI
Figure 5.8: HPV oncoprotein expression in primary cells expressing mPKI
Figure 5.9: E6 protein in HPV18 genome containing HFKs
Figure 5.10: 14-3-3ζ protein expression in HPV genome containing keratinocytes
Figure 5.11: The consequences of E6 phosphorylation
LIST OF TABLES

Tables are found alongside the text

Chapter 1
Table 1.1: Biological properties associated with different HPV genus ........................................ 2
Table 1.2: HPV types associated with human disease ................................................................. 5
Table 1.3: IARC classification of HPV carcinogenic risk ............................................................ 6
Table 1.4: High-risk HPV ORFs ...................................................................................................... 14
Table 1.5: Protein partners of E6 .................................................................................................. 21
Table 1.6: Cell polarity and adhesion complex components .................................................... 30
Table 1.7: Classes of PDZ-binding motif ..................................................................................... 32
Table 1.8: PDZ proteins in human disease .................................................................................... 36
Table 1.9: Virus interactions with PDZ proteins .......................................................................... 41
Table 1.10: PDZ targets of high-risk HPV E6 proteins .............................................................. 54

Chapter 2
Table 2.1: Cell lines ....................................................................................................................... 57
Table 2.2: qRTPCR primers ......................................................................................................... 72
Table 2.3: DLG1 primers .............................................................................................................. 75
Table 2.4: Restriction enzymes used to assess HPV genome status ........................................ 79
Table 2.5: Primary antibodies used in western blotting and immunofluorescence ................ 90
Table 2.6: Secondary antibodies used in western blotting and immunofluorescence ............. 91

Chapter 3
Table 3.1: PDZ proteins associated with high-risk HPV E6 proteins ......................................... 99
Table 3.2: Affymetrix microarray analysis of E6 PDZ target expression in HPV16 and HPV18 genome containing keratinocytes ................................................................. 103
Table 3.3: Summary of PDZ target qRTPCR analysis .............................................................. 107
Table 3.4: Comparison of microarray and qRTPCR analyses .................................................... 108
Table 3.5: Summary of PDZ protein analysis ............................................................................. 121

Chapter 4
Table 4.1: Insertions in DLG1 variants .......................................................................................... 134
Table 4.2: DLG1 primer and array probe positions ..................................................................... 143
Table 4.3: Exon and intron comparison to consensus splice nucleotide sequence ..................... 145
Table 4.4: Potential protein interactions with I2-I6 sequence ................................................... 147
Table 4.5: DLG1 Antibodies ......................................................................................................... 148
Table 4.6: Microarray and qRTPCR analysis of DLG1 expression ........................................... 152

Chapter 5
Table 5.1: Examples of PBM phosphorylation .......................................................................... 168
Table 5.2: Consequences of cellular PBM Phosphorylation .................................................... 172
### LIST OF ABBREVIATIONS

Standard international (SI) units are used throughout this thesis.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td>Anchorage independent growth</td>
</tr>
<tr>
<td>AKAP</td>
<td>A-kinase anchor protein</td>
</tr>
<tr>
<td>AMPA</td>
<td>α-amino-3-hydroxy-5-methyl-4-isoxazole proprionic acid</td>
</tr>
<tr>
<td>AMPAR</td>
<td>AMPA receptor</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variants</td>
</tr>
<tr>
<td>AP-1</td>
<td>Activator protein-1</td>
</tr>
<tr>
<td>APC</td>
<td>Adenomatous polyposis coli protein</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulphate</td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia telangiectasia mutated protein</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenine triphosphate</td>
</tr>
<tr>
<td>B2M</td>
<td>β2-Microglobulin</td>
</tr>
<tr>
<td>BCC</td>
<td>Basal cell carcinoma</td>
</tr>
<tr>
<td>BPV</td>
<td>Bovine papillomavirus</td>
</tr>
<tr>
<td>BrdU</td>
<td>5-bromo-2-deoxyuridine</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CAL</td>
<td>Conductance regulator-associated ligand</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CAMK</td>
<td>Ca²⁺/Calmodulin dependant protein kinase</td>
</tr>
<tr>
<td>CARD</td>
<td>Caspase activation and recruitment domain</td>
</tr>
<tr>
<td>CASK</td>
<td>Calcium/calmodulin-dependent serine protein kinase</td>
</tr>
<tr>
<td>CEM</td>
<td>Complete E media</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin-dependent protein kinase</td>
</tr>
<tr>
<td>CIN</td>
<td>Cervical intraepithelial neoplasia</td>
</tr>
<tr>
<td>CK2</td>
<td>Casein kinase 2</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>Co-IP</td>
<td>co-Immunoprecipitation</td>
</tr>
<tr>
<td>CR</td>
<td>Conserved region</td>
</tr>
<tr>
<td>crPV</td>
<td>Cottontail rabbit papillomavirus</td>
</tr>
<tr>
<td>DABCO</td>
<td>1,4-diazabicyclo{2,2,2}octane</td>
</tr>
<tr>
<td>DAPI</td>
<td>4',6'-diamino-2-phenylindole</td>
</tr>
<tr>
<td>DbcAMP</td>
<td>Dibutyllycyclic adenosine monophosphate</td>
</tr>
<tr>
<td>DLG</td>
<td>Discs large</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide tri-phosphate</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DV</td>
<td>Dengue virus</td>
</tr>
<tr>
<td>E6AP</td>
<td>E6-associated protein</td>
</tr>
<tr>
<td>E6BP</td>
<td>E6-binding protein</td>
</tr>
<tr>
<td>EB</td>
<td>Ezrin binding</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr virus</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>EV</td>
<td>Epidermodysplasia verruciformis</td>
</tr>
<tr>
<td>FBS</td>
<td>Foetal bovine serum</td>
</tr>
<tr>
<td>FCS</td>
<td>Foetal calf serum</td>
</tr>
<tr>
<td>FERM</td>
<td>4.1, ezrin, radixin, moesin</td>
</tr>
<tr>
<td>FK</td>
<td>Forskolin</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GK</td>
<td>Guanylate kinase</td>
</tr>
<tr>
<td>GKAP</td>
<td>Guanylate kinase associated protein</td>
</tr>
<tr>
<td>GMP</td>
<td>Guanosine monophosphate</td>
</tr>
<tr>
<td>GPCR</td>
<td>G-protein coupled receptor</td>
</tr>
<tr>
<td>GRB2</td>
<td>Growth factor receptor-bound protein 2</td>
</tr>
<tr>
<td>GRID2IP</td>
<td>Glutamate receptor ionotrophic, delta 2 interacting protein</td>
</tr>
<tr>
<td>GRIP</td>
<td>Glutamate receptor interacting protein</td>
</tr>
<tr>
<td>GRB2</td>
<td>Growth factor receptor-bound protein 2</td>
</tr>
<tr>
<td>GRK5</td>
<td>G protein coupled receptor kinase 5</td>
</tr>
<tr>
<td>GukH</td>
<td>GKHolder protein</td>
</tr>
<tr>
<td>Gy</td>
<td>Gray</td>
</tr>
<tr>
<td>H89</td>
<td>Dihydrochloride</td>
</tr>
<tr>
<td>H and E</td>
<td>Haematoxylin and Eosin</td>
</tr>
<tr>
<td>HAT</td>
<td>Histone acetyltransferase</td>
</tr>
<tr>
<td>HBSS</td>
<td>Hank’s balanced salt solution</td>
</tr>
<tr>
<td>HBV</td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HIER</td>
<td>Heat-induced epitope retrieval</td>
</tr>
<tr>
<td>HFK</td>
<td>Human foreskin keratinocyte</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HINGS</td>
<td>Heat-inactivated goat serum</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HMAST</td>
<td>Microtubule associated protein kinase</td>
</tr>
<tr>
<td>HNSCC</td>
<td>Head and neck squamous cell carcinoma</td>
</tr>
<tr>
<td>HPV</td>
<td>Human papillomavirus</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>HTK</td>
<td>Human tonsil keratinocyte</td>
</tr>
<tr>
<td>HTLV</td>
<td>Human T-lymphotrophic virus</td>
</tr>
<tr>
<td>IBMX</td>
<td>3-Isobutyl-1-methylxanthine</td>
</tr>
<tr>
<td>IARC</td>
<td>International agency for research on cancer</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitation</td>
</tr>
<tr>
<td>IRF</td>
<td>Interferon response factor</td>
</tr>
<tr>
<td>JNK</td>
<td>Jun N-terminal kinase</td>
</tr>
<tr>
<td>K18</td>
<td>Keratin 18</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>Kir 2.3</td>
<td>Potassium inwardly rectifying channel</td>
</tr>
<tr>
<td>KSHV</td>
<td>Kaposi’s sarcoma herpes virus</td>
</tr>
<tr>
<td>L27</td>
<td>Lin2-7</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LncRNA</td>
<td>Long non-coding RNA</td>
</tr>
<tr>
<td>LRP</td>
<td>Lipoprotein receptor related protein</td>
</tr>
<tr>
<td>mAmp</td>
<td>Milliamp</td>
</tr>
<tr>
<td>NFκB</td>
<td>Nuclear factor κ-light chain enhancer of activated B cells</td>
</tr>
<tr>
<td>NMSC</td>
<td>Non-melanoma skin cancer</td>
</tr>
<tr>
<td>MAGI</td>
<td>Membrane associated guanylate kinase homology with an inverted domain</td>
</tr>
<tr>
<td>MAGUK</td>
<td>Membrane-associated guanylate kinase</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MC</td>
<td>Methycellulose</td>
</tr>
<tr>
<td>MCC</td>
<td>Mutated in colorectal cancer</td>
</tr>
<tr>
<td>MCM7</td>
<td>Minichromosome maintenance complex component 7</td>
</tr>
<tr>
<td>MCV</td>
<td>Merkel cell polyomavirus</td>
</tr>
<tr>
<td>MDCCK</td>
<td>Madin-Darby Canine Kidney</td>
</tr>
<tr>
<td>MG132</td>
<td>z-Leu-Leu-Leu-CHO</td>
</tr>
<tr>
<td>MG115</td>
<td>z-Leu-Leu-Nva-H</td>
</tr>
<tr>
<td>MLV</td>
<td>Murine leukemia virus</td>
</tr>
<tr>
<td>MUPP</td>
<td>Multiple PDZ protein</td>
</tr>
<tr>
<td>NET</td>
<td>Neuroepithelioma transforming gene</td>
</tr>
<tr>
<td>NHERF1</td>
<td>Na+/H+ exchange regulatory factor 1</td>
</tr>
<tr>
<td>NMDA</td>
<td>N-methyl-D-aspartate</td>
</tr>
<tr>
<td>NMDAR</td>
<td>NMDA receptor</td>
</tr>
<tr>
<td>NMSC</td>
<td>Non-melanoma skin cancer</td>
</tr>
<tr>
<td>NR2B</td>
<td>NMDAR subtype 2B</td>
</tr>
<tr>
<td>NS</td>
<td>Non-structural protein</td>
</tr>
<tr>
<td>OC</td>
<td>Open circle</td>
</tr>
<tr>
<td>OPC</td>
<td>Oropharyngeal carcinoma</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>P</td>
<td>Passage</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PAK1</td>
<td>p21 activated kinase 1;</td>
</tr>
<tr>
<td>PALS1</td>
<td>Pals associated with lin-7</td>
</tr>
<tr>
<td>PARD</td>
<td>Partitioning defective</td>
</tr>
<tr>
<td>PATJ</td>
<td>Pals-1 associated tight junction protein</td>
</tr>
<tr>
<td>PBK</td>
<td>PDZ binding kinase</td>
</tr>
<tr>
<td>PBM</td>
<td>PDZ-Binding domain motif</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PBST</td>
<td>Phosphate buffered saline-Tween20</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDGF</td>
<td>Platelet derived growth factor receptor</td>
</tr>
<tr>
<td>PDZ</td>
<td>PSD/DLG/ZO</td>
</tr>
<tr>
<td>PH</td>
<td>Pleckstrin homology</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphatidylinositol 3-kinase</td>
</tr>
<tr>
<td>PICK1</td>
<td>Protein interacting with C kinase 1</td>
</tr>
<tr>
<td>PIP2</td>
<td>4,5-Bisphosphate</td>
</tr>
<tr>
<td>PKA</td>
<td>Protein kinase A</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein kinase C</td>
</tr>
<tr>
<td>PKI</td>
<td>Protein kinase A inhibitor</td>
</tr>
<tr>
<td>PSD</td>
<td>Post-synaptic density protein</td>
</tr>
<tr>
<td>PSF</td>
<td>Protein associated splicing factor</td>
</tr>
<tr>
<td>PTEN</td>
<td>Phosphatase and tensin homologue</td>
</tr>
<tr>
<td>PTP</td>
<td>Protein tyrosine phosphatase</td>
</tr>
<tr>
<td>PTPN</td>
<td>Protein tyrosine phosphatase non-receptor</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>qRTPCR</td>
<td>Quantitative RT PCR</td>
</tr>
<tr>
<td>Rb</td>
<td>Retinoblastoma</td>
</tr>
<tr>
<td>rhPV</td>
<td>Rhesus papillomavirus</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
</tbody>
</table>
RTPCR - Reverse transcription
polymerase chain reaction
SAP97 - Synapse-associated protein 97
SARS - Severe acute respiratory syndrome
SC - Supercoiled
SCC - Squamous cell carcinoma
SCC - Saline sodium citrate
SCRIB - Scribble protein
SDS - Sodium dodecyl sulphate
SFM - Serum free media
SH3 - Src Homology 3
SI - Standard international
SIV - Simian immunodeficiency virus
SP-1 - Stimulating protein-1
SRPK - Serine-arginine protein kinase
SRSF - Serine/arginine rich splicing factor
T3 - Tri-iodo-1L-thyronine
TCR - T-cell receptor
TE - Tris-EDTA
TEMED -
N,N,N',N'Tetramethylene-diamine
TBE - Tris/Borate/EDTA
TBEV - Tick borne encephalitis virus
TBS - Tris-buffered saline
TBST - Tris-buffered saline-Tween20
TI - Trypsin inhibitor
TIP - Tax interacting protein
TJ - Tight junction
T_m - Primer melting temperature
TNF - Tumour necrosis factor
TNFR - Tumour necrosis factor receptor
UTB - Urea, Tris, β-mercaptoethanol
UTR - Untranslated region
UV - Ultraviolet
V - Volts
VLP - Virus like particle
VP - Viral protein
WHO - World Health Organisation
WT - Wild type
ZA - Zonula adherens
ZO - Zona ocludens protein
ZONAB - ZO-1 associated nucleic acid binding protein
CHAPTER 1: INTRODUCTION

1.1 Papillomaviruses

Papillomaviruses were first discovered in 1907, when the Italian scientist Gieseppe Ciuffo inoculated himself with the cell-free extract of warts to prove that the growths were caused by a transmissible agent (Ciuffo, 1907). More than twenty years later, while Francis Peyton Rous was investigating transmissible agents causing tumours in chickens, virologists Richard Shope and E. Weston Hurst were studying warts taken from wild cottontail rabbits. In the soluble extracts of these warts, they found a transmissible agent, the cottontail rabbit papillomavirus (CRPV), that was capable of inducing benign growths when introduced to uninfected rabbits (Shope and Hurst, 1933). The pair passed on tissue to the aforementioned Rous, who used it to infect a domestic species of rabbit. These rabbits failed to form warts, which would have been evidence of normal successful viral replication, but instead developed aggressive skin carcinomas (Rous and Beard, 1935).

Rous went on to establish retroviruses as cancer-causing agents in chickens, but it would be more than 30 years until human papillomaviruses (HPVs) were linked to human cancers and accepted as a human tumour virus.

Papillomaviruses, named for the papilla (buds in Latin) that infected cells form, are a highly diverse group of small, double-stranded DNA viruses. Their diversity is reflected in the wide range of hosts that they infect, which includes reptiles, birds and mammals (Herbst et al., 2009; Chen et al., 1982; Schulz et al., 2009; Bernard et al., 2010). There are more than 300 known papillomaviruses, 170 of which are officially human-specific, with other types pending classification (Bernard et al., 2010; de Villiers, 2013; van Doorslaer et al., 2013). The viruses are classified by the nucleotide sequence encoding the conserved L1 open reading frame (ORF); this is the most conserved gene within the genome. Members of the same
papillomavirus genus are denoted by Greek letters (i.e. α, β, γ, µ) and share more than 60% nucleotide sequence identity in the L1 ORF (Bernard et al., 2010; de Villiers, 2013; de Villiers et al., 2004). These are sub-divided into HPV types, which are numbered (e.g. HPV16, HPV18; both α-papillomaviruses) and defined by an L1 region containing DNA sequence that differs by more than 10% to the closest known type (Figure 1.1, Table 1.1). Virus variants are further classified (e.g. HPV16 114/B, HPV16 114/K), by containing a less than 2% difference in the L1 region (Bernard et al., 2010; de Villiers, 2013; de Villiers et al., 2004).

Table 1.1: Biological properties associated with different HPV genus

<table>
<thead>
<tr>
<th>Genus</th>
<th>Biological Properties</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha-papillomavirus</td>
<td>Mucosal and cutaneous lesions in humans and primates Many high-risk types under this genus</td>
<td>HPV6, HPV11, HPV16, HPV18</td>
</tr>
<tr>
<td>Beta-papillomavirus</td>
<td>Cutaneous lesions in humans Predominantly latent infection, re-activated in immunosuppressed conditions Close association with Epidermodysplasia verruciformis (EV)</td>
<td>HPV8 HPV12</td>
</tr>
<tr>
<td>Gamma-papillomavirus</td>
<td>Cutaneous lesions in humans</td>
<td>HPV60</td>
</tr>
<tr>
<td>Nu-papillomavirus</td>
<td>Benign and malignant cutaneous lesions</td>
<td>HPV41</td>
</tr>
<tr>
<td>Mu-papillomavirus</td>
<td>Cutaneous lesions</td>
<td>HPV1</td>
</tr>
</tbody>
</table>

The major human-tropic genus are mentioned here. Adapted from de Villiers et al., 2004.

HPVs infect cutaneous or mucosal epithelial (Table 1.1), causing a variety of lesions; from benign plantar, flat and common warts and genital warts, to cervical neoplasia and cancer (Table 1.2). As such, HPV types can be classified as high- or low-risk, according to their capacity to promote malignant growth.

1.2 Benign HPV infection

A number of HPV types cause benign growths in mucosal and cornified epithelia. Infection with low-risk types of HPV can cause condylomata acuminata (genital warts) or
Figure 1.1: Human papillomavirus classification
Classification of papillomavirus types is based on the DNA sequence of the L1 gene; HPV types differ by more than 10% in the DNA sequence of the L1, whereas HPV variants differ by less than 2% in L1 sequence. Members of the same genus share >60% nucleotide sequence identity in the L1. HPVs are contained within evolutionary groups. The outermost semicircular symbols indicate papillomavirus genera, e.g. alpha-papillomavirus. The number at the inner semicircular symbol refers to papillomavirus species e.g. HPV16. The HPV types that infect the cervix come from the Alpha group which contains over 60 members. HPV types from the Beta, Gamma, Mu and Nu groups primarily infect cutaneous sites. Adapted from de Villiers, 2004.
 verrucae (skin warts). Genital warts are the most prevalent sexually-transmitted infection worldwide (Baseman and Koutsky, 2005), with the HPV types 6 or 11 found in 100% of cases and co-infection with high-risk types associated with 25% of cases (Lacey, 2005; Brown et al., 1999; Greer et al., 1995). Genital warts are highly infectious, with a transmission rate of about 65% and an incubation period of between 3 weeks and 8 months, with the majority developing warts at around 2–3 months (Oriel, 1971). After regression, it has been suggested that asymptomatic infection persists for life (Stanley, 2010).

Verrucae, caused by HPV infection of cornified epithelium, are most common in children and adolescents, the majority of whom clear the growths spontaneously (Williams et al., 1993). The many and varied types of warts are classified based upon the site of infection and pathology (Table 1.2). Painful and unsightly warts can be treated by removal, but there is still a requirement for the immune system in disease clearance (Sterling et al., 2001).

Although infection with low-risk viruses is largely benign, immunocompromised individuals can experience more damaging effects. Patients with the rare heritable disease, epidermodysplasia verruciformis (EV), are susceptible to β-HPV infection and are at risk of developing non-melanoma skin cancer (NMSC); 30-60% of the patients develop carcinoma (Akgül et al., 2006). EV patients exhibit an inability to control HPV infection, resulting in the development of scaly growths, particularly on the hands and feet. This is attributed to a defect in the immune system’s ability to reject the HPV-positive keratinocytes (Majewski and Jabłońska, 1995; Majewski and Jabłonska, 2002; Baadsgaard et al., 1990). This is reinforced by the fact that β-HPVs are ubiquitous in the general population and EV-like conditions develop in HIV infected individuals as well as other immunosuppressive conditions (Majewski and Jabłońska, 1995). In addition, genetic screening has revealed that loss of function mutations in the EVER1 and EVER2 genes,
which encode zinc-transporting proteins, are responsible in a majority of cases (Ramoz et al., 1999; 2000; 2002). It has been suggested that the EVER genes affect the virus life cycle by limiting the accessibility of Zn$^{2+}$ ions to the viral zinc binding proteins E6 and E7, and that somatic mutations remove this anti-HPV activity (Lazarczyk et al., 2008). It is also thought that the loss of EVER gene function impairs T-cell activation and proliferation, by impairing regulation of lymphocyte Zn$^{2+}$ concentration, thus impairing immunity against HPV in EV patients (Lazarczyk et al., 2012). UV light is also influential in the development of EV associated squamous cell carcinoma (SCC); although warts develop on all body sites, carcinomas occur exclusively on sites exposed to the sun (Majewski and Jabłońska, 1995). Thus, the pathology of EV demonstrates the interplay between multiple factors, genetic and environmental in HPV mediated carcinogenesis.

1.3 HPV infection and cancer

It was not until the late 1980s that papillomaviruses were established as the etiological agent of cervical cancer in humans, and now HPVs are associated with more human cancers than any other virus (WHO). Infection with high-risk HPV is responsible for 4.8% of the world’s cancer burden, causing over 610,000 cases per year worldwide (de Martel et al., 2012; Forman et al., 2012). The discovery was made by Nobel laureate Harald zur Hausen, who was studying the genital condylomata-causing virus in women (Hausen et al., 1974). HPV has subsequently been shown as the causative agent in a number cancers of the anogenital tract, as well as important factors in the development of NMSC and in carcinomas of the oral and oropharyngeal mucosa (Parkin and Bray, 2006; Lowy and Schiller, 2012). To date, twelve HPV types have been defined by the World Health Organisation (WHO) as ‘high-risk’ with respect to causing cancer (IARC Group 1), and a
further ten types are recognised as ‘probably’ and ‘possibly’ cancer-causing (IARC Group 2A and 2B, Table 1.3, Bouvard et al., 2009).

**Table 1.2: HPV types associated with human disease**

<table>
<thead>
<tr>
<th>Disease</th>
<th>Most frequently associated HPV types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common warts</td>
<td>HPV2, 4, 7; occasionally other types in immunosuppressed (e.g. HPV 75, 76, 77)</td>
</tr>
<tr>
<td>Flat plane warts</td>
<td>HPV3, 10, occasionally types 26–29 and 41</td>
</tr>
<tr>
<td>Plantar warts</td>
<td>HPV1, 2, 4</td>
</tr>
<tr>
<td><strong>Epidermodysplasia verruciformis (EV)</strong></td>
<td></td>
</tr>
<tr>
<td>Plane warts</td>
<td>HPV3, 10</td>
</tr>
<tr>
<td>Scale-like plaques</td>
<td>HPV5, 8; less commonly 9, 12, 14, 15, 17, 19, 20, 21–25, 36–39, 47, 49</td>
</tr>
<tr>
<td>Squamous cell carcinomas of sun-exposed skin</td>
<td>HPV5, 8, less commonly 14,17,20 and 47</td>
</tr>
<tr>
<td>External warts</td>
<td>HPV6, 11, 40, 42, 43, 44, 54, 61, 72, 81, 89</td>
</tr>
<tr>
<td><strong>Anogenital Warts</strong></td>
<td></td>
</tr>
<tr>
<td>Buschke–Lowenstein tumour</td>
<td>HPV6</td>
</tr>
<tr>
<td>Bowenoid papulosis</td>
<td>HPV16, 55</td>
</tr>
<tr>
<td>IARC Group 1: Carcinogenic</td>
<td>HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66</td>
</tr>
<tr>
<td>IARC Group 2A: Probably carcinogenic</td>
<td>HPV68</td>
</tr>
<tr>
<td>IARC Group 2B: Possibly carcinogenic</td>
<td>HPV26, 53, 64, 65, 66, 67, 69, 70, 73, 82</td>
</tr>
<tr>
<td><strong>Oral Lesions</strong></td>
<td></td>
</tr>
<tr>
<td>Oral papillomas</td>
<td>HPV2, 6, 7, 11, 16, 18, 32, 57</td>
</tr>
<tr>
<td>Laryngeal papillomas</td>
<td>HPV6, 11</td>
</tr>
<tr>
<td>Focal hyperplasia (Heck's disease)</td>
<td>HPV13, 32</td>
</tr>
<tr>
<td>Oropharyngeal carcinoma</td>
<td>Predominantly HPV16</td>
</tr>
</tbody>
</table>

Adapted from Cubie, 2013.

1.3.1 Anogenital cancers

Altogether, there are more than 530,000 new cases and 275,000 deaths from cervical cancer per year (Ferlay et al., 2012). In identifying HPV as a cause of cervical cancer, Harald zur
Hausen determined the presence of HPV16 and 18 DNA in cancer tissue; together these α-HPVs contribute to more than 70% of cervical cancer cases worldwide (Li et al., 2011). HPV-induced cancer of the cervix is the most characterized; stages of cervical neoplasia are clinically categorized by cytological and histological examination; pre-malignant phases are termed cervical intraepithelial neoplasia (CIN) and are graded 1 to 3 (Woodman et al., 2007). These pre-malignant changes represent a spectrum ranging from mild dysplasia (CIN1), moderate dysplasia (CIN2), to severe dysplasia and carcinoma (CIN3) (Figure 1.2). These progressively malignant changes are associated with the increased expression of HPV encoded oncoproteins E6 and E7, which together provide the primary transforming activity of high-risk HPVs (Section 1.4.2). The most common HPV types found in cervical cancers are HPVs 16, 18, 58, 52, 31 and 33 (Clifford et al., 2006). In the cervix, HPV-driven cancers arise from squamous cells and adenomatous glandular cells. Squamous cell carcinomas (SCC) account for 80% of cervical cancer cases, of which 62% are HPV16 associated (de Sanjose et al., 2010; Walboomers et al., 1999). By contrast, adenocarcinomas are less frequent and are more commonly associated with HPV18 (de Sanjose et al., 2010; Green et al., 2003).

Table 1.3: IARC classification of HPV carcinogenic risk

<table>
<thead>
<tr>
<th>IARC Group</th>
<th>HPV types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 - carcinogenic</td>
<td>16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59</td>
</tr>
<tr>
<td>Group 2A - probably carcinogenic</td>
<td>68</td>
</tr>
<tr>
<td>Group 2B - possibly carcinogenic</td>
<td>26, 53, 66, 67, 70, 73, 5, 8, 6, 11</td>
</tr>
</tbody>
</table>

Adapted from Bouvard et al., 2009.

Other anogenital malignancies associated with high-risk HPV include penile and vulvo-vaginal cancers, where 40% of incidences are attributed to HPV, and anal carcinomas, where 90% of incidences are attributed to HPV (Parkin and Bray, 2006; Ferlay et al., 2012). The high proportion of HPV associated with anogenital cancers indicates the necessity for
Figure 1.2: Stages of cervical intraepithelial neoplasia

HPV is thought to access basal cells via micro-abrasions in the cervical epithelium. Following infection, viral DNA replicates and exists as episomes (purple nuclei). In the upper layers of epithelium, the viral genome is replicated further, and progeny virions form in the nucleus. Low-grade intraepithelial lesions support productive viral replication. An unknown number of high-risk HPV infections progress to high-grade cervical intraepithelial neoplasia (CIN). These pre-malignant changes represent a spectrum of histological abnormalities ranging from CIN1 (mild dysplasia), CIN2 (moderate dysplasia), to CIN3 (severe dysplasia/carcinoma in situ). The progression of untreated lesions to microinvasive and invasive cancer is associated with the integration of the HPV genome into the host chromosomes (red nuclei). Adapted from Woodman et al. 2007.
viral infection in cancer development. However, not all infections with HPV lead to invasive disease; it has been suggested that 80 to 90% of HPV infections persist for 12 to 18 months and are cleared by the immune system (Bodily and Laimins, 2011; Stanley, 2012). As with EV, a number of co-factors have been reported to aid the development of high-risk HPV-driven anogenital carcinoma, including hormones, smoking and immunosuppression (Green et al., 2003; Kjellberg et al., 2000; Coutlee et al., 2012; Hildesheim et al., 1990; Harper et al., 1994). Smoking has been shown to be the most significant cofactor in cervical cancer development; current smokers are twice as likely as non-smokers to be diagnosed with CIN (Collins et al., 2010; Kjellberg et al., 2000). Chemical carcinogens found in cigarette smoke are thought to influence carcinogenesis by causing DNA damage and impairing the immune system; there is no evidence to suggest that HPV infection is prolonged by the effects of smoking (Collins et al., 2010; Castellsagué and Muñoz, 2003). As with low-risk types, immunosuppressed patients are susceptible to the development of HPV-driven cancers. For example, HPV infection is more common and more likely to persist in HIV-positive compared to HIV-negative women (Clifford et al., 2006). Whilst the exact mechanisms are unclear, co-factors are proposed to influence progression to carcinogenesis via suppression of the immune system, increased viral persistence, and/or inducing DNA damage.

1.3.1.1 Screening and vaccination

Rates of cervical cancer have fallen in developed countries over the last few decades, thanks to the implementation of screening programs, which allow the early detection and treatment of pre-cancerous cervical lesions (Feldman, 2014). In the UK, the Papanicolaou (Pap) test of exfoliated cervical cells was introduced nationally in 1988, since which mortality rates of cervical cancer have decreased by 80% (Peto et al., 2004). Despite the
decrease in SCC rates, the rates of adenocarcinoma have been reported to increase, which is attributed to cervical screening being less efficient at the detection of adenocarcinoma; the disease presents minor cytological abnormalities compared to SCC and so screening is less likely to identify and interrupt disease progression (Clifford et al., 2003; Muñoz et al., 2003; Feldman, 2014; Woodman et al., 2007; 2001).

A recent development in the protection against cervical carcinoma has been achieved in the licensing of two prophylactic vaccines against HPV infection: the quadrivalent vaccine Gardasil®, which protects against HPV types 16, 18, 6 and 11 and the bivalent vaccine Cervarix®, which protects against HPV16 and 18. Both vaccines show a level of cross-protection against infection with other HPV types, including types HPV31 and HPV45 (Bonanni et al., 2009; Toft et al., 2014). Cervical cancer screening is an expensive preventative health care measure; in England, it costs the NHS around £175 million a year, including treatment (http://www.cancerscreening.nhs.uk/). A vaccination program may alleviate some of this cost, but screening will still be required for those who are missed by the vaccination program, and to catch cancers induced by the high-risk HPV types not targeted by the vaccines (Villa, 2011). As the incidence of cervical cancer decreases in the developed world due to a better understanding and improved screening protocols, it remains constant in the developing world where screening and treatment are difficult and expensive to implement (Forouzanfar et al., 2011; Parkin and Bray, 2006). The HPV-associated cancer burden is disproportionally spread geographically; more than 76% of cervical cancers occur in developing countries, where there is a 80% lower chance of successful treatment, compared to women living in more developed countries (Baseman and Koutsky, 2005; Parkin et al., 2001; de Martel et al., 2012). As in the case of screening programs, poor infrastructure and high costs make implementation of vaccination in
developing countries unfeasible, and so better methods of prevention are still required in countries where disease burden is greatest (de Martel et al., 2012).

Interestingly, vaccine development is being explored as a treatment against HPV-positive tumour cells; using peptide sequences corresponding to E6 and E7 to induce an immune response against these proteins and thus against E6 and E7-expressing tumor cells. This is currently under development, but has been effective in inducing a response from lymphocytes during a study of HPV-positive vulvar intraepithelial neoplasia (VIN) patients (Le Buanec et al., 1999; Wieking et al., 2012; Wang et al., 2000).

1.3.2 Non-genital cancers

1.3.2.1 Non-melanoma skin cancer

NMSC is the most prevalent malignancy in the Caucasian population worldwide (Alam and Ratner, 2001). The term includes both basal cell carcinomas (BCCs) and squamous cell carcinomas (SCCs); the former being more common and contributing to 80% of NMSC (Rubin et al., 2005). Ultraviolet (UV) irradiation is a well characterised risk factor in the development of NMSC and molecular studies have identified HPV as a co-factor also; HPV DNA has been identified in 30% to 50% of NMSCs in the non-immunosuppressed population (Meyer et al., 2001; Berkhout et al., 1995; Pfister, 2003). A number of HPV types have been associated with NMSC; notably types 5, 8, 14, 17, 20 and 47 (Harper et al., 2013; Jemal et al., 2013). Unlike in anogenital cancers, the number of copies of HPV genomes per cell in NMSCs is low; only 1 in 20-50,000 cells contain HPV DNA (Pfister, 2003; Weissenborn et al., 2005). Both the requirement for UV radiation in carcinogenesis and the low viral DNA copy number suggests that HPVs promote the initiation and progression of NMSCs, but are not essential for the maintenance of a malignant phenotype (Majewski and Jabłońska, 1995). The influence of HPV in promoting NMSC has not been fully elucidated;
in contrast to mucosal HPVs, the E6 and E7 oncoproteins of cutaneous HPVs have weak or no transforming properties \textit{in vitro} (Akgül \textit{et al.}, 2005). However, E6 and E7 from HPV49 have been shown to efficiently deregulate p53 and Rb pathways in a similar fashion to their mucosal equivalents (Cornet \textit{et al.}, 2012). Furthermore, current thinking suggests that the E6 and E7 proteins of cutaneous HPV types inhibit the normal DNA damage repair response or UV-induced apoptosis (Muschik \textit{et al.}, 2011; Jackson and Storey, 2000; Underbrink \textit{et al.}, 2008; Viarisio \textit{et al.}, 2011). According to this hypothesis, infected cells are less able to efficiently repair any genetic errors induced by UV radiation, leading to the accumulation of mutations, changes in cellular phenotype and the development of cancer (Hanahan and Weinberg, 2011). The E6 proteins of HPV5 and 18 prevent the translocation of apoptosis-inducing factors to the nucleus (Leverrier \textit{et al.}, 2007) and HPV38 E6 has been shown to interact with the histone acetyltransferase (HAT) p300 to inhibit p53 acetylation, and thus inhibit p53-mediated growth arrest and apoptosis (Muench \textit{et al.}, 2009; Muschik \textit{et al.}, 2011). Similarly, E6 proteins of HPV types 5, 10 and 77 have been shown to prevent UV-induced apoptosis, by abrogating the anti-apoptotic functions of Bak (Underbrink \textit{et al.}, 2008; Holloway \textit{et al.}, 2014; Simmonds and Storey, 2008). As with other HPV infections, the immune status of the host is an important co-factor in the development of NMSC; organ-transplant recipients have a 10-fold increased risk of BCC and up to a 100-fold increased risk of SCC (Pfister, 2003). The immunomodulatory functions of E6 and E7 have been suggested to play a role in NMSC progression; it is argued that the secretion of anti-inflammatory cytokines by just a few HPV-positive cells can impair the cancer cell elimination by the host immune system (Harwood and Proby, 2002; Pfister, 2003).
1.3.2.2 Oropharyngeal cancers

Oropharyngeal carcinoma (OPC) is one of the many cancers associated with smoking and alcohol consumption, and results in more than 300,000 deaths per year worldwide (Blot et al., 1988; Ang and Sturgis, 2012; Canova et al., 2010). Since HPV DNA has been found to be more prevalent in OPC tumour than in adjacent normal tissue, the virus is considered an etiological agent for disease (D’Souza et al., 2007). In both the USA and Europe, the incidence of HPV-related OPC has been increasing since the early 1990s, although the proportion of OPC caused by HPV varies widely; in North America 40 to 80% of OPCs are associated with HPV, whereas in Europe the proportion varies from 20 to 90% (Kreimer et al., 2005a; Hansson et al., 2005; Applebaum et al., 2007; Pintos et al., 2008; Smith et al., 2004; Tachezy et al., 2009; Termine et al., 2008; Du et al., 2011; Attner et al., 2009; Mehanna et al., 2012; Chaturvedi et al., 2008). Unlike in cervical carcinomas, HPV16 is more commonly found than HPV18, with 60% of HPV associated OPC in the USA containing HPV16 genomes (Kreimer et al., 2005b; Marur et al., 2010; Herrero et al., 2003). Perhaps counter-intuitively, patients with HPV-positive OPC have a better prognosis than those without; presence of HPV is associated with better tumor control and survival outcomes following radiotherapy, chemotherapy and combined radiation with cisplatin (Lassen et al., 2011; Fakhry et al., 2008; Posner et al., 2011; Ang et al., 2010; Rischin et al., 2010). The mechanism for this is not known, although HPV-positive tumours are on average smaller than HPV-negative carcinomas at diagnosis and it has been suggested that the virus promotes sensitivity to therapy. Furthermore, HPV-mediation of OPC development has been suggested to follow a similar pathway as anogenital cancers (Oh et al., 2013). Ongoing studies aim to identify driver mutations and predictive biomarkers of this disease in order to improve treatment outcome (Ang and Sturgis, 2012).
1.4 The HPV life cycle

Compared to many other viruses, the HPV genome is small, comprising of 8 kb double-stranded DNA and encoding 8 open reading frames (ORFs, Figure 1.3, Table 1.4). These ORFs encode 6 early proteins, E1, E2, E4, E5, E6 and E7, and 2 late structural proteins, L1 and L2, which form the icosahedral capsid of the virus (Buck et al., 2013; Finnen et al., 2003). With the exception of the helicase E1, these ORFs do not encode enzymatic functions, and so viral replication requires cellular machinery. As such, the papillomavirus life cycle is closely tied to the biology of keratinocytes.

Whether or not a productive HPV life cycle is completed depends upon the nature of the epithelial site where infection occurs. A specific population of cells in the cervix (at the squamo-columnar junction) have been identified as having the potential of developing HPV-related CIN, while cells located elsewhere do not (Herfs et al., 2012). As previously indicated, HPV infection can also depend upon the presence of external factors such as hormones and cytokines (Gariglio et al., 2009; de Jong et al., 2004).

Initial HPV infection occurs via micro-abrasions in the epithelium, enabling viral access to the basal lamina (Kines et al., 2009). Viral entry into basal keratinocytes is mediated by L1 and L2; L1 interacts with heperin sulphate molecules, as well as laminin, on the surface of the basal lamina (Combita et al., 2001; Giroglou et al., 2001; Finnen et al., 2003; Johnson et al., 2009; Culp et al., 2006; Joyce et al., 1999). L1 binding induces structural changes in the virion capsid, exposing L2 for cleavage by furin and thus facilitating virus interaction with a secondary receptor on the keratinocyte surface and subsequent internalization of virus (Richards et al., 2006; Johnson et al., 2009; Kines et al., 2009; Surviladze et al., 2012). The mechanism of virus internalisation has not been fully resolved for all HPV types and findings are disparate. Most recently, it has been suggested that entry of HPV types 16, 18
Figure 1.3: HPV genome organisation
The HPV genome has a circular double stranded DNA genome, encoding 8 ORFs. Splicing of mRNA creates a number of gene products. Viral genes are transcribed in a clockwise direction to produce 6 early, non structural proteins (E1, E2, E4, E5, E6 and E7), and 2 late transcripts encoding structural proteins (L1 and L2, light blue). The major oncoproteins E6 and E7 are highlighted in green, and the transcriptional control region (long control region; LCR, highlighted in grey). The LCR contains a DNA replication origin and functions as the regulator for the DNA replication. This region contains binding sites for cellular transcription factors (e.g. SP1, AP1, Oct1), and for the viral E2 protein, which control viral gene expression. The HPV genome contains two promoter elements (P Early and P Late, indicated in red) which regulate the expression of differentially-spliced mRNAs. The genome also contains two polyadenylation sites, (Poly A Early and Late, indicated in red).
Adapted from Kajitani et al. 2012 and Doorbar et al. 2012.
and 31 are all independent of clatherin, caveolin, and dynamin, but require actin polymerization and the tetraspanin CD151 (Spoden et al., 2013).

Following internalisation, the viral capsid disassembles in the endosome compartment and L2 mediates relocation of the viral genome from the endosomal compartment to the nucleus, where genomes are replicated and maintained at a low copy number in basal cells (Richards et al., 2006; Day et al., 2003; Pyeon et al., 2009; Parish et al., 2006; McBride, 2008). The initial round of viral DNA replication occurs independently of the cell cycle, and results in amplification of the viral genome to around 50–100 copies per cell (Moody and Laimins, 2010). Here, the viral transcription factor E2 is important for the initiation of viral replication and segregation of viral genomes into infected daughter cells (Lehman and Botchan, 1998; Sanders and Stenlund, 2000; Mohr et al., 1990). There are multiple E2 binding sites in the viral LCR, which enable E2 to recruit the helicase E1 along with cellular transcriptional machinery to position it at the viral origin of replication and promote HPV gene transcription (Stubenrauch et al., 1996; Mohr et al., 1990; Sanders and Stenlund, 2000). At this stage, HPV genes are transcribed from the early promoter (Figure 1.3). E2 also acts to tether HPV genomes to mitotic chromosomes via interaction with Brd4 and is able to activate or repress cellular gene transcription, depending on the context of the binding sites and nature of the associated cellular factors (Bernard et al., 1989; Dostatni et al., 1991; Dong et al., 1994; You et al., 2004).

Following the initial replication stage, known as establishment, the infected cell leaves the basal compartment and migrates upwards in the epithelia. During a normal, uninfected, life cycle, basal keratinocytes divide and migrate upwards, differentiating as they do so (Figure 1.4). Basal cells divide asymmetrically; one daughter cell migrates, leaving a basal layer of dividing epithelial stem cells (Figure 1.4). However, entering the differentiation program requires exit from the cell cycle and poses an obstacle to HPV genome replication, which
The right-hand side depicts the structure of uninfected differentiated epithelia and the main characteristics of each cell layer. Basal cells in the cervical epithelium rest on the basal lamina, which is supported by the dermis. Basal cells divide asymmetrically, and a daughter cell migrates upwards as differentiation occurs, in which cells exit the cell cycle and progressively accumulate cytoplasmic keratin.

The left-hand side depicts the HPV infectious cycle. Initially, HPV infects basal cells via microabrasions in epithelia. Infected basal cells form the reservoir of virus. In these cells, the early genes E1, E2, E4, E5, E6 and E7 are expressed and the viral DNA establishes as episomes within the nuclei of cells. As cells migrate upwards, the viral episomes are maintained at a copy number of 100-1000 genomes per cell. The late productive life cycle occurs in upper layers of the epithelia, where the viral genome is amplified, and the late genes L1 and L2, and E4 are expressed. Virions are assembled and released from cornified keratinocytes as they slough off, and are able to produce a new infection.
depends upon host replication substrates and machinery. To overcome this, HPV expresses the early proteins E6 and E7. The combined actions of these proteins reactivates cellular DNA synthesis in non-dividing cells, inhibits apoptosis, and uncouples the differentiation program of the infected keratinocyte from the cell cycle (Section 1.4.1).

Table 1.4: High-risk HPV ORFs

<table>
<thead>
<tr>
<th>HPV ORF</th>
<th>Function(s)</th>
<th>Target Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>E6</td>
<td>Reactivation of cellular replication mechanisms</td>
<td>p53, ADA3, p300/CBP, E6AP, SP1, c-myc, NFX1-91, TERT, FAK, FADD, Caspase 8, BAX, BAK, IRF3, PDZ domain proteins</td>
</tr>
<tr>
<td></td>
<td>Proliferation, immortalisation, inhibition of apoptosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maintenance of viral genome</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Co-operates with E7 to immortalise primary keratinocytes</td>
<td></td>
</tr>
<tr>
<td>E7</td>
<td>Reactivation of cellular replication mechanisms</td>
<td>RB, p107, p130, HDAC, E2F6, p21, p27, CDK/cyclin, ATM, ATR, gamma-tubulin</td>
</tr>
<tr>
<td></td>
<td>Proliferation, immortalisation, inhibition of apoptosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maintenance of viral genome</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Binds retinoblastoma protein (pRb), deregulates the G1/S checkpoint</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Co-operates with E6 to immortalise primary keratinocytes</td>
<td></td>
</tr>
<tr>
<td>E4</td>
<td>Interacts with cytoskeletal proteins, allows viral assembly</td>
<td>Cytokeratin 8/18</td>
</tr>
<tr>
<td>E5</td>
<td>Weak transforming activity</td>
<td>EGFR, PDGFR, V-ATPase, MHC1, TRAIL receptor, FAS receptor</td>
</tr>
<tr>
<td></td>
<td>Upregulates growth factor receptors</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>Replication of viral genome: Helicase, ATPase, ATP-binding protein</td>
<td>RPA, Topoisomerase, polymerase alpha-primase</td>
</tr>
<tr>
<td></td>
<td>Viral genome replication, transcription and maintenance: Acts as a transcription factor, facilitates initiation of HPV DNA replication, important for genome encapsidation</td>
<td>Brd4, ChlR1</td>
</tr>
<tr>
<td>E2</td>
<td></td>
<td>HPV E1</td>
</tr>
<tr>
<td>L1</td>
<td>Major capsid function</td>
<td>-</td>
</tr>
<tr>
<td>L2</td>
<td>Minor capsid function</td>
<td>-</td>
</tr>
</tbody>
</table>

Adapted from Stanley, 2001 and Kajitani et al., 2012.

Next follows a phase where viral gene expression is minimal and, importantly, the expression of the oncogenes E6 and E7 is tightly control by E2 (Gammoh et al., 2009). As infected keratinocytes enter the differentiating compartment of the epithelia, there is an expansion in viral gene expression, including abundant expression of E6 and E7 (Chow et al., 1987) (Figure 1.4). Transcription occurs from the differentiation-dependent late
promoter (P late) located within the E7 ORF, and terminates at the late polyadenylation site (Figure 1.3), amplifying viral genomes to more than 1000 copies per cell (Lambert, 1991; Bedell et al., 1991). This promoter drives the expression of E1, E2, E4 and E5 proteins as well as the late proteins L1 and L2 (Hummel et al., 1992; Ozbun and Meyers, 1998; 1997), Figure 1.4). Both E1^E4 (a fusion protein comprising the first amino acids of E1 and the E4 ORF) and E5 proteins contribute to this differentiation-dependent phase of the viral life cycle (Wilson et al., 2005; Peh et al., 2004). Cleavage of E1 by cellular caspases and activation of the ATM (ataxia telangiectasia mutated protein)-mediated DNA damage pathway are necessary for viral genome amplification as cells enter G2 (Moody et al., 2007; Moody and Laimins, 2009). It is suggested that the subsequent induction of the homologous recombination pathway recruits cellular DNA repair factors to facilitate HPV DNA replication, thus enabling the virus to replicate in an S phase independent manner, in the differentiating compartment of stratified epithelia (Gillespie et al., 2012).

Little is known about the actions of E5 during the HPV life cycle, although it has been shown to downregulate MHC class I expression to promote viral immune evasion, as well as activating epidermal growth factor (EGF) signal cascades to induce keratinocyte proliferation (Petti et al., 1991; Belleudi et al., 2011; Pim et al., 1992; Crusius et al., 1998; Straight et al., 1993). The sequences of the high-risk HPV E5 proteins are conserved, and it has been shown to have weak transforming properties (Bouvard et al., 1994a; Straight et al., 1993; Stöppler et al., 1996). The E5 protein has also been shown to form pore structures, thought to allow the maturation of endosomes during viral entry and/or release (Wetherill et al., 2012). Similarly, the function of E4 in the upper layers of the epithelia has not been fully elucidated although it is abundantly expressed (Figure 1.4). The E4 protein is translated from spliced transcripts which includes the first 5 amino acids of E1 to generate E1^E4 fusion proteins (Chow et al., 1987). Several studies have shown that E4 is able to
arrest transition of the cell cycle from G2 to M, and inhibit entry into S phase as well as inhibiting cellular DNA replication (Knight et al., 2004; Davy et al., 2009; Peh et al., 2004). However Knight et al. (2011) showed that G2 arrest was not associated with viral genome amplification. Together, the functions of E4 produce an environment which is suitable for viral DNA replication without competition from the host, as well as optimising viral genome amplification and the expression of capsid proteins (Wilson et al., 2007; 2005; Nakahara et al., 2005).

In the final stages of the viral life cycle, the L1 and L2 proteins encapsulate viral genomes to form infectious virus particles (Figure 1.4). Virus capsids are assembled in the nucleus from 360 molecules of L1 and variable number of L2 molecules, arranged into 72 pentameric units to form an icosahedral structure, stabilized by intermolecular disulphide bonds between L1 proteins (Buck et al., 2008; Finnen et al., 2003). Infectious virions are released during the natural sloughing of keratinocytes, a process which may be facilitated by E4-mediated disruption of the keratin cytoskeleton (Roberts et al., 1997; Doorbar et al., 1991). The entire life cycle takes 2-3 weeks in vivo; the time taken for a basal keratinocyte to migrate up the epithelial layers and differentiate (Figure 1.4).

1.4.1 HPV E6 and E7 in the HPV life cycle

The E6 and E7 proteins of both high and low-risk viruses are translated from a bicistronic E6E7 mRNA (Tang et al., 2006). Together they target a number of cellular factors, including those involved in cell cycle regulation, gene expression, DNA replication and cell signalling, culminating in the uncoupling of keratinocyte differentiation from division to promote an environment favorable for viral replication.
1.4.1.1 HPV E7

The functions of E7 induce cellular DNA synthesis, bypass cellular growth arrest, prevent cellular senescence and modulate immune responses.

The interaction of E7 with the retinoblastoma protein (pRb) family is a well-characterised example of E7 acting upon the cell cycle (Figure 1.5). The interaction occurs via a conserved LXCXE motif at the N-terminus of high-risk HPV E7 proteins (Münger et al., 1989b). The pRb family of proteins, consisting of p105, p107 and p130, regulate the transition of the cell cycle from G1 to S phase by acting upon the activity of the transcription factor family E2F (Dyson et al., 1989). In quiescent cells, pRb is hypophosphorylated, which promotes its ability to bind to E2F (Neuman et al., 1994; Hiebert et al., 1992; Morkel et al., 1997). The pRb:E2F heterodimer prevents binding of the transcription factor to the promoter region of genes required for G1 to S phase transition, thus repressing transcription of E2F-responsive genes (Jones et al., 1997; DeGregori and Johnson, 2006; Zwicker et al., 1996). In the presence of mitogenic signals, pRb is progressively phosphorylated during G1 by active CDK:cyclin complexes CyclinD1:CDK4, CyclinD1:CDK6 and CDKE:cyclin2, which causes the complex to dissociate thus allowing transcription of E2F-responsive genes. These genes are required in nucleotide synthesis, DNA replication and progression of the cell cycle into S phase. For example, cyclin A is transcribed, which binds to CDK2 to activate the kinase, allowing the phosphorylation of target proteins to activate S phase specific processes, including DNA replication (Jeffrey et al., 1995; Harper et al., 1993). HPV E7 mimics the phosphorylation of pRb, disrupts the pRb:E2F complex and causes proteasomal degradation of pRb (Figure 1.5, Dyson et al., 1989; Huh et al., 2007). This enables the constitutive expression of E2F-responsive genes; inducing infected cells to enter S phase despite absence of mitogenic signals (Cheng et al., 1995; Chellappan et al., 1992).
Figure 1.5: E7 acting upon the cell cycle

E2F transcription factors form heterodimeric complexes with members of the pRb family and regulate the transcription of several genes during the cell cycle. In quiescent cells, pRb is present in a hypophosphorylated form and associates with E2F molecules, thereby inhibiting their transcriptional activity. When quiescent cells are exposed to mitogenic signals, genes encoding the G1 specific D-type cyclins (D1, D2 and D3) are activated. Subsequently, cyclins associate with a cyclin-dependant kinase, CDK4 or 6, transport into the nucleus and phosphorylate pRb in mid-G1 phase. Phosphorylation of pRb releases active E2F complexes and progression into S phase. E7 binding to pRb mimics its phosphorylation, promoting its proteasomal degradation. Thus, E7 expressing cells can enter S phase in the absence of a mitogenic signals.
High-risk $\alpha$-HPV E7 affects CDK2 via a number of mechanisms to promote cell cycle progression from $G_1$ to $S$ phase. E7 binds to the CDK2 inhibitors p21 and p27, preventing their inhibitory activity, and E7 expression induces increased CDC25 transcription; the phosphatase dephosphorylates and consequently activates CDK2 (Zerfass-Thome et al., 1996; Funk et al., 1997; Jones and Münger, 1997; Nguyen et al., 2002). Similarly, E7 upregulates the transcription of cyclin E to promote cell cycle progression and increases transcription of the senescence inhibitor protein DEK (Wise-Draper et al., 2005). Moreover, E7 affects $S$ phase-specific gene expression by interacting with transcription factors; E7 binds to E2F to act as a transcriptional activator (Hwang et al., 2002).

The viral protein is able to modulate transcription directly via the zinc finger domain. This domain allows E7 to bind and recruit chromatin modifiers such as histone deacetylases (HDACs) to repress the activity of specific promoters, including E2F responsive ones (Brehm et al., 1998; Longworth et al., 2005; Longworth and Laimins, 2004a). Furthermore, E7 is able to facilitate the removal of HDAC at other promoters, in order to activate transcription of growth-promoting genes (Brehm et al., 1999). This interaction is important in $S$ phase promotion but has also been shown to influence the maintenance of episomal genomes (Longworth et al., 2005; Longworth and Laimins, 2004b).

The transcriptional upregulation of interleukin-8 (IL-8) induced by E7 not only prevents cellular senescence, but also enables infected cells to evade immune detection (Walker et al., 2011). Furthermore, E7 expression has recently been shown to increase the levels IL-18 binding protein (IL-8BP), an antagonist of the proinflammatory cytokine IL-18 (Richards et al., 2014). This enables infected keratinocytes to evade immune detection by interfering with IL-18-mediated lymphocyte activation (Richards et al., 2014). The multifarious nature of E7 can also be seen in the inhibition of tumour necrosis factor $\alpha$ (TNF$\alpha$)-mediated cell cycle arrest by the oncoprotein, where the proteins ability to interrupt translation also
inhibits TNFα-mediated differentiation and apoptosis and aids immune evasion (Basile et al., 2001; Thompson et al., 2001; Boccardo et al., 2010). All of which enables the virus to evade detection as well as promote cell survival despite the activation of apoptotic signals. During the viral life cycle, E7 expression and activity is regulated by E2; HPV16 E2 protein directly binds to E7 to inhibit degradation of pRb (Gammoh et al., 2009). Disruption of the cell cycle by E7 induces p53-dependent apoptosis and stabilises p53; degradation of pRb leads to an increase p53 protein and would trigger apoptotic pathways and end viral propagation, were it not for the functions of E6 (Demers et al., 1994).

1.4.1.2 HPV E6

Full-length E6 protein is transcribed from the bicistronic E6E7 mRNA to produce a protein of approximately 150 amino acids in length, which resides in the cytoplasm and nucleus of infected cells (Howie et al., 2009; Tao et al., 2003). It has been suggested that E6 forms a dimer in high concentrations of salt and protein but remains predominantly monomeric under physiological conditions (Lipari et al., 2001). Like E7, E6 contains two zinc finger motifs (Figure 1.6) and the protein is known to interact with a number of cellular targets, via which it affects a range of cellular processes, including apoptosis, cellular differentiation, cell-cell adhesion, polarity and proliferation (Figure 1.7, Barbosa et al., 1989; Grossman and Laimins, 1989).

The association between HPV E6 proteins and the cellular ubiquitin ligase E6AP occurs via an LXLL motif within the oncoprotein and allows E6 to promote the proteasomal degradation of a number of targets, as well as increasing E6 stability (Huibregtse et al., 1993b; Tomaić et al., 2009; Daniels et al., 1998b). One of the most well known targets is the DNA damage response protein, p53, which ordinarily is not targeted for degradation via E6AP (Camus et al., 2003). However, E6 binding both E6AP and p53 facilitates
Figure 1.6: Structure of E6

(A) Schematic of HPV E6 protein domain structure. E6 proteins are approximately 150 amino acids in length and contain two zinc-finger domains and a C terminus PDZ binding motif (PBM, highlighted in pink). Numbers correlate to the amino acid sequence of HPV18 E6.

(B) The 3D structure of E6 from Zanier et al. 2013, shows that the PBM (pink) is disordered in NMR structural studies and adopts a β-strand confirmation upon binding to the PDZ domain of MAGI-1 (red ribbon structure). The right-hand image shows the protein rotated by 180 degrees relative to the left-hand image. Region in green indicates E6AP binding (LXXLL) motif, with the interacting helix of E6AP shown (green). Blue indicates region of E6 involved in self-association, with a ribbon stricture indicating a second E6 molecule. The white region indicates residues which are potentially available to interact with p53.
Figure 1.7: E6 interacts with a number of cellular proteins

From Howie et al. 2009. E6 alters numerous cellular pathways through the binding of proteins. The interactions predominantly function to promote cell-cycle progression for the propagation of virus, but promote transformation of infected cells when the protein is deregulated, aiding the progression of cancer and highlighting the role of E6 as a viral oncoprotein. Examples of proteins involved are indicated below the affected pathways.
ubiquitination and subsequent degradation of the cellular tumour suppressor (Huibregtse et al., 1993a; Werness et al., 1990; Crook et al., 1991; Muench et al., 2010). Interestingly, the association between HPV16 E6 and E6AP is stronger than between E6AP and HPV18 E6, and drives the degradation of p53 more efficiently than HPV18 E6 (Scheffner et al., 1990). E6 proteins also disrupts the function of p53 directly; binding to the protein interferes with the DNA-binding activity and blocks transcription of p53-dependent genes (Lechner and Laimins, 1994). E6 further disrupts p53 activity by binding to the histone acetyltransferases p300 and CREB binding protein (CBP), inhibiting p53 acetylation, which would otherwise stabilize the cellular protein, and increasing E6 stability (Patel et al., 1999; Zimmermann et al., 1999). A decrease in p53 provides the infected cell with resistance to apoptosis; E6-expressing cells are able to continue promoting cell proliferation without disruption. In this respect, the HPV oncoproteins E6 and E7 act synergistically; where E7 puts the cell at risk of p53-mediated apoptosis, E6 intervenes removing p53 and allowing the cell to continue providing substrates for DNA synthesis and viral propagation (Figure 1.8). Moreover, loss or abrogation of E6-directed p53 degradation impairs the ability of HPV16 to maintain genomes and silencing E6 in the context of HPV18 negatively affects viral genome amplification (Park and Androphy, 2002; Wang et al., 2009c; 2009b; 2009a; Kho et al., 2013). Similarly, HPV11 encoding a non-functional E6 were unable to maintain episomal genomes, indicating the importance of HPV E6 in viral replication (Oh et al., 2004).

HPV E6 is also able to affect p53-independent apoptotic pathways to prevent virally induced cell-death. This was first observed in transgenic mouse models expressing HPV16 E6 and E7 in the mouse eye lens; expression of E7 alone induced apoptosis, which was partially rescued by knocking down p53 (Pan and Griep, 1994; Griep et al., 1993). Although both high and low-risk HPV E6 proteins promote the degradation of pro-apoptotic factor Bak, only high-risk E6 has been shown to bind to procaspase 8, an event which induces
Figure 1.8: HPV oncoprotein cooperation in cervical carcinogenesis

The ability of E6 and E7 to target cellular regulators of proliferation, apoptosis, immortalization and genomic stability collectively bestows cells with a growth advantage and an increased propensity for transformation and malignant progression. The viral proteins work synergistically; the induction of hyperproliferation by the E7 protein triggers cellular apoptotic pathways, which is blocked by the actions of the E6 protein. Cells are therefore immortalized via the cooperative actions of E6 and E7, although these functions are potentiated by the actions of E5. Adapted from Moody and Laimins 2010.
conformational changes that interrupt downstream signalling and the cells response to pro-
apoptotic stimuli (Thomas and Banks, 1999; 1998; Underbrink et al., 2008; Tungteakkhun
and Duerksen-Hughes, 2008; Filippova et al., 2004).

Table 1.5: Protein partners of E6

<table>
<thead>
<tr>
<th>Function</th>
<th>E6 Binding partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>E3 ubiquitin ligase</td>
<td>E6AP</td>
</tr>
<tr>
<td>Mediators of apoptosis</td>
<td>p53, Bak, c-myc, procaspase-8, FADD, TNFR1</td>
</tr>
<tr>
<td>Transcriptional regulation</td>
<td>CBP/p300, NFκB</td>
</tr>
<tr>
<td>Immune regulation</td>
<td>IRF3</td>
</tr>
<tr>
<td>Chromosomcal stability</td>
<td>Mcm7, hTERT</td>
</tr>
<tr>
<td>Epithelial differentiation and organisation</td>
<td>E6-BP, fibulin-1, zyxin, Notch pathway</td>
</tr>
<tr>
<td>Cell polarity, adhesion and proliferation</td>
<td>PDZ proteins*</td>
</tr>
</tbody>
</table>

Adapted from Howie et al., 2009; Vande Pol and Klingelho, 2013. *Cell polarity proteins are discussed in more detail in Section 1.6

During HPV infection, keratinocyte differentiation must be tightly controlled in order to
balance the promotion of viral genome amplification in the spinous layer of epithelia with
keratinocyte differentiation in the granular and cornified layers (Figure 1.4). High-risk E6
proteins modulate keratinocyte differentiation and it is likely that these functions of E6 are
temporally regulated (Alfandari et al., 1999; Sherman et al., 1997). In cell culture and mouse
models, E6 has been shown to enhance the activity of the Wnt signalling pathway, which is
involved in cell proliferation, in an E6AP dependent manner (Lichtig et al., 2010; Bonilla-
Delgado et al., 2012). Moreover, E6 activates the mTOR receptor tyrosine kinase cascade
by enhancing phosphorylation of the kinase mTOR which activates downstream signalling
pathways, culminating in the stimulation of protein synthesis (Spangle and Münger, 2010;
Spangle et al., 2012). The oncoprotein also acts on the Notch pathway, which is central to
epithelial differentiation; cutaneous HPV E6 proteins are able to downregulate the pathway
by binding to MAML1 protein, and E6 proteins have been shown to repress pathway
components transcriptionally (Brimer et al., 2012; Meyers et al., 2013; Henken et al., 2012; Tan et al., 2012). However, HPV has been shown to potentiate the Notch pathway in keratinocytes via transcription factor NFX-123 (Vliet-Gregg et al., 2013; 2015). Interestingly, a recent study indicated that NFκB activation by E6 is cell context specific; whereas E6-mediated activation of the pathway in ectocervical cells increased proliferation, it inhibited growth in cells derived from the endocervix (Vandermark et al., 2012). Further complexity is added by the addition of a temporal aspect; it is likely that E6 differentially modulates Notch signalling during the viral life cycle, repressing and enhancing signalling in the spinous and upper layers respectively (Brimer et al., 2012; James et al., 2006; Nees et al., 2001; Yuan et al., 2005; Havard et al., 2005; D'Costa et al., 2012; Vandermark et al., 2012). In addition, the C-terminus of high-risk α-HPV E6 proteins contains a conserved motif that is able to interact with a number of PDZ-domain containing proteins involved in differentiation and the regulation of epithelial polarity. The relationship between high-risk HPV E6 and these polarity proteins is discussed in more depth in Section 1.5.4.

By interacting with transcription factors, such as p300/CBP and c-myc, E6 is able to influence the regulation of gene transcription (Crook et al., 1991; Desaintes et al., 1992; Ronco et al., 1998; Patel et al., 1999). By activating c-myc, high-risk HPV E6 is able to regulate the expression of hTERT a telomerase complex which prevents telomeric DNA loss during repeated cell replication cycles (Oh et al., 2001; Kiyono et al., 1998). This is beneficial to the viral life cycle as it extends the replicative potential of infected cells, however chronic activation leads to an immortal phenotype, promotes chromosomal instability and mutation accumulation, all of which are oncogenic phenotypes (Figure 1.5). Furthermore, the expression of HPV16 E6 causes downregulation of a number of genes involved in keratinocyte differentiation, including IFN-responsive genes, including IFNα and IFNβ (Nees et al., 2001; Vandermark et al., 2012). The same study observed an
upregulation in expression of proliferation promoting genes, including IL-8 and NFκB pathway components, in the presence of E6 alone (Nees et al., 2001). Moreover, it has recently been observed that expression of high-risk E6 alone is sufficient to cause transcriptional upregulation of *APOBEC3B*, which encodes a DNA deaminase (Vieira et al., 2014). It has been suggested that this is an effect of E6-mediated p53 abrogation, rather than a direct interaction with cellular DNA, but it must be noted that E6 is able to effect gene expression more directly. For example, E6 proteins have been shown to bind and target the HAT ADA3 for degradation; as such, interaction with chromatin remodeling machinery is a mechanism by which E6 can alter host gene expression (Kumar et al., 2002). The effect of E6 upon transcription contributes to the proteins ability to delay cell differentiation and promote a favourable environment for viral genome replication.

Another feature of high-risk HPV E6 proteins is their ability to regulate cellular microRNA (miRNA) expression (Martinez et al., 2008; Wang et al., 2009c; Wald et al., 2011; McKenna et al., 2014; Yablonska et al., 2013). For example, HPV16 E6 downregulates miR34a, which targets cell cycle control genes (Wang et al., 2009c). The mechanism by which E6 regulates miRNAs is not known, but it has been suggested that the reduction in p53 and the increase in E2F activity when E6 and E7 are present plays a role (Yablonska et al., 2013).

1.4.1.2.1 Spliced forms of E6

Variations of the E6 protein, denoted E6*I and E6*II, are produced via alternative splicing (Smotkin and Wettstein, 1986). The splice site is highly conserved, and causes expression of the first 42 or 44 amino acids of HPV16 or HPV18 E6 respectively followed by a variable sequence (Comerford et al., 1995). The functions of E6 spliced variants remain unclear, although interactions with both E6AP and full-length E6 have been observed (Pim et al., 1997). The overexpression of HPV18 E6*I promotes the proteasome-dependent degradation of junctional proteins DLG1, MAGI-1 and PATJ which are also targets of
full-length E6 (Pim et al., 2009; Storrs and Silverstein, 2007). Moreover, both HPV16 and HPV18 E6*I have been shown to target the HAT TIP60 for degradation (Jha et al., 2010). Conversely, the kinase AKT has been identified as a target of E6*I, but is not targeted by the full-length protein (Pim et al., 2009).

The actions of E6*I appear to antagonize some functions of the full-length protein. For example, in vitro HPV18 E6*I is able to inhibit E6-mediated p53 degradation and stabilize procaspase 8 (Pim et al., 1997; Tungteakkhun et al., 2010; Pim and Banks, 1999). It has been suggested that the antagonistic functions of E6* controls the activity of full-length E6 during the virus life cycle by preventing complete removal of cellular p53 (Mantovani and Banks, 2001). Interestingly, the full-length and E6*I versions bind to different regions of procaspase 8 (Tungteakkhun et al., 2010; Filippova et al., 2007).

1.4.2 HPV oncogenesis

HPVs are well adapted to their host; in most instances the virus manages to proliferate and is maintained in the host population without causing any apparent disease (Woodman et al., 2007; Liaw et al., 1999; Wallin et al., 1999). However, malignant growth can be a side effect of infection, with the time from initial exposure to invasive growth spanning a period of 15-30 years (Woodman et al., 2007; Bekkers et al., 2004). The progression of infected cells from neoplastic lesions to invasive cancer is associated with failure of the immune system to clear persistent infection and the integration of the HPV genomes into host chromosomes (Cricca et al., 2006; Andersson et al., 2005; Peitsaro et al., 2002; Berumen et al., 1995; Fuchs et al., 1989; Cooper and McGee, 1992; Das et al., 1992; Hausen, 2002; Collins et al., 2009). Although integrated HPV genomes are present in the majority of cervical cancers and the frequency of integration increases with severity of cervical neoplasia, it should also be noted that integration of viral genomes is an early event (Wentzensen et al., 2004; Xu et al., 2013; Collins et al., 2009). While no integration hotspots
have been confirmed, it has been suggested that preferential integration occurs at a site near the c-myc gene and that genomes integrate at naturally occurring regions of host genomic instability, known as fragile sites (Ziegert et al., 2003; Ferber et al., 2003; Akagi et al., 2014). Integration of the genome is influential in HPV oncogenesis; in epithelial cell culture, integration of HPV16 genomes is associated with a selective growth advantage and an increased proliferative capacity (Jeon et al., 1995; Jeon and Lambert, 1995). HPV genome transcription is regulated by the viral E2 protein, which represses early gene expression and regulates copy number (Bouvard et al., 1994b). Integration disrupts the viral genome, often interrupting E1 and E2 expression and thus removing the regulatory influence of E2 over the expression of early viral genes (Pett and Coleman, 2007). This was demonstrated in culture, when reintroduction of E2 into HPV18-positive cancer cells containing integrated genomes caused cellular senescence (Goodwin and DiMaio, 2000; Francis et al., 2000). Furthermore, the expression of E6 and E7 oncoproteins is retained in cancers, and the transcripts expressed from integrated genomes are more stable than from episomes (DeFilippis et al., 2003; Hausen, 1999; Schwartz, 2013; Jeon et al., 1995; Jeon and Lambert, 1995). Replication of integrated genomes also results in the activation of host DNA repair and recombination systems, which increases the likelihood of acquiring cellular mutations, increased genomic instability and, eventually, malignant progression (Kadaja et al., 2009). It has also been suggested that the coexistence of episomal genomes with integrated copies may be important in carcinogenesis; Kadaja et al. (2009) propose that expression of HPV E1 and E2 from episomal genomes can initiate DNA replication from integrated viral origins, resulting in the induction of chromosomal abnormalities as well as viral gene transcription, leading to genomic instability. Thus, there is a significant role for HPV genome integration in the transformation of infected cells to cervical cancer, and the viral oncoproteins E6 and E7 are important in maintaining a transformed phenotype.
1.4.2.1 HPV E6 and E7 in oncogenesis

There is a wealth of evidence demonstrating the ability of HPV oncoproteins E6 and E7 to drive immortalisation of cells. Expression of the E6 protein from high-risk HPV types 16, 18, and 31 have been shown to extend keratinocyte lifespan (Münger et al., 1989a; Hudson et al., 1990; Sedman et al., 1991; Woodworth et al., 1989). Expression of high-risk HPV E7 immortalises human primary keratinocytes in culture at low frequency, but is unable to immortalise mammary cells alone (Hawley-Nelson et al., 1989). Moreover, mouse models of cervical cancer have shown that E6 and E7 are able to induce dysplasia independently, with E7 expression generating a more invasive phenotype than E6 (Riley et al., 2003). Mice expressing E6 under the control of a keratin promoter develop epithelial hyperplasia (Herber et al., 1996; Song and Pitot, 1999). In this transgenic mouse model, expression of E7 alone is sufficient to induce high-grade cervical dysplasia and invasive cervical malignancies (Arbeit et al., 1996). Similarly, expression of HPV16 E7 in transgenic embryonic mouse lens led to induction of hyperplastic growth (Howes et al., 1994; Pan and Griep, 1994). When E6 was induced alongside E7 in this system, an increased susceptibility to tumour development was observed (Howes et al., 1994; Pan and Griep, 1994; 1995). In raft culture, expressing both E6 and E7 induced cellular changes are comparable to changes seen in the development of CIN (McCance et al., 1988). Although the oncoproteins have transformative properties alone, concurrent expression produced larger and more extensive cancers, highlighting the cooperative nature of E6 and E7 in driving tumour development (Riley et al., 2003).

Despite the ability of high-risk E6 and E7 to immortalise primary cells efficiently, HPV immortalised cells are not tumorigenic in nude mouse models (Hawley-Nelson et al., 1989; Münger et al., 1989a; Kaur and McDougall, 1988). In order to develop tumourigenic capabilities, HPV-immortalised cells require the coexpression of additional oncogenes,
such as v-ras, or extensive passaging in tissue culture (Durst et al., 1989; Pei et al., 1998; Hurlin et al., 1991). Transgenic mice expressing high-risk E6 and E7 in epithelial cells develop squamous carcinomas when treated with oestrogen, highlighting hormone levels as a cofactor in cervical cancer development (Arbeit et al., 1996). Furthermore, murine experiments have suggested that, during the development of HPV-associated cancers, E6 contributes to the later stages of oncogenesis, whereas the functions of E7 contribute to the early stages (Song et al., 2000; Simonson et al., 2005).

The same functions that promote cell growth during the HPV life cycle, when deregulated, allow E6 and E7 to drive cell proliferation and survival, as well as promoting genetic instability and immune evasion, all of which contribute to HPV-driven oncogenesis (Figure 1.8). For example, the activation of hTERT by E6 leads to an immortal phenotype, promotes chromosomal instability and mutation accumulation (Oh et al., 2001). Furthermore, there is a strong correlation between the ability of HPV E6 to activate hTERT and the ability of the HPV type with association cancer (van Doorslaer and Burk, 2012). Interestingly, whilst low-risk HPV E6 proteins are able to associate with E6AP as well as high-risk types, this does not result in p53 degradation, indicating that this interaction plays an important role in E6-driven transformation (Brimer et al., 2007; Lechner and Laimins, 1994). Similarly, low-risk HPV E7 proteins associate with pRb, but are less efficient than high-risk proteins at activating E2F responsive genes (Longworth and Laimins, 2004a). Moreover, only high-risk HPV E7 proteins have been identified as able to bind to HDACs via conserved sequences, indicating that this mechanism of controlling gene transcription is significant to the contribution of E7 to transformation (Figure 1.9).
Figure 1.9: E6 and E7 affect a number of cellular targets in carcinogenesis
By targeting cellular proteins, E6 and E7 are able to affect a number of cellular processes (boxed) culminating in transformation of E6 and E7 expressing cells. Factors targeted by E7 are indicated in purple, at the left-hand side of the cellular processes and factors affected by E6 are indicated in green at the right-hand side. Adapted from Yugawa and Kiyono, 2009.
1.4.2.2 HPV E5 in oncogenesis

Although E6 and E7 provide the primary functions in HPV driven oncogenesis, the functions of E5 have also been shown to contribute to tumour progression (Stöppler et al., 1996). E5 promotion of cell proliferation has been observed in tissue culture when expressed with E6 and E7, and to a lesser extent when expressed alone (Petti et al., 1991; Valle and Banks, 1995; Bouvard et al., 1994a). In transgenic mice expressing HPV16 E5 in the skin, epithelial hyperproliferation is observed, eventually leading to tumour formation and oestrogen treatment of mice induces cervical cancers when E5 is expressed (Maufort et al., 2010; 2007). It is suggested that the protein provides a supporting role in tumour progression and the involvement of this protein during the HPV life cycle and oncogenesis has yet to be fully elucidated.

1.5 Epithelial polarity

Epithelial tissues comprise the foundation of the majority of human organs, and 90% of cancers are epithelia-derived carcinomas (Tanos and Rodriguez-Boulan, 2008). Epithelia line the cavities of the body and are specialised to perform secretion, absorption, protection and sensory functions. These functions depend largely upon the establishment of tissue polarity and integrity of cell-cell attachments, which enable cells to form a sheet-like structure, and to distinguish between the surfaces of epithelial sheets (Figure 1.10). Individually, all cell types exhibit polarity, in that structures and membrane domains are organised asymmetrically. Such cell shape is fundamental during the development of an organism, in defining tissue architecture and maintaining tissue homeostasis. The deregulation of mechanisms regulating cell polarity can lead to developmental disorders, tissue degeneration or cancer (Humbert et al., 2003).
Figure 1.10: Epithelial polarity and junction complexes

(A) Epithelial differentiation, as depicted in Figure 1.4. Basal cells rest on the basal lamina, which is supported by the dermis, and divide asymmetrically. A daughter cell migrates upwards as differentiation occurs, in which cells exit the cell cycle and progressively accumulate cytoplasmic keratin.

(B) Cells themselves are polar and epithelia also exhibits planar and apical-basal polarity, allowing the trafficking of molecules and specialized functions on different faces of the tissue.

(C) There are three main complexes involved in the organization of epithelial cell-cell junctions. Tight junctions coordinate the zona occludens, a belt-like structure which ties planes of epithelia together.

(D) The Crumbs complex (red) consists of PATJ, PALS and Crumbs proteins. Tight-junctions (green) involve the Par complex, made up of PAR3, PAR6 and atypical PKC, as well as associated tight junction PDZ-domain containing proteins MAGI-1, MUPP1, ZO-1 and ZO-2. Adherens junctions (purple) contain the Scribble complex, made up of SCRIB, DLG1 and LGL.
As a tissue, epithelial polarity requires the vertical organisation of individual cells, such that specialized structures (e.g. adherens junctions) are orientated along the apical-basolateral axis, known as apical-basolateral polarity (ABP), as well as the positioning of cells within the plane of epithelial tissues, known as planar cell polarity (PCP, Figure 1.10 B). Furthermore, epithelial polarity is retained during cell division via asymmetrical replication; in basal cells, orientation of the mitotic spindle enables division to occur in parallel to or perpendicular to the epithelial sheet (Lechler and Fuchs, 2005). Division of basal cells within the plane of the epithelium functions to expand cell number, whereas division perpendicular to the plane produces stratified epithelia (Kim et al., 2009).

The correct regulation of both apico-basal polarity (ABP) and planar cell polarity (PCP) is accomplished via the precise localisation of polarity proteins and other molecules such as lipids (Knust and Bossinger, 2002; Wodarz and Näthke, 2007). Central to ABP regulation are the Scribble, Par (Partitioning defective) and Crumbs protein complexes (Yamanaka and Ohno, 2008; Assémat et al., 2008, Table 1.6). These complexes link cell polarity with the regulation of cell–cell contact via junctional complexes, as well as with regulation of cell proliferation, migration and apoptosis. The main junctional complexes involved in epithelial organisation and integrity are the tight junctions (TJs) and the adherens junctions (AJs), formed at the apical and basolateral domains of a cell, respectively (Table 1.6, Figure 1.10).

1.5.1 Polarity complexes

In order to distinguish between the outer and inner surfaces of epithelial sheets, the membrane of epithelial cells is segregated into apical and basolateral domains with an asymmetric distribution of trafficking machinery and membrane proteins (St Johnston and Ahringer, 2010). Within a cell, the apical and basolateral regions are separated by the zonula adherens (ZA), a belt-like structure circling the apical side of the cell (Figure 1.10 C). The
apical and basolateral membrane regions are defined by the distinct lipid components, creating distinct functions, and polarity is further determined and regulated by a number of interacting protein complexes forming scaffolds, signalling complexes and trafficking systems (Figure 1.10).

Table 1.6: Cell polarity and adhesion complex components

<table>
<thead>
<tr>
<th>Complex</th>
<th>Components</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell polarity complexes</td>
<td>Scribble, CRIB, DLG1 and LGL</td>
<td>Apical-Basal polarity</td>
</tr>
<tr>
<td></td>
<td>Crumbs, CRB, PALS1 and PATJ</td>
<td>Apical plasma membrane</td>
</tr>
<tr>
<td></td>
<td>Par, PARD3, PAR6, CDC42 and aPKC</td>
<td>Apical-Lateral polarity</td>
</tr>
<tr>
<td>Cell adhesion complexes</td>
<td>Adherens junctions, Cadherins, catenins, nectins and afadins</td>
<td>Adherens junction</td>
</tr>
<tr>
<td></td>
<td>Tight junctions, ZO-1, ZO-2, ZO-3, Occludin, Claudin and JAM</td>
<td>Tight junction</td>
</tr>
</tbody>
</table>

The Scribble complex was first identified in *Drosophila melanogaster* and consists of Scribble (SCRIB), Discs large (DLG1) and Lethal giant larvae (LGL1) proteins. In epithelial cells, the Scribble complex is located at the basolateral region and is required for membrane organisation and maintenance (Figure 1.10, (Woods et al., 1996; Yates et al., 2013; Humbert et al., 2012). The Crumbs complex was also identified in *Drosophila*. The complex consists of Crumbs, Pals1 (Protein associated with Lin seven 1) and Pals1-associated tight junction (PATJ, also known as INAD1) proteins (Figure 1.10 D). The partitioning defective (Par) complex was initially identified in *Caenorhabditis elegans*, and in humans consists of the proteins PARD3, PAR6 and an atypical protein kinase C (aPKC, Figure 1.10 D, (Kemphues et al., 1988). The Par complex is localised to the apical region of epithelial cells and, together with Crumbs, regulates the maintenance of the apical membrane. The complexes do not exist in isolation; Par and Crumbs act in a mutually antagonistic fashion.
with the Scribble complex (Bilder et al., 2003). In the Par complex, aPKC negatively regulates LGL in the Scribble complex by phosphorylation, which prevents the protein’s apical localisation (Betschinger et al., 2003). Conversely, the kinase is required for the stabilization of the Crumbs complex, possibly via PAR6 binding or directly by phosphorylation of Crumbs proteins (Sotillos et al., 2004).

Both the polarity and junctional complexes of epithelia involve proteins containing the PDZ (PDZ/PSD/DLG/ZO) interaction domain, forming polarity complexes and defining the polarity required for epithelial functions.

1.5.2 The PDZ domain

PDZ domains are one of the most widely distributed protein-protein interaction domains; the human genome encodes over 320 PDZ-domain containing proteins (Fan and Zhang, 2002). The PDZ domain is an evolutionarily conserved domain of approximately 90 amino acids folded into six β-sheets and two α-helices (Figure 1.11). A groove between the second β-strand and the second α-helix forms a hydrophobic pocket that interacts with specific C-terminal sequence motifs of target proteins (Songyang et al., 1997; Niethammer et al., 1998). The structure of the PDZ domain does not change upon ligand binding; the PDZ binding motif (PBM) peptide serves as an extra β-strand and participates in the extensive hydrogen-bonding pattern with main chain PDZ domain residues (Doyle et al., 1996). Crystal structures of PDZ domains in complex with peptides reveal that a loop at the end of the binding groove, which contains a conserved motif (R/K-X-X-X-G-L-G-F), is involved in positioning the target peptide within the binding groove (Figure 1.11). In the PBM:PDZ complex, the terminal COOH of the PBM is coordinated by a network of hydrogen bonds to main-chain amide groups in this loop (Tochio et al., 1999; Daniels et al., 1998a). When bound to the PDZ domain groove, the side chain of the terminal and -3
Figure 1.11: The PDZ domain structure
The secondary (A) and tertiary (B) protein structure of the PDZ domain. A groove between the second $\beta$-strand ($\beta B$) and the second $\alpha$-helix (indicated with *) forms a hydrophobic pocket that interacts with specific target proteins. The ten core domain binding sites are highlighted in blue and the bound peptide (orange). A schematic of the bond formation between residues in the PDZ binding groove and peptide is shown in C. Residues in the binding pocket are shown in black and the peptide is shown in blue. Hydrogen bonds are indicated by red dotted lines, and hydrophobic packing is indicated by green arcs. Adapted from Subbaiah et al. 2011, Hui et al. 2013 and Harris and Lim, 2001.
residues of the PBM peptide point towards the base of the peptide-binding groove. As such, interactions with these residues provide much of the PBM binding specificity (Songyang et al., 1997; Doyle et al., 1996). For example, the preference of the first PDZ domain of NHERF1 for a C-terminal PBM leucine is attributed to its large hydrophobic pocket (Karthikeyan et al., 2001). In some cases, recognition of the PBM involves residues outside of the four amino acid consensus motif, enabling specificity of interactions unique to individual PDZ domains (Niethammer et al., 1998). In terms of PBM sequence, there are three classes, with different C-terminal consensus sequences (Table 1.7). Although PDZ domains are mainly known for their ability to bind to the short C-terminal peptides of their target proteins, they have also been shown to recognize internal sequences that form a beta hairpin fold, other PDZ domains, and membrane phospholipids (Zimmermann et al., 2002; Zhang et al., 2009; Wong et al., 2003; Hillier et al., 1999; Brenman et al., 1998; Gee et al., 1998; Penkert et al., 2004). Lipid binding further enables PDZ proteins to locate to specific membrane domains lipid rafts. For example, Syntenin-1 binding to phosphoinositide containing lipids localises the PDZ protein to the cell membrane and phosphatase and tensin homologue (PTEN) binding to membrane lipids aids the clustering of the Par complex proteins to cell membrane junctions (Feng et al., 2008).

### Table 1.7: Classes of PDZ-binding motif

<table>
<thead>
<tr>
<th>Class</th>
<th>Consensus</th>
<th>Position</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>X-(Y_1)-X-(Y_2)</td>
<td>C-terminal</td>
<td>DLG4 binding shaker type K(^+) channel (Kim et al., 1995)</td>
</tr>
<tr>
<td>Class II</td>
<td>X-Y-X-(Y_2)</td>
<td>C-terminal</td>
<td>CASK binding Neurexin (Songyang et al., 1997)</td>
</tr>
<tr>
<td>Class III</td>
<td>X-X-C</td>
<td>C-terminal</td>
<td>SITAC binding L6 antigen (Borrell-Pagès et al., 2000)</td>
</tr>
</tbody>
</table>

Adapted from Harris and Lim, 2001.

As a protein-protein interaction module, PDZ-domain containing proteins often comprise more than one PDZ domain and/or more than one type of interaction module (Figure 1.12). DLG1 encodes three PDZ domains alongside an SH3, L27 and GUK domains, all
of which are involved in protein-protein interactions and, by contrast, MUPP1 consists solely of 13 PDZ domains (Figure 1.12). The presence of modular domains within a single protein allows PDZ proteins to form interactions with many partner proteins simultaneously and act as a scaffold; organising large protein complexes at specific subcellular locations (Giallourakis et al., 2006).

Cell-cell junction complexes comprise a number of PDZ domain containing proteins (Figure 1.10). These proteins form complexes at the cell membrane and form interactions with junctional complexes in adjacent cells, providing integrity and connecting extracellular space to intracellular components, allowing communication via signalling between the two (Figure 1.10, Laura et al., 2002; Balda and Anderson, 1993; Stucke et al., 2007; Yates et al., 2013). For example PATJ is essential for TJ formation, it is important for localisation of PARD3 to the junction complex, and also in localising PARD3 to the leading edge of migrating cells (Lemmers, 2002). The expression of PATJ in turn is regulated by fellow PDZ protein MUPP1, indicating the wide-ranging effects and diverse roles of PDZ domain containing proteins (Assémat et al., 2013). PDZ domain containing proteins are also involved in the destination and recycling of target proteins; for example, DLG4 is crucial for endocytic recycling of NMDA receptors and β2-adrenergic receptors at the synapse and in muscle cells, affecting the cells response to signals such as adrenaline (Lavezzari et al., 2003; Xiang et al., 2002). As well as acting at the cell membrane, PDZ-containing proteins have been implicated in targeting proteins to intracellular organelles and protein trafficking around the cell. The PDZ interactor with cellular kinase 1 (PICK1) protein targets PKC to mitochondria (Wang et al., 2003), thus affecting the access of the kinase to target proteins and downstream signalling.
Figure 1.12: PDZ domain containing protein domain structure
Schematic to show the organization of several human PDZ domain containing proteins. MAGUK family proteins grouped at the top, contain an inactive guanylate kinase domain. Known PDZ targets of high-risk HPV E6 proteins are indicated by a star (*).

SH3, Src Homology 3; PDZ, PSD/DLG/ZO; GK, Guanylate kinase; L27, Lin2-7; CARD, caspase activation and recruitment domain; CAMK, Ca\(^{2+}\)/Calmodulin dependant protein kinase; ww, tryptophan fold; LRR, Leucine rich repeat; EB, Ezrin binding; PTP, Protein tyrosine phosphatase; FERM, 4.1/Ezrin/Radixin/Moesin.
Thus, there is a huge range and number of interactions that PDZ proteins can form and yet they are able direct interactions with many target proteins in a specific fashion, enabling PDZ domain containing proteins to perform roles in organizing multimeric signalling complexes, forming cell-cell interactions and organizing junctional complexes (Figure 1.10).

1.5.3 Alternative splicing

PDZ protein diversity is further extended via alternative processing of mRNA (Figure 1.13). Alternative splicing generates a variety of protein isoforms expressed from a single gene. Notable examples of PDZ protein splicing include the Erbb2 interacting protein (ERBIN), which has 7 known mRNA variants, including some with a truncated PDZ domain, and the Enigma homolog protein, which has more than 100 predicted variants (Yamazaki et al., 2010; Favre et al., 2001). Differential expression of spliced variants may depend upon cellular context; in the case of PDZ proteins, alternative splicing has been shown to generate isoforms specific to cell context and to developmental stages (Sierralta and Mendoza, 2004). Both MAGI-1 and MAGI-3 encode three spliced variants that differ in the C-terminal sequence and are expressed in a tissue specific fashion; two of the MAGI-1 isoforms are brain specific (Dobrosotskaya et al., 1997; Hirao et al., 2000; Laura et al., 2002). Furthermore, expression of alternative exons not only defines tissue specificity variant, but also cellular localisation; in a cellular context one MAGI-1 isoform (MAGI-1c) is predominantly nuclear, whereas the other two are absent from the nucleus (MAGI-1a and b), when expressed exogenously (Laura et al., 2002). Different spliced isoforms may have different roles or interacting partners in different tissues; whilst all of the PARD3 isoforms associate with tight junctions in epithelial cells, they are differentially expressed in epithelial and haematopoietic cells and have different binding affinities with other Par complex components, PAR6 and aPKC (Gao et al., 2002; Zhou et al., 2008).
Figure 1.13: Patterns of alternative splicing

The inclusion or exclusion of exons can occur in many splicing patterns. Consensus splice site sequence (A) derived from Srebrn et al. 2006, Ghinga et al. 2008 and Harris and Senapathy, 1990. Solid rectangles represent exons, and dotted lines the pattern of inclusion. Intron retention (B), alternative 5’ splice site selection (C), alternative 3’ splice-site selection (D), cassette-exon inclusion or skipping (E) and mutually exclusive exon inclusion (F).
Deregulation of PDZ protein expression can lead to inappropriate isoform expression and promotion of carcinogenesis, for example spliced variants of DLG2 have been identified as transcriptionally upregulated in kidney cancer precursor growths (Zubakov et al., 2006). With such diversity in form and function, PDZ domain containing proteins have significant oncogenic potential when deregulated at a transcriptional, translational or post-translational level. Alternative splicing of DLG1 will be discussed in more depth in Chapter 4.

1.5.4 PDZ domain containing proteins in cancer

Epithelial cancers are characterized by epithelial-mesenchymal transition (EMT, Huber et al., 2005). EMT is caused by a loss of ABP and cell–cell adhesion; cell-to-cell junctions become disrupted so that neighboring cells begin to dissociate from each other, no longer forming tight junctions or the interface that characterises their function (Moreno-Bueno et al., 2008; Huber et al., 2005). Since cell polarity is crucial for tissue integrity and homeostasis, it is logical that loss results in tissue disorganisation and facilitates both initiation and progression of cancer (Figure 1.14). Deregulation of epithelial polarity leads to abnormal cellular trafficking, signalling and migration, all of which contribute to the development of carcinomas and progression to an invasive phenotype (Tanos and Rodriguez-Boulan, 2008; Hanahan and Weinberg, 2011). As PDZ proteins have extensive roles in cell polarity and signalling, their deregulation has been implicated in a number of human diseases (Table 1.8).

In the Scribble complex, a reduction in DLG1 protein expression has been described in the later stages of a number epithelial cancers, including cervical and colon (Watson et al., 2002; Cavatorta et al., 2004; Lin et al., 2004). Furthermore, mutations in the DLG1 gene which are suggested to affect the binding of DLG1 to adenomatous polyposis coli (APC) and PTEN proteins, are associated with mammary carcinomas (Navarro et al., 2005; Fuja et al., 2004).
The localization and abundance of PDZ proteins is regulated in normal tissues by transcriptional, post-transcriptional and post-translational mechanisms. Several tumours exhibit mislocalisation of PDZ proteins, with gradual loss of the protein at cell–cell contacts, and increased protein in the cytoplasm during early stages of tumorigenesis. Downregulation of polarity proteins is usually associated with a more invasive phenotype, and late stages of carcinogenesis. The light blue area represents the basal membrane. TFs, transcription factors. From Facciuto et al. 2012.
### Table 1.8: PDZ proteins in human disease

<table>
<thead>
<tr>
<th>PDZ Protein</th>
<th>Function</th>
<th>Disease associations</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLG1 (SAP97)</td>
<td>Cell polarity, proliferation, migration and adhesion, cell division, synapse formation</td>
<td>Breast, cervical and colon cancers, schizophrenia, Chron's disease</td>
</tr>
<tr>
<td>DLG2 (PSD93)</td>
<td>Cell polarity, proliferation, synapse formation</td>
<td>Retinoblastoma</td>
</tr>
<tr>
<td>DLG3 (SAP102)</td>
<td>Cell polarity, proliferation, synapse formation</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>DLG4 (PSD95)</td>
<td>Receptor clustering</td>
<td>Schizophrenia, Alzheimer’s, epilepsy</td>
</tr>
<tr>
<td>DLG5</td>
<td>Cell polarity, proliferation, invasion and migration, cell division</td>
<td>Breast and prostate cancer, Crohn's disease, inflammatory bowel disease</td>
</tr>
<tr>
<td>DLG6 (MPP4)</td>
<td>Cell polarity, proliferation, synapse formation</td>
<td>Retinopathies</td>
</tr>
<tr>
<td>PALS1 (MPP5)</td>
<td>Cell polarity, proliferation, synapse formation</td>
<td>Prostate cancer</td>
</tr>
<tr>
<td>SCRIB</td>
<td>Cell polarity and integrity, proliferation, apoptotic signalling</td>
<td>Breast, colon, endometrial and prostate cancers</td>
</tr>
<tr>
<td>PTPN13 (FAP-1)</td>
<td>Fas-associated phosphatase, cell growth signalling, apoptotic signalling</td>
<td>Pancreatic, ovarian and colon cancers, autoimmune thyroid disease</td>
</tr>
<tr>
<td>PARD3 (PAR3)</td>
<td>Cell polarity, asymmetrical cell division</td>
<td>Esophageal and breast cancers, glioblastomas</td>
</tr>
<tr>
<td>PAR6</td>
<td>Cell polarity, asymmetrical cell division</td>
<td>Breast cancer</td>
</tr>
<tr>
<td>CASK</td>
<td>Cell polarity and migration, synapse formation</td>
<td>Renal failure, cerebellar hypoplasia, idiopathic ataxia</td>
</tr>
<tr>
<td>ZO-1</td>
<td>Cell polarity and migration</td>
<td>Breast and liver cancer, kidney disease, cardiopathies</td>
</tr>
<tr>
<td>ZO-2</td>
<td>Cell polarity and migration</td>
<td>Breast cancer, kidney disease, Pancreatic adenocarcinoma</td>
</tr>
<tr>
<td>ZO-3</td>
<td>Cell polarity and migration</td>
<td>Kidney disease</td>
</tr>
<tr>
<td>MAGI-1</td>
<td>Cell polarity, synapse formation, apoptosis</td>
<td>Leukemia, hepatocellular carcinoma</td>
</tr>
<tr>
<td>MAGI-2</td>
<td>Cell polarity, synapse formation, apoptosis</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>MAGI-3</td>
<td>Cell polarity, synapse formation, apoptosis</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>MUPP1 (MPP1)</td>
<td>Cell polarity, proliferation</td>
<td>Retinopathies, inflammatory bowel disease</td>
</tr>
<tr>
<td>NHERF1</td>
<td>Regulation of GPCRs, cytoskeletal anchoring</td>
<td>Hypouricemia</td>
</tr>
</tbody>
</table>

Adapted from Facciuto et al., 2012. Commonly used alternative protein names in brackets.
Similarly, loss of SCRIB has been demonstrated in breast carcinomas, and loss of MAGI-1 is observed in colon and liver carcinomas (Navarro et al., 2005). In Madin-Darby canine kidney (MDCK) cells, expression of MAGI-1 suppresses cell invasion by recruiting PTEN to cell–cell contacts (Kotelevets et al., 2005). In the Par complex, PAR6 has been shown to be overexpressed in breast cancer, and is thought to activate growth signalling cascades and promote cell proliferation independently of growth factor signalling (Aranda et al., 2008; Nolan et al., 2008). Downregulation of PARD3 has been implicated in oesophageal cancer, breast cancers and glioblastomas (Zen et al., 2009; Tokes et al., 2012; Rothenberg et al., 2010), this is suggested to be due to the disassembly of the Par complex that occurs when PARD3 is decreased (Tokes et al., 2012).

Furthermore, colon cancers exhibit a loss of the non-junctional PDZ protein PTPN13, whilst protein levels are increased in aggressive breast cancers (Wang et al., 2004; Glondu-Lassis et al., 2010). This apparent dissimilarity reiterates the fact that the role of PTPN13 (and indeed PDZ proteins) is cell context specific; often it is not clear whether to define PDZ proteins as an oncogene or a tumour suppressor, as their roles depend upon cellular localisation and tissue type (Roberts et al., 2012). A change in abundance and localisation of PDZ proteins disturb the balance of protein function and has severe consequences, namely carcinogenesis.

1.5.5 Virus interactions with PDZ domain containing proteins

A number of human viruses encode proteins containing PDZ-binding motifs (Table 1.9). Given that viral genomes are limited in size, the inclusion of a PBM indicates that the ability to interact with cellular PDZ proteins plays an important role in viral propagation. Computational screening has identified a C-terminal PBM in over 99% of Influenza A virus NS1 proteins (Obenauer et al., 2006). The importance of this PBM to the virulence of
Influenza A was demonstrated by studies showing that deletion or masking of the motif attenuated the virus (Soubies et al., 2010a; Cauthen et al., 2007; Li et al., 2006). Removal of the PBM decreases efficacy of influenza H1N1 virus transmission, as well as virus replication, and increases interferon expression in infected cells (Kim et al., 2014; Soubies et al., 2013; Thomas et al., 2011). Moreover, switching the PBM for the sequence of an avian-specific strain increased virus virulence (Jackson et al., 2008; Zielecki et al., 2010). The NS1 PBM of the human strain H5N1 has been shown to interact with the PDZ domains of MAGI-1, MAGI-2, MAGI-3, DLG1 and SCRIB (Golebiewski et al., 2011; Liu et al., 2010). Despite the homology of the MAGI proteins, the strength of NS1 binding to each is not equal, indicating that viral PBM interactions are highly specific, and are determined by both the PDZ domain and the PBM sequences (Liu et al., 2010; Thomas et al., 2011). The swine-specific strain H6N6 NS1 alters the subcellular localisation of polarity regulator MAGI-1 and H5N1 NS1 colocalises with DLG1 and SCRIB, disrupts cellular tight junctions and produces a similar phenotype to cells in which DLG1 and SCRIB have been knocked down (Golebiewski et al., 2011). It has been suggested that such disruption of cell junctions may benefit viral dissemination within the host, and/or spread to other hosts (Golebiewski et al., 2011; Liu et al., 2010). Furthermore, targeting SCRIB abrogates the pro-apoptotic function of the protein and may therefore protect infected cells from viral induction of apoptotic pathways, thus supporting viral propagation (Liu et al., 2010; Thomas et al., 2011). The PBM has also been suggested to aid viral immune evasion, by downregulating the cellular interferon response (Soubies et al., 2010b; 2013; Kumar et al., 2012).

Propagation of the RNA viruses tick borne encephalitis virus (TBEV) and dengue virus (DV) is dependent upon two virally-encoded PBMs within the NS5 proteins; mutations in the PBM negatively affect viral replication (Melik et al., 2012; Ellencrona et al., 2009). Interestingly, both viruses target the PDZ protein ZO-1, but via different approaches;
where TBEV NS5 is stabilised at the cell membrane by the interaction, DV NS5 colocalises with ZO-1 in the nucleus (Ellencrona et al., 2009). The viruses have also been shown to interact with IL-16, a chemotactic cytokine which contains a PDZ domain (Werme et al., 2008). A number of PDZ proteins have been identified as able to interact with TBEV NS5 protein with varying affinities; the viral factor binds strongly to ZO-2, and less strongly to TIP2, CASK (calcium/calmodulin-dependent serine protein kinase) and GRIP2 (glutamate receptor interacting protein 2, Werme et al., 2008). Furthermore, TBEV NS5 has a high affinity for SCRIB, an interaction which has been suggested to be important in viral protein localisation and in preventing the host anti-viral interferon response, via the JAK-STAT pathway (Werme et al., 2008).

Expression of the core protein of fellow flavivirus hepatitis C virus (HCV) leads to a decrease in DLG1 protein and mislocalisation of SCRIB from cell-cell contacts in liver and kidney epithelial cells (Awad et al., 2013). A redistribution of PDZ proteins E-cadherin, claudin and ZO-1 has also been observed in in HCV infected cells, which leads to disruption of cell polarity and aids the cell-cell transmission of virus (Mee et al., 2010).

Interaction with cellular PDZ-domain containing proteins has been reported to promote the entry of human immunodeficiency virus 1 (HIV-1); knockdown of the cytoskeletal protein PDZD8 inhibits HIV-1 infection and negatively affects viral capsid stability (Henning et al., 2010; Guth and Sodroski, 2014). The PDZ protein Syntenin-1 has been implicated in HIV-1 entry; Gordón-Alonso et al. (2012) suggest that the PDZ protein is recruited to the membrane by HIV-1 protein Env and clusters HIV-1 entry receptors on the cell membrane. Furthermore, overexpression of PDZD8 enhances infection by HIV-1 and related retroviruses murine leukemia virus (MLV), simian immunodeficiency virus (SIV, Henning et al., 2011; 2010). HIV also acts upon the tight junction proteins ZO1,
occludin and claudin, disrupting their localisation and epithelial integrity, as well as downregulating their expression at a transcriptional level (Nazli et al., 2010).

Rhinovirus also appears to affect tight junction proteins at a transcriptional level; ZO-1, claudin, occludin and E-cadherin gene expression is decreased by rhinovirus in cultured epithelia (Yeo and Jang, 2010). This may be the cause of the loss of ZO-1 from tight junctions observed in Rhinovirus infected cells (Sajjan et al., 2008). It is not yet clear which viral factors are responsible for this, but it is suggested that the function promotes viral transmission (Yeo and Jang, 2010). Similarly, the consequences of PDZ binding during the infectious cycle of Hepatitis B virus (HBV) have yet to be fully characterized, although it has been suggested that PTPN3 suppresses the expression of viral genes and HBV core protein has been shown to interact with TIP-2 (Razanskas and SASNAUSKAS, 2010).

SARS coronavirus E protein interacts with Crumbs complex protein PALS1 and Syntenin-1 (Teoh et al., 2010; Jimenez-Guardeño et al., 2014). The former interaction has been shown to effect epithelial polarity and tight junction formation (Teoh et al., 2010). SARS E protein is crucial to viral virulence and the PBM is a determinant of virus pathogenicity; deletion of the PBM had no effect on viral growth in cell culture but attenuated viral infection of mice, where less lung damage and no mortality occurred in mice infected with a PBM-deletion virus compared to wild-type virus (Jimenez-Guardeño et al., 2014; DeDiego et al., 2007; 2008).

Rabies pathogenicity is associated with the ability of the envelope G protein to interact with PDZ proteins PTEN, PTPN4 and MAST2 (Prehaud et al., 2010; Terrien et al., 2012). Both attenuated and virulent strains of Rabies exhibit an affinity for PDZ proteins, but it has been suggested that virulence is not associated solely with the ability to bind PDZ proteins generally, but rather the ability to bind and target proteins specifically (Prehaud et al., 2010).
Intriguingly, the F11 protein of vaccinia virus encodes both a C-terminal PBM and a central PDZ-like domain (Handa et al., 2013). Removal of the PBM results in the production of less infectious virus, and the PDZ domain has been shown to act as a scaffold protein and enhance viral spread (Handa et al., 2013).

**Table 1.9: Virus interactions with PDZ proteins**

<table>
<thead>
<tr>
<th>Virus/viral protein</th>
<th>Interacting PDZ protein</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenovirus E4ORF1</td>
<td>DLG1, MAGI-1, MUPP1, ZO-2</td>
<td>Glaunsinger et al., 2000; Lee et al., 1997</td>
</tr>
<tr>
<td></td>
<td>Tight junction disruption, loss of cell polarity</td>
<td></td>
</tr>
<tr>
<td>Influenza A NS1</td>
<td>Interactions with DLG1, SCIB, MAGI-1, -2, -3 and PDLIM2</td>
<td>Golebiewski et al., 2011; Liu et al., 2010</td>
</tr>
<tr>
<td>Rotavirus NSP4 VP8</td>
<td>Localisation of ZO-1 disrupted</td>
<td>Tafazoli et al., 2001; Nava et al., 2004</td>
</tr>
<tr>
<td>Hepatitis B core protein</td>
<td>TIP-2</td>
<td>Razanskas and Sasnauskas, 2010</td>
</tr>
<tr>
<td>SARS coronavirus envelope protein E</td>
<td>PALS1, Syntenin-1</td>
<td>Jimenez-Guardeño et al., 2014; Teoh et al., 2010</td>
</tr>
<tr>
<td>Rhinovirus</td>
<td>ZO-1 and E-cadherin mRNA and protein downregulation</td>
<td>Yeo and Jang, 2010; Sajjan et al., 2008</td>
</tr>
<tr>
<td>Tick borne encephalitis virus NS5</td>
<td>SCRIB, TIP2, CASK, ZO-1</td>
<td>Melik et al., 2012; Werme et al., 2008; Ellencrona et al., 2009</td>
</tr>
<tr>
<td>HTLV-1 Tax</td>
<td>DLG1 and MAGI-3 mislocalisation, pro-IL16</td>
<td>Suzuki et al., 1999; Rousset et al., 1998; Wilson et al., 2003; Ohashi et al., 2004</td>
</tr>
<tr>
<td>*High-risk α-HPV E6</td>
<td>DLG1, SCIB, MAGI-1, -2, -3, PTPN13, NHERF1, etc.</td>
<td>Kiyono et al., 1997; Kranjec and Banks, 2011; Thomas et al., 2002; Spanos et al., 2008b; Accardi et al., 2011; Nakagawa and Huibregtse, 2000</td>
</tr>
<tr>
<td>rhPV1 E7</td>
<td>PARD3</td>
<td>Tomaic et al., 2009</td>
</tr>
<tr>
<td>CRPV</td>
<td>DLG1</td>
<td>Du et al., 2005</td>
</tr>
<tr>
<td>HIV-1 gp120</td>
<td>Inhibit localisation of ZO-1, occludin and claudin to TJs</td>
<td>Nazli et al., 2010</td>
</tr>
<tr>
<td>Rabies G protein</td>
<td>DLG1, PTPN13, MUPP1</td>
<td>Prehaud et al., 2010</td>
</tr>
</tbody>
</table>

Adapted from Banks et al., 2012; Javier, 2008. *A comprehensive list of PDZ targets of HPV E6 is shown in Table 1.10.

A number of human viruses encode PBMs, which have been implicated in viral entry, propagation and immune evasion. They appear to do so via common PDZ targets proteins,
namely DLG1, SCRIB and MAGI-1, -2 and -3, which is highly suggestive of the important roles that these proteins play in cellular polarity, proliferation, immune evasion and apoptosis, the interruption of which are fundamental in viral replication.

1.5.6 PDZ domain containing proteins in viral oncogenesis

The oncogenic viruses HTLV-1 and adenovirus 9, as well as tumourigenic papillomaviruses rhPV1 (rhesus papillomavirus), CRPV (cottontail rabbit papillomavirus) and HPVs have also been shown to affect PDZ domain containing proteins, and this function has been linked to the oncogenic potential of the virus as well as their infectious cycles (Figure 1.15, Weiss et al., 1997; Rousset et al., 1998).

1.5.6.1 HTLV-1 Tax

Mutational studies have demonstrated the importance of the Tax PBM in viral persistence and to the induction of proliferation in infected cells (Xie et al., 2006). These studies also revealed the PBM is not essential to HTLV-1 mediated immortalisation, although its presence enhances the ability of HTLV-1 to transform rodent cells and enables Tax to transform mouse lymphocytes in cooperation with NFκB (Hirata et al., 2004; Xie et al., 2006). In human cells, the Tax PBM is also required for Tax-induced proliferation of T-cells (Xie et al., 2006; Higuchi et al., 2007). In vitro assays show that the protein kinase CK2 phosphorylates Tax within the PBM, at threonine 351, suggesting that phosphorylation of the PBM can regulate viral protein binding to PDZ proteins (Bidoia et al., 2010).

Yeast two-hybrid screening experiments identified a number of PDZ domain containing Tax-interacting proteins, including DLG4, TIP-1, TIP-2 and PAR6 (Rousset et al., 1998). Additionally, in cell culture, overexpressed Tax was shown to interact with MAGI-3 whereas the PBM mutant protein did not (Ohashi et al., 2004). The interaction results in altered localisation of the viral protein from the nucleus to the cytoplasm and increased
Figure 1.15: Viral proteins target common cellular PDZ proteins

A number of PDZ proteins involved in the formation and maintenance of cell polarity are commonly targeted by viruses. The oncoproteins Tax, E4ORF1 and E6, from oncogenic viruses HTLV-1, adenovirus and HPV respectively, specifically target a number of PDZ substrates.

From Javier and Rice, 2011.

Flu A, influenza A virus; AJ, adherens junction; TJ, tight junction
Tax stability (Ohashi et al., 2004). MAGI-3 is involved in cell polarity and in cell survival signalling; the ability of Tax to target such pathways is suggested to promote viral replication, as well as promoting aberrant growth and transformation of infected cells (Ohashi et al., 2004). Tax was also identified as interacting with PDZ domain containing protein Erbin during screening experiments (Ress and Moelling, 2006). This interaction was confirmed by immunoprecipitation of endogenous Erbin in epithelial cells with Tax protein encoding an intact PBM, but not when the PBM was absent (Song et al., 2009). The first PDZ domain of the T-cell specific interleukin-16 precursor (Pro-IL16) binds to the Tax PBM in both human T-cells and immortalized monkey fibroblasts, where both proteins colocalise in the nucleus (Wilson et al., 2003). The presence of an intact PBM enables Tax-expressing cells to overcome pro IL-16-driven cell cycle arrest, suggesting a role for the interaction in overcoming the growth inhibitory activity of pro IL-16 and promoting cell cycle progression (Wilson et al., 2003). Such progression is favourable for viral propagation and is likely to contribute to the overproliferative phenotype of transformed cells.

The first target of HTLV-1 Tax to be identified was DLG1 (Lee et al., 1997). Expression of wild-type but not PBM-mutant Tax protein hampers DLG1 inhibition of S phase progression in mouse fibroblasts (Suzuki et al., 1999). Furthermore, Tax causes a redistribution in DLG1 protein from the soluble to the insoluble compartment of T-cells in a PBM dependent manner (Hirata et al., 2004). Sequestering of DLG1 has been suggested to be a mechanism of DLG1 inactivation by Tax (Tsubata et al., 2005; Ishioka et al., 2006). However, depletion of DLG1 does not complement transformation in the absence of the Tax PBM, although it does augment transformation by wild-type virus, altogether suggesting that HTLV-1 transformation does not depend on the Tax PBM binding and inactivating DLG1 (Tsubata et al., 2005; Ishioka et al., 2006). Yeast two-hybrid
screening experiments revealed that HTLV-1 Tax binds to fellow Scribble complex member SCRIB (Arpin-Andre and Mesnard, 2007). This interaction was confirmed \textit{in vivo} by the detection of a Tax and SCRIB complex in infected human T-cells, where Tax caused SCRIB mislocalisation (Aoyagi \textit{et al.}, 2010). Both DLG1 and SCRIB block T-cell receptor (TCR)-induced activation \textit{via} NFAT a key regulator of T-cell activation and differentiation (Round \textit{et al.}, 2007b; Xavier \textit{et al.}, 2004). In wild-type Tax expressing T-cells, this activity of SCRIB was inhibited, whereas inhibition still occurred when the Tax PBM was mutated (Arpin-Andre and Mesnard, 2007). DLG1 is important in the formation and organisation of the immunological synapse (Cullinan \textit{et al.}, 2002). HTLV-1 infected T-cells fail to establish proper apical-basal polarity (Barnard \textit{et al.}, 2005). Together, this suggests that the interaction between Tax and DLG1 or SCRIB disrupts T-cell polarity, migration, and signalling as well as immune synapse formation, resulting in the stimulation of cell proliferation and prohibition of immune activation by the virus, thus aiding HTLV-1 persistence (Ludford-Menting \textit{et al.}, 2005; Rebeaud \textit{et al.}, 2007; Rincon and Davis, 2007; Round \textit{et al.}, 2005; 2007b).

Interestingly, the association between the Tax oncoprotein and the PDZ domain containing ubiquitin E3 ligase PDLIM2 does not occur between the viral PBM and the PDLMI2 PDZ domain (Fu \textit{et al.}, 2010). Instead, Tax binds to PDLIM2 \textit{via} multiple regions in Tax and an \(\alpha\)-helical motif in PDLIM2, unlike the multiple PDZ targets which have been identified to interact \textit{via} the PBM of the viral protein (Fu \textit{et al.}, 2010). The interaction causes transport of the viral protein to the nucleus and subsequent proteasome-mediated degradation, thereby suppressing Tax’s transforming activity (Yan \textit{et al.}, 2009b).
1.5.6.2 Adenovirus E4ORF

Adenovirus type 9 causes benign eye infections in humans, but is capable of causing mammary tumours in rats in the presence of high levels of oestrogen and of transforming cells in culture (Graham et al., 1984; Horwitz, 2001). Adenovirus 9-induced transformation is driven by the viral factor E4ORF1 and, more specifically, the PBM at the C-terminus of E4ORF1 has been shown to be crucial in this oncogenesis (Weiss and Javier, 1997). Furthermore, expression of the E4ORF1 causes tight junction defects in epithelial cells in a PBM-dependent manner, but does not affect the structure of adherens junctions (Latorre et al., 2005).

DLG1 was the first identified target of E4ORF1, shown to interact via PDZ domains 1 and 2 (Lee et al., 1997). Subsequently, it was demonstrated that E4ORF1 was able to bind to the cellular PDZ proteins MUPP1, PATJ, MAGI-1, and ZO-2 (Glaunsinger et al., 2000; Lee et al., 2000; Latorre et al., 2005; Glaunsinger et al., 2001). Interestingly, there are differences in PDZ protein binding between the monomeric and trimeric forms of the E4ORF1 protein; where monomeric E4ORF1 interacts with MAGI-1, PATJ, MUPP1 and ZO-2, the trimer only binds to DLG1 (Chung et al., 2008). E4ORF1 expression causes the sequestering of MUPP1, PATJ, MAGI-1, and ZO-2 away from tight junctions into cytoplasmic membrane vesicles, disrupting cell adhesion and polarity (Glaunsinger et al., 2000; Lee et al., 2000; Glaunsinger et al., 2001; Latorre et al., 2005). E4ORF1 trimers promote DLG1 translocation to the plasma membrane, which may contribute to the junctional defects observed in adenovirus infected cells (Laprise et al., 2004; Stucke et al., 2007). However, E4ORF1 monomers exert a gain-of-function upon DLG1, as the interaction constitutively activates the PI3K signalling pathway via the PDZ protein (Frese et al., 2003). The association of the E4ORF1 with DLG1 facilitates adenovirus activation of the phosphatidylinositol 3-kinase (PI3K) signalling cascade via oncogene RAS, which
culminates in an increase in cellular metabolism and thus increases cellular substrates available for viral replication (Kong et al., 2014; Frese et al., 2006; Kumar et al., 2014). This interaction also induces anchorage independent growth (AIG), a characteristic of transformed cells (Frese et al., 2003; 2006).

It has been suggested that both Tax and E4ORF interfere with the interaction between DLG1 and the cellular tumor suppressor adenomatous polyposis coli protein (APC) by competing for binding to the same PDZ domain (Javier and Rice, 2011; Frese et al., 2006). The APC:DLG1 interaction has growth-inhibitory consequences; viral protein binding DLG1 would prevent APC binding and thus lead to an increase in cell proliferation, due to increased growth promoting signalling (Ishidate et al., 2000; Matsumine et al., 1996; Javier, 2008).

1.5.6.3 Papillomavirus oncoproteins

Both animal and human PVs are able to target PDZ proteins. The E6 protein of CRPV has also been shown to interact with cellular DLG1, underlining a commonality in viral PDZ protein targeting (Du et al., 2005). Although expression of CRPV E6 is able to immortalise rabbit keratinocytes, the protein does not bind cellular p53 (Muench et al., 2010). Furthermore, CRPV is able to induce skin cancer in rabbits without additional co-factors (Ganzenmueller et al., 2008; Jeckel et al., 2002; Muench et al., 2010). Together this suggests that DLG1 binding may contribute to CRPV E6 mediated immortalization, in the absence of this tumorigenic interaction.

Interestingly, the oncogenic rhesus papillomavirus type 1 (RhPV1) E7 oncoprotein, but not the E6 protein, contains a C-terminal PBM, via which it interacts with the polarity protein PARD3 (Tomaic et al., 2009). RhPV1 drives mucosal neoplasia in rhesus macaques, in an
analogous fashion to high-risk HPVs, again indicating a potential carcinogenic role for the E6 PBM.

The β-papillomaviruses do not encode an E6 PBM, but do encode a conserved C-terminal hydrophobic motif (YXDW). In HPV5, this motif interacts with the adhesion protein β1-integrin and disrupts normal signalling (Du et al., 2005). Furthermore, the β-papillomavirus HPV8 encodes an alternative mechanism to regulate the PDZ protein Syntenin-2 by influencing transcription (Lazic et al., 2012). Both HPV5 and HPV8 are associated with the development of skin cancer, and a level of control over adhesion proteins may be important to the cancer-causing capability of viruses as well as functioning during normal papillomavirus life cycle.

The commonality in viral PDZ targets further highlights oncogenic potential of the proteins as well as their importance in cellular proliferation (Figure 1.15). Viral PBMs function to promote a favourable cellular environment for viral replication, but are also important in driving oncogenesis. By activating cellular proliferation and anti-apoptotic pathways, PBMs contribute to transformation of infected cells which, when combined with additional mutations and carcinogenic environmental influences, can lead to cancerous growth.

1.6 High-risk HPV E6 and cellular PDZ proteins

Like Tax and E4ORF1, the E6 oncoproteins of high-risk HPVs encode a C-terminal PBM (Figure 1.16). To date, more than 15 PDZ domain containing proteins have been shown to interact with high-risk HPV E6 oncoproteins. The proteins identified encompass a variety of roles in cell polarity, tissue integrity, scaffold formation, signalling and intracellular trafficking (Table 1.10). The biological consequences of HPV E6 PBM binding PDZ
proteins are varied, and have been investigated in cell-based models of papillomavirus oncogenesis and life cycle.

1.6.1 The E6 PBM in viral life cycle

The E6 protein functions to decouple keratinocyte differentiation from proliferation (Section 1.4.1.2). The interaction of the E6 PBM with a subset of cellular PDZ proteins is influential in the life cycle of high-risk HPVs, where E6 is expressed at physiological levels. The motif may contribute to the overall function of E6 by promoting disruption of cell polarity and enabling proliferation in cells that would otherwise differentiate.

Studies in mice have revealed that the E6 PBM promotes DNA replication; in mouse epithelia the interaction between E6 and PDZ proteins correlates with the induction of suprabasal DNA synthesis by the oncoprotein (Nguyen et al., 2003). Deletion of the PBM in the context of whole HPV genomes has shown that the motif contributes to HPV-driven keratinocyte proliferation; where stratified cells containing HPV18 or HPV31 genomes exhibited a thicker spinous layer, PBM mutant genome containing cells exhibited a morphology that resembled normal keratinocytes (Lee and Laimins, 2004; Delury et al., 2013). In addition, the HPV31 E6 PBM is necessary for the induction of keratinocyte proliferation in both undifferentiated and stratified cell culture and keratinocyte growth is lessened in cells containing a HPV18 E6 PBM mutant compared to cells containing the wild-type genome or to normal keratinocytes (Lee and Laimins, 2004; Delury et al., 2013).

Similarly, in both undifferentiated and stratified primary keratinocytes, Choi et al. (2013) demonstrated that the PBM of HPV16 E6 is required for the increased growth seen in cells expressing E6 as well as in cells containing both E6 and E7 oncoproteins.

Deletion within the context of the whole HPV genome affects the viral genome; the loss of the E6 PBM of HPV16, HPV18 and HPV31 reduces the capacity of the virus to maintain episomes (Nicolaides et al., 2011; Delury et al., 2013; Lee and Laimins, 2004; Brimer and
Vande Pol, 2014). Indeed, the PBM is essential to viral propagation; its absence results in a lack of genome amplification and late protein expression (Delury et al., 2013).

It has been suggested that removal of the PBM destabilizes the E6 protein; Nicolaides et al., (2011) demonstrated an enhanced proteasomal degradation rate of HPV16 E6 in the absence of a PBM, as well as an increase in the levels of E6 protein when the PDZ target SCRIB was overexpressed. However, deletion of the HPV18 and HPV31 E6 PBM does not affect E6 stability in primary keratinocytes (Lee and Laimins, 2004; Delury et al., 2013). Furthermore, phosphorylation of the E6 PBM has been implicated in E6 stability, suggesting a more complex interplay (Boon and Banks, 2012b).

The PBM has been implicated in protecting cells from TNF-induced apoptosis; HPV16 E6 activates nuclear factor kappa B (NFκB) in human airway epithelial cells and subsequently inhibits downstream apoptotic signalling in a PBM dependent manner (James et al., 2006). A number of known PDZ targets of E6 have roles in apoptotic signalling, including PTPN13 and SCRIB, and it may be that these interactions have anti-apoptotic consequences, but it is likely that such activation also promotes epithelial proliferation and thus viral replication (Nakagawa and Huibregtse, 2000; Spanos et al., 2008b; James et al., 2006).

Thus there is evidence that the PBM function contributes to E6-mediated promotion of viral propagation. Such functions are also influential in the transformation of infected cells, under conditions where E6 is deregulated.

1.6.2 The E6 PBM in viral oncogenesis

The E6 PBM is found exclusively in the high-risk type HPV s (Figure 1.16), and in culture this PBM is required for E6-mediated transformation of rodent cells (Kiyono et al., 1997; Spanos et al., 2008a). In primary human tonsil keratinocytes (HTKs) containing HPV16,
Figure 1.16: C-terminal alignment of high- and low-risk HPV E6 proteins

E6 protein To date, twelve HPV types have been defined by the WHO as high-risk with respect to causing cancer - these are designated as IARC Group 1- and a further ten types are recognised as ‘possibly’ (Group 2A) and ‘probably’ cancer-causing (Group 2B). HPV6 and 11 are included as examples of low risk virus. HPV8 is associated cutaneous SCC, and has been shown to effect the transcription of PDZ protein Syntenin-2. The canonical PBM sequence (X₅ᵢ,X₅ᵢ₊₁) is shown above the highlighted region, which indicates the E6 PBM sequence. Phospho acceptor sites within the PBM are highlighted in bold. Adapted from Delury et al. 2013.
PBM deletion depleted the efficiency of HPV-driven immortalisation of cells, and removed the ability of HPV16 E6 to induce AIG (Spanos et al., 2008a). Similarly, the expression of HPV18 E6 in human keratinocytes led to disrupted cell-cell junctions and promotion of a fibroblast-like morphology, phenotypes which are characteristics of EMT and these were reduced when the PBM was deleted (Watson et al., 2003).

As in cell culture, E6 is able to induce epithelial hyperplasia in transgenic mice independently of interactions with p53 (Song et al., 2000; Kiyono et al., 1997). In concordance with this, mouse models expressing HPV16 E6 under a keratin promoter have demonstrated that the PBM is key to the development of hyperplasia of the skin, the cervix and the eye, although the E6 PBM is not required for the immortalisation of other keratinocytes (Nguyen et al., 2003; 2005; Shai et al., 2010). Furthermore, murine studies have shown that the E6 PBM promoted greater invasive properties and tumour size, and that the cooperation of E6 and E7 in driving tumour growth was shown to be dependent upon the E6 PDZ binding function (Nguyen et al., 2003; Simonson et al., 2005).

A number of PDZ proteins have been identified to interact with the high-risk HPV E6 PBM (Table 1.10), and have been investigated in systems that resemble a cancer-like cellular environment, highlighting interactions that may be crucial in promoting the HPV life cycle.

1.6.3 PDZ targets of high-risk E6

As with other oncogenic viruses, the PDZ proteins DLG1 and SCRIB are common targets of both high-risk HPV types 16 and 18 (Kiyono et al., 1997; Nakagawa and Huibregtse, 2000). Furthermore, the different HPV types display differing affinities for targets; despite the PBM differing by a single amino acid, HPV16 preferentially binds SCRIB over DLG1, and HPV18 DLG1 over SCRIB, these affinities can be altered within the context of whole E6 by switching the PBM sequences (Pim et al., 2000; Thomas et al., 2001). In addition,
DLG4 has been shown to be a stronger target for HPV18 than for HPV16, whilst GOPC was found to bind more strongly to HPV16 E6 than HPV18 (Handa et al., 2007; Kranjec and Banks, 2011; Jeong et al., 2007).

The binding of PDZ proteins themselves are formed in a select fashion; despite the presence of more than one PDZ domain per cellular protein, only one domain is thought to interact with E6. Specificity is determined by the sequence of the viral PBM, sequence upstream of the PBM and the target PDZ domain (Thomas et al., 2001; 2002). The first two PDZ domains of DLG1 (PDZ1 and PDZ2) are important for HPV18 E6 binding (ETQV) and the threonine residue of the PBM is essential for the interaction; substitutions with either neutral or acidic residues abolishes binding and subsequent ubiquitin ligase recruitment (Gardiol et al., 1999). Structural analysis confirmed this, and indicated that the arginine residues proximal to the PBM form additional contacts with PDZ2 to strengthen the interaction (Liu et al., 2007). Similarly, the structure of MAGI-1 PDZ1 bound to HPV16 E6 confirmed that residues located inside and outside the binding groove contribute to binding, notably lysine 499 of the junctional protein is important in the specificity of the E6 interaction (Fournane et al., 2011). E6 PDZ targeting is also specific to cellular location; in the case of DLG1 and MAGI-1, HPV18 E6 preferentially induces the degradation of the cytoplasmic and nuclear forms (Massimi et al., 2004; Kranjec and Banks, 2011; Narayan et al., 2009).

Although hundreds of PDZ domains exist in a human keratinocyte, HPV E6 has been shown to associate with seventeen (Table 1.10). Recent screening studies have identified an additional number of PDZ targets which have yet to be confirmed in vivo (Belotti et al., 2013). The identified targets of the E6 PBM comprise a range of cellular proteins involved in epithelial cell polarity, tissue integrity and proliferation (Table 1.10). PDZ targets of E6 have been identified from the Scribble, Par and Crumbs polarity complexes (Figure 1.15),
as have a number of cellular PDZ proteins involved in cell signalling. Specific E6 PDZ targets are discussed in more detail in Chapter 3.

1.6.3.1 Kinase recognition of HPV E6 PBM

The PBM of high-risk HPV types lies in close proximity to a canonical serine/threonine-specific kinase recognition site (Figure 1.16). X-ray crystal structures predict that phosphorylation of the PBM would sterically hinder the ability of E6 to fit within the PDZ binding groove (Ainsworth et al., 2008; Charbonnier et al., 2011; Liu et al., 2007). In vitro experiments have shown that HPV18 E6 PBM phosphorylation reduces DLG1 binding and activation of the PKA pathway in cell culture decreases degradation of DLG1 by E6, which suggests that phosphorylation of the PBM negatively affects E6 PDZ binding ability (Kühne et al., 2000). Furthermore, in vitro protein kinase A (PKA) phosphorylates the E6 PBM of high-risk HPV types 16 and 18, and a site proximal to the PBM of HPV31 (Boon et al., 2014). Moreover, different E6 proteins exhibit different susceptibilities to phosphorylation. HPV18 E6 shows a higher susceptibility to PKA phosphorylation than HPV16, and HPV31 is only weakly phosphorylated by PKA (Kühne et al., 2000; Boon et al., 2014). The kinase AKT (also known as protein kinase B, PKB) is a serine/threonine kinase family member which has been also shown to act upon both HPV16, 58 and 31 E6 PBMs in vitro (Boon et al., 2014). It is therefore likely that other kinases recognize these sites, and perhaps recognize certain E6 PBMs better than others.

Investigations by the Roberts lab have indicated that E6 PBM phosphorylation may function to regulate interactions with PDZ proteins during the virus life cycle (Delury et al., 2013). By reversibly interrupting PDZ binding, it is hypothesized that phosphorylation may provide a mechanism of regulating PDZ-PBM interactions during the viral life cycle and during epithelial differentiation. In addition to interrupting PDZ interactions,
phosphorylation has been shown to introduce a novel interaction; phosphorylation of the HPV18 E6 PBM facilitates binding of the phospho-tyrosine binding scaffold protein 14-3-3ζ (Boon and Banks, 2012b). This indicates that E6 PBM function depends on the phosphorylation status, mediating interactions with either PDZ domain containing substrates or 14-3-3 family members.
Table 1.10: PDZ targets of high-risk HPV E6 proteins

<table>
<thead>
<tr>
<th>PDZ protein (Gene Symbol)</th>
<th>Alternative Name(s)</th>
<th>Cellular Functions</th>
<th>Effect of Interaction with E6</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLG1 SAP97</td>
<td>Cell polarity, cell proliferation, migration and adhesion</td>
<td>Degradation, mislocalisation</td>
<td>Kiyono et al., 1997; Gardiol et al., 1999</td>
<td></td>
</tr>
<tr>
<td>SCRIB Scribble, Vartul, LAP4</td>
<td>Cell polarity and integrity, proliferation, apoptotic signalling</td>
<td>Degradation</td>
<td>Nakagawa and Huibregtse, 2000</td>
<td></td>
</tr>
<tr>
<td>INADL PATJ</td>
<td>Tight junction formation, cell polarity</td>
<td>Degradation</td>
<td>Latorre et al., 2005; Storrs and Silverstein, 2007</td>
<td></td>
</tr>
<tr>
<td>SLC9A3R1 NHERF1, EBP50</td>
<td>GPCR regulation, cytoskeleton anchoring</td>
<td>Degradation</td>
<td>Accardi et al., 2011</td>
<td></td>
</tr>
<tr>
<td>PTPN13 FAP-1, PTPL-1, PTPL-BAS, PTPLBL</td>
<td>Fas-associated phosphatase, cell growth signalling, apoptotic signalling</td>
<td>Degradation</td>
<td>Spanos et al., 2008b</td>
<td></td>
</tr>
<tr>
<td>MPDZ MUPP1</td>
<td>Receptor clustering, scaffold protein</td>
<td>Degradation</td>
<td>Lee et al., 2000</td>
<td></td>
</tr>
<tr>
<td>MAGI-1 AIP3, WWP3</td>
<td>Cell polarity, invasiveness, synapse formation, apoptosis</td>
<td>Degradation</td>
<td>Glaunsinger et al., 2000</td>
<td></td>
</tr>
<tr>
<td>MAGI-2 AIP1</td>
<td>Cell polarity, invasiveness, synapse formation, apoptosis</td>
<td>Degradation</td>
<td>Thomas et al., 2002</td>
<td></td>
</tr>
<tr>
<td>MAGI-3 -</td>
<td>Cell polarity, invasiveness, synapse formation, apoptosis</td>
<td>Degradation</td>
<td>Thomas et al., 2002</td>
<td></td>
</tr>
<tr>
<td>DLG4 PSD95, SAP90, TIP15</td>
<td>Receptor clustering</td>
<td>Degradation</td>
<td>Handa et al., 2007</td>
<td></td>
</tr>
<tr>
<td>PTPN3 PTPH1</td>
<td>Phosphatase, cell growth signalling, apoptotic signalling</td>
<td>Degradation</td>
<td>Jing et al., 2007; Töpffer et al., 2007</td>
<td></td>
</tr>
<tr>
<td>TAX1BP3 TIP-1</td>
<td>Wnt pathway inhibitor</td>
<td>-</td>
<td>Hampson et al., 2004</td>
<td></td>
</tr>
<tr>
<td>GIPC1 Synectin, TIP-2</td>
<td>Cell surface receptor expression and trafficking, TGFβ signalling</td>
<td>Degradation</td>
<td>Favre-Bonvin et al., 2005</td>
<td></td>
</tr>
<tr>
<td>GOPC CAL</td>
<td>Intracellular trafficking</td>
<td>Degradation</td>
<td>Jeong et al., 2007</td>
<td></td>
</tr>
<tr>
<td>PARD3 PAR3</td>
<td>Cell polarity, asymmetrical cell division</td>
<td>Mislocalisation</td>
<td>Faccio et al., 2014</td>
<td></td>
</tr>
<tr>
<td>SDCBP2 Syntenin-2</td>
<td>Cell adhesion, growth factor signalling</td>
<td>Decreased transcription</td>
<td>Lazic et al., 2012</td>
<td></td>
</tr>
<tr>
<td>PDZRN3 LNX3</td>
<td>E3 ubiquitin ligase</td>
<td>Degradation</td>
<td>Thomas and Banks, 2014</td>
<td></td>
</tr>
</tbody>
</table>

E6* is a truncated isoform of E6. Underlined symbols indicate nomenclature used in this thesis.
1.7 Hypothesis and Aims

The E6 PBM is a conserved feature of high-risk α-HPV types and a number of cellular PDZ proteins have been identified as interacting with this motif. In many cases, the E6:PDZ interaction leads to proteasome-mediated degradation of the cellular protein. However, the effect of HPV genome replication upon PDZ protein expression has not been fully characterized. Moreover, the physiologically relevant targets and the consequences of the E6:PDZ interaction remains unclear. Furthermore, the PBM overlaps with a kinase recognition site, which may function to regulate PDZ interactions during the viral life cycle. Given the large number of cellular PDZ domain containing proteins and the specificity of PDZ:PBM binding, it is likely that E6 targets specific proteins at certain subcellular locations and at specific times during the HPV life cycle. Thus, this thesis aims to investigate PDZ protein expression in a primary keratinocyte model of the HPV life cycle.

1. To investigate PDZ domain containing protein expression during the life cycle of high-risk HPV types 16 and 18 (Chapter 3).

2. To characterize the expression of E6 PDZ target DLG1 in epithelia and during the HPV16 and 18 life cycles (Chapter 4).

3. To determine the effect of PKA phosphorylation upon the HPV18 E6 PDZ binding domain motif (Chapter 5).
CHAPTER 2: MATERIALS AND METHODS

Unless otherwise stated, materials and chemicals were obtained from commercial sources (Section 2.5). For cell culture reagents and stock solution recipes, see Appendix III.

2.1 Cell culture

Media and solutions were prepared as described below, or purchased from indicated suppliers. All culture procedures were performed under sterile conditions in a Holton Laminair S2000 1.2 microbiological safety cabinet (Thermo Fisher Scientific, UK). Culture media components were filter sterilised using 0.22 µm hydrophilic polyvinylidene fluoride (PVDF) membrane filters. Cells were cultured in sterile dishes (BD Biosciences, UK) in the appropriate growth medium, and were incubated in a cell culture incubator at 37°C and with 5% CO₂ in air.

2.1.1 Culture of cell lines

The cancer-derived cell lines HeLa, Caski, HEK293T, H1299 and SiHa were all cultured in Dulbecco’s modified eagle medium (DMEM, Sigma-Aldrich, UK) supplemented with 4 mM L-glutamine (Gibco, UK) and either 10% v/v foetal bovine serum (FBS, Gibco) or foetal calf serum (FCS, Gibco, see Table 2.1). The HPV-negative cervical cancer cell line C33A was cultured in RPMI-1640 (Roswell Park Memorium Institute) medium (Sigma-Aldrich) supplemented with 10% v/v FCS and 4 mM L-glutamine (Gibco). The head and neck cancer derived cell lines Vu40T, Vu147, SCC040 and SCC154 were cultured in DMEM supplemented with 10% v/v FBS, 1% v/v penicillin streptomycin (PenStrep, Sigma-Aldrich) and 4 mM L-glutamine. Primary cervical epithelial cells, Cx129 were cultured in serum-free media (SFM, Sigma-Aldrich). For subculture, cells were first washed with 5 mL sterile phosphate-buffered saline (PBS, Sigma-Aldrich) and then incubated with
1 mL TrypLE (Life Technologies, UK) until cells were no longer adherent. The TrypLE was quenched with media and the entire suspension transferred to a sterile tube for centrifugation at 10000 × g for 5 min. The cell pellet was resuspended in the appropriate culture media, cells were counted and cells were seeded at the desired density into 10 cm dishes with 10 ml of warmed medium.

Table 2.1: Cell lines

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Tissue of origin</th>
<th>Media requirements</th>
<th>HPV status</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa</td>
<td>Cervical adenocarcinoma</td>
<td>DMEM, 10% v/v FBS, 4mM L-glutamine</td>
<td>HPV18 integrated</td>
</tr>
<tr>
<td>Caski</td>
<td>Cervical carcinoma</td>
<td>DMEM, 10% v/v FCS, 4mM L-glutamine</td>
<td>HPV16 positive</td>
</tr>
<tr>
<td>SiHa</td>
<td>Cervical carcinoma</td>
<td>DMEM, 10% v/v FBS, 4mM L-glutamine</td>
<td>HPV16 positive</td>
</tr>
<tr>
<td>SCC154</td>
<td>Oropharyngeal SCC</td>
<td>DMEM, 10% v/v FBS, 4mM L-glutamine, 1% v/v PenStrep</td>
<td>Positive</td>
</tr>
<tr>
<td>Vu147</td>
<td>HNSCC</td>
<td>DMEM, 10% v/v FBS, 4mM L-glutamine, 1% v/v PenStrep</td>
<td>Positive</td>
</tr>
<tr>
<td>H1299</td>
<td>Lung carcinoma</td>
<td>DMEM, 10% v/v FCS, 4mM L-glutamine</td>
<td>Negative</td>
</tr>
<tr>
<td>Cx129</td>
<td>Primary cervical epithelia</td>
<td>SFM</td>
<td>Negative</td>
</tr>
<tr>
<td>C33A</td>
<td>HPV negative cervical epithelia</td>
<td>RPMI-1640, 10% v/v FCS, 4mM L-glutamine</td>
<td>Negative</td>
</tr>
<tr>
<td>HEK293T</td>
<td>Adenovirus transformed embryonic kidney</td>
<td>DMEM, 10% v/v FCS, 4mM L-glutamine</td>
<td>Negative</td>
</tr>
<tr>
<td>SCC040</td>
<td>Oropharyngeal SCC</td>
<td>DMEM, 10% v/v FBS, 4mM L-glutamine, 1% v/v PenStrep</td>
<td>Negative</td>
</tr>
<tr>
<td>Vu40T</td>
<td>HNSCC</td>
<td>DMEM, 10% v/v FBS, 4mM L-glutamine, 1% v/v PenStrep</td>
<td>Negative</td>
</tr>
</tbody>
</table>

For storage of stocks in liquid nitrogen, two million cells were counted, pelleted and resuspended in cryopreservation media consisting of media (DMEM or SFM as
appropriate), supplemented with 20% (v/v) FCS or FBS and 10% (v/v) dimethyl sulfoxide (DMSO, Sigma-Aldrich).

2.1.2 Cell counting
During cell harvest, disposable Glasstic™ slides (Hycor Biomedical Ltd., UK) were used to determine cell concentration in culture. Ten microlitres of cell suspension was pipetted into a well on the Glasstic™ slide. Using an inverted light microscope, all live (bright spherical) cells contained within three large squares of the Glasstic™ grid were counted. This number was divided by three to obtain an average count of cell number, which was multiplied by $10^4$ to calculate the number of cells per ml of culture medium. Typically, cells were then made up to a density of $2 \times 10^6$ cells per ml.

2.1.3 Mycoplasma testing
All cell lines were periodically tested using the MycoAlert mycoplasma detection kit (Lonza, UK), as per the manufacturer’s instructions and were found to be consistently negative for mycoplasma.

2.1.4 Generation and maintenance of HPV genome-containing cell lines

2.1.4.1 Maintenance of primary foreskin keratinocytes prior to transfection
Normal primary human foreskin keratinocytes (HFKs) were isolated from neonate foreskin by Dr Sally Roberts following the protocol described by Meyers and Laimins (1994). Foreskin tissues were obtained from circumcisions (carried out by Dr Joseph Spitzer) with informed parental written consent (Regional Ethics Committee approval number 06/Q1702/45, Roberts). Also, HFKs were purchased from Lonza (Clonetics). Throughout this thesis, the donor number assigned to cells derived from the same background is consistent (Donor 1-4). HFKs were cultured on collagen-coated tissue culture dishes in
serum-free keratinocyte growth medium (Invitrogen, UK) and media was replaced every two days. Cells were grown to 80% confluency, at which point they were passaged by removal of media followed by a wash with sterile PBS and the addition of 1 ml of 0.05% trypsin/EDTA (Gibco). This was incubated for 5 minutes at 37°C, in a similar procedure as described for cell lines. Trypsin was inactivated by the addition of 1 ml trypsin inhibitor (TI; 0.25 mg/ml soyabean trypsin inhibitor in Dulbecco’s PBS without calcium or magnesium, pH 7.2, Invitrogen). Cells were then transferred to a 15 ml conical tube and pelleted by centrifugation at 10000 × g for 5 minutes at room temperature. The supernatant was subsequently removed and the pellet resuspended in SFM. Typically, a 10 cm tissue culture dish was seeded with 2×10^5 cells in 10 ml SFM.

2.1.4.2 Cryopreservation of HFKs

Cells were frozen at a density of between 1 and 2×10^6 cells per vial. Freezing media consists of SFM supplemented with 10% (v/v) FCS and 10% (v/v) DMSO. HPV genome-containing HFKs were frozen in E-medium supplemented with 10% (v/v) Hyclone™ FCS (Thermo Fisher Scientific) and 20% (v/v) glycerol (Thermo Fisher Scientific).

2.1.4.3 Generation of HPV genome-containing HFKs

Generation of HFK cell lines was undertaken according to the methods described by Wilson and Laimins, (2005, see Figure 2.1). In preparation, HFKs were seeded into uncoated 5 cm tissue culture dishes and allowed to grow. Once the cells reached 50-60% confluency, recircularised HPV genomes (Section 2.2.3.6) were co-transfected into HFKs alongside pCDNA3.1 (Invitrogen), a plasmid that carries the neomycin resistance gene. Control transfections carried out in parallel include; the pCDNA3.1 (Neo) alone, E green fluorescent protein (EGFPe1; Clontech, France) reporter expression vector, and a no plasmid vector DNA. For each transfection, 94 µl of SFM was dispensed into a
Figure 2.1: Transfection of primary keratinocytes with HPV genomes

Primary foreskin keratinocytes were grown to 50-60% confluency in 5 cm tissue culture dishes. For transfection with recircularised HPV genomes, media was replaced with serum-free media. Genomes were co-transfected into cells alongside a neomycin resistance plasmid using Fugene6 transfection reagent. After incubation at 37°C, 5% CO₂ for 48 h, transfected keratinocytes were re-plated onto 10 cm tissue culture dishes previously seeded with 2×10⁶ γ-irradiated J2-3T3 fibroblasts and cultured in complete E medium. A G418 selection process was applied over the next 8 days; addition of selection-containing media and 1×10⁶ irradiated J2-3T3 cells was alternated by day. This selection period began two replenishments with CEM containing 100 μg/ml G418 followed by two replenishments of 200 μg/ml G418. Following the last day of G418 treatment, media was replaced with normal CEM and 2×10⁶ irradiated J2-3T3 cells were added. Once colonies reached approximately 2 cm in diameter, typically 1 month after completion of the selection process, the cells were pooled and transferred onto a single dish pre-seeded with γ-irradiated J2-3T3 cells. Upon reaching 80% confluence, HFKs were passaged onto five 10 cm tissue culture dishes.
polypropylene conical tube (BD Biosciences), followed by the addition of 6 µl FuGene 6 Transfection Reagent (Roche, UK). In a separate tube, 100 µl of SFM, 1 µg of HPV genomic plasmid DNA and 1 µg of neomycin resistance plasmid were mixed. The contents of two tubes were combined and then incubated at room temperature for 30 min. Media from the HFKs was replaced with SFM media (Gibco), and the transfection mixture was added to the dish dropwise, rocked gently to mix and incubated at 37°C and 5% CO₂ for 24 hrs.

Transfection efficiency was measured by observation of the GFP control transfection, using a Zeiss Axiovert 100 microscope, with GFP-expressing cells typically representing 5% of the total population. Transfected keratinocytes were then removed from the dish by trypsin treatment and re-plated onto 10 cm tissue culture dishes previously seeded with 2×10⁶ γ-irradiated J2-3T3 fibroblasts in complete E medium (CEM, Section 2.1.4.5). Over the next 8 days the HFKs underwent a selection process, with the alternate addition of selection-containing media and 1×10⁶ irradiated J2-3T3 cells. The selection period began with the replacement of CEM with CEM containing 100 µg/ml G418 (PAA Laboratories, UK). After two replenishments with 100 µg/ml G418 CEM (typically 4 days), selection was increased to 200 µg/ml G418 for two replenishments (typically 4 days). Following the last day of G418 treatment, media was replaced with normal CEM and 2×10⁶ irradiated J2-3T3 cells were added. In the event that cells reached 80% confluence during the selection process, the dish was split equally onto two dishes previously seeded with 2×10⁶ irradiated J2-3T3 fibroblasts in the presence of normal CEM and the selection process was resumed the following day.

Once colonies reached approximately 2 cm in diameter, typically observed 1 month after completion of the selection process, the dishes were pooled and transferred onto a single dish pre-seeded with γ-irradiated J2-3T3 cells. Upon reaching 80% confluence, HFKs were
passaged onto five 10 cm tissue culture dishes. Once 70-80% confluence was reached, cell pellets were harvested by trypsin treatment, washed with ice-cold saline and prepared for RNA, DNA and protein analysis, as well as samples cryopreserved in liquid nitrogen.

2.1.4.4 Maintenance of HFKs following transfection of HPV genomes

Genome-containing HFKs were cultured on a feeder layer of $2 \times 10^6$ irradiated J2-3T3 fibroblast feeder cells in CEM. CEM consists of Dulbecco’s modified eagle media (DMEM, 60% v/v, Gibco), Ham’s F12 (32% v/v, Gibco), the antibiotics penicillin and streptomycin (2% v/v PenStrep, Gibco), hydrocortisone (0.1% v/v, Sigma-Aldrich), Hyclone™ FCS, (10% v/v, Thermo Fisher Scientific), mouse epidermal growth factor (EGF, 0.5% v/v, BD Biosciences), L-glutamine (2% v/v, Gibco), cholera toxin A (0.1% v/v, MP Biomedicals, USA), insulin (0.2% v/v, Sigma-Aldrich), transferrin (0.2% v/v, Sigma-Aldrich), tri-iodo-1L-thyronine (T3, 4 nM, Sigma-Aldrich) and adenine (36 µM, Sigma-Aldrich). The media was replaced every two days and the cells grown to a confluency of no greater than 80% prior to passaging. Irradiated J2-3T3 cells were seeded onto 10 cm dish containing 10 ml E-medium at least 2 h, or 24 h, prior to addition of HFKs. HFKs were passaged by first removing the J2-3T3 feeder layer, using 0.5 mM EDTA (ethylene diamine tetraacetic acid) in PBS. Once the feeder cells were detached, a transfer pipette was used to aspirate the EDTA and the dish was washed twice with 5 ml of sterile PBS. HFKs were then removed with 2 ml TrypLE and incubated at 37°C for 5-10 minutes with regular checks for cell detachment. The TrypLE was quenched with CEM and dishes washed with 10 ml media, all of which was transferred to a sterile conical tube.

HFKs were pelleted by centrifugation at $538 \times g$ for 5 minutes at room temperature. The supernatant was removed and the cells resuspended in warmed CEM. HFKs were counted and seeded out at a density of $2 \times 10^5$ cells per dish onto $2 \times 10^6$ irradiated J2-3T3 cells. The
plates were gently rocked several times to ensure thorough mixing and the cells incubated at 37°C and 5% CO₂. The media was replaced every 2 days, with cells taking typically 5-7 days to reach a confluency of 80%. At each passage, aliquots of 5×10⁵ cells were pelleted by centrifugation, washed with phosphate buffered saline (PBS) and stored at -80°C for further analysis. The establishment and maintenance of HPV episomes was confirmed by blot analysis of total DNA (Section 2.2.2.3).

Cell lines containing the HPV16 114/B, HPV16 114/K (Kanda et al., 1988; Roden et al., 1996) or HPV18 (Wilson et al., 2007) episomal genomes were cultured alongside untransfected HFKs derived from the same donor.

2.1.4.5 Maintenance and irradiation of J2-3T3 fibroblasts

Following transfection of primary keratinocytes with HPV genomes, cells were cultivated with a feeder layer of irradiated NIH J2-3T3 mouse fibroblasts. J2-3T3 cells were cultured in DMEM supplemented with L-glutamine (2% v/v) and newborn calf serum (10% v/v, Gibco) and were grown to a confluency of no greater than 90% before passaging. For subculture, 3T3 cells were first washed with 5 mL sterile PBS and then incubated with 1 mL TrypLE until cells were no longer adherent (typically 5 min). Trypsin was quenched and cells washed off of the dish with medium. The suspension transferred to a sterile tube for centrifugation at 10000 × g for 5 min. Cells were resuspended in culture media, counted and seeded into 10 cm dishes with 10 ml of prewarmed medium. J2-3T3 cells were grown to no more than passage 25 and then replaced with earlier passage J2-3T3 cells. When J2-3T3 cells were required as feeder cells, they were harvested by trypsinisation and resuspended in CEM at 2×10⁶ cells/ml. Cells were irradiated with 50 Grays of a Caesium-137 source and were stored in CEM at 4°C for up to 5 days.
For cryopreservation, $2 \times 10^6$ J2-3T3 cells were resuspended in 1 ml DMEM supplemented with 20% (v/v) newborn calf serum, and 10% (v/v) DMSO.

2.1.4.6 Organotypic raft culture

Organotypic raft cultures were prepared as described by Wilson and Laimins (2005). Each collagen plug was prepared in 3 cm dishes using 2.4 ml Rat-tail collagen (BD Biosciences, UK) diluted to 4mg/ml with 0.02 M acetic acid. Per plug, $2 \times 10^6$ 3T3-J2 cells were reconstituted in 0.3 ml 10× filtered DME and 0.3 ml 10× filtered reconstitution buffer (0.05 M NaOH, 260 mM NaHCO$_3$, 200 mM Hepes; all Sigma-Aldrich). To this, the acetic acid and collagen were added and mixed using an ice-cold strippette. The pH of plug mixture was adjusted by addition of 1 M NaOH dropwise, before the addition of 3 ml per 3 cm dish. Plugs were incubated overnight at 37°C 5% CO$_2$. Typically, $2 \times 10^6$ untransfected or genome-containing HFKs were grown on a collagen plug pre-seeded with $2 \times 10^6$ J2-3T3 fibroblasts in a 3 cm tissue culture dish. Media was replaced daily, and cells were allowed to expand on the plugs until the media turned yellow (24-48 hr), after which the collagen plug was carefully lifted and placed on top of a wire mesh platform (known as a grid), and placed in a 10 cm tissue culture dish. This was then filled CEM lacking EGF, in order to create an air/liquid interface (see Figure 2.2). Rafts were grown for 13 days; the media was replaced every 2 days. Six hours before fixing, 20 µM of the thymidine analogue 5-Bromo-2’-deoxyuridine (BrdU, BD Biosciences) was added to media. BrdU is incorporated into the newly synthesized DNA of replicating cells, and thus it allows the detection of replicating cells. In preparation for downstream immunofluorescent analysis (Section 2.3), rafts were fixed by flooding the raft-containing dish with 4% (w/v) paraformaldehyde in DMEM. Rafts were paraffin embedded and sections (4 microns) were prepared for staining by Propath Ltd, UK.
Figure 2.2: Organotypic raft culture

(A) Schematic of the organotypic raft culture system. HFKs containing HPV genomes are seeded onto collagen plugs containing fibroblasts in a 3 cm tissue culture dish and cultured in complete E media.

(B) When HFKs reach confluence, indicated by media turning yellow, the collagen plugs are lifted onto stainless steel support grids. This is cultured with E media supplemented with L-glutamine, forming an air-liquid interface at the raft grid. This induces differentiation of the HFKs, leading to the production of a stratified epithelium.
The metal grids were prepared by cutting squares of wire mesh approximately 4 x 4 cm and bending two opposite edges down evenly. Use of a spirit level confirmed the creation of a level surface. Grids were then sterilized by incubation for 2 hours in concentrated (24 M) sulphuric acid (Thermo Fisher Scientific). For safety, this took place in a glass beaker in a fume cupboard. The grids were rinsed by placement of the beaker under cold running water overnight, after which they were dried on paper, placed in a Pyrex beaker and autoclaved.

2.1.4.7 Protein extraction from raft cultures

Protein was extracted from raft cultures by peeling off the stratified HFKs from the collagen layer followed by homogenization in 1 ml of UTB lysis buffer (8 M urea, 50 mM Tris-HCl pH 8, 0.15 M β-mercaptoethanol, Sigma-Aldrich) supplemented with protease and phosphatase inhibitors (Sigma-Aldrich) in a glass Dounce homogeniser. Lysate was incubated for 30 minutes on ice before being transferred by transfer pipette into a pre-cooled microcentrifuge tube. This was centrifuged at 16000 × g for 15 minutes at 4°C. The supernatant was subsequently removed and designated as the ‘soluble fraction’. The pellet was thus designated the ‘insoluble fraction’, and was solubilised in 250 µl of Laemmlli buffer (BioRad, UK) supplemented with 5% (v/v) β-mercaptoethanol. The soluble fraction was aliquoted at a 1:1 ratio into Laemmlli buffer supplemented with 5% (v/v) β-mercaptoethanol, heated at 95°C for 5 minutes prior to storage at -20°C and/or SDS-PAGE and subsequent analysis.

2.1.4.8 RNA extraction from raft cultures

Similar to the extraction of protein, RNA extraction from raft culture began with removal of the HFKs from the collagen layer using sterile forceps. This was then placed into a 15 ml collection tube and homogenized in 100 µl of RNA Stat-60 (AMS Biotechnology, UK)
using a glass Dounce homogenizer. HFKs were incubated for 5 minutes at room temperature in the RNA Stat-60, which is a monophase solution containing phenol and guanidinium thiocyanate. Once 200 µl of chloroform was added and the sample shaken vigorously for 15 seconds, the homogenate separated into two phases, the aqueous and the organic. After centrifugation at 12000 × g for 5 minutes at 4°C, the total RNA in the aqueous phase was transferred to a fresh collection tube, leaving any DNA and proteins in the organic phase. Addition of 500 µl isopropanol and incubation at room temperature for 5 minutes allowed precipitation of the RNA from the aqueous phase. This was centrifuged at 12000 × g for 10 minutes at 4°C and the pelleted washed by resuspension in 1 ml 75% (v/v) ethanol and centrifugation at 7500 × g for 5 minutes at 4°C. Once the pellet had been dried briefly, it was resuspended in 40µl of nuclease-free water.

2.1.5 Treatment and manipulation of cells

2.1.5.1 Subcellular fractionation

Subcellular fractionation was carried out on HeLa cells and HFKs using the ProteoExtract® subcellular proteome extraction kit (EMD Millipore, USA), and was optimised for use on untransfected and genome-containing HFKs grown in monolayer. All buffers were supplemented with protease and phosphatase inhibitors, and Buffer III was also supplemented with the endonuclease benzonase®. Firstly, supporting 3T3-J2 fibroblasts were first removed and cells washed with ice-cold PBS. Cells were then washed with 1 mL Wash Buffer before incubation in 500 µl extraction Buffer I with gentle agitation at 4°C for 15 min. All traces of the buffer were removed and transferred into a pre-cooled microcentrifuge tube and designated fraction 1 (F1), representing the cytoplasmic fraction of cell contents. Next, 500 µl extraction Buffer II was added to the dish of cells and incubated at 4°C for 30 minutes with gentle agitation. The buffer was then
transferred to a pre-cooled microcentrifuge tube and designated fraction 2 (F2), the membrane fraction. A final incubation in 250 µl extraction Buffer III at 4°C for 10 minutes with gentle agitation was transferred to a pre-cooled 1.5 ml microcentrifuge tube and designated fraction 3 (F3), the nuclear fraction. The remaining fraction, the cytoskeletal fraction, was resuspended in 250 µl extraction Buffer IV, transferred to a pre-cooled 1.5 ml microcentrifuge tube and designated fraction 4 (F4). For subsequent SDS-PAGE and subsequent protein analysis, 10 µl of each fraction was aliquotted into 10 µl Laemmlli buffer supplemented with 5% (v/v) β mercaptoethanol, and treated as a western blot (Section 2.2.4.2).

2.1.5.2 Proteasome inhibition

Untransfected or HPV genome-containing HFKs were seeded into a 6 cm dishes containing 2×10⁵ irradiated J2-3T3 fibroblasts as a feeder layer. HFKs were allowed to grow for 48 h before the addition of the proteasome inhibitor MG132 (z-Leu-Leu-Leu-CHO), which was dissolved in DMSO and applied at a final concentration of 40 µM in cell culture medium. As a control, vehicle (DMSO) treated cells were cultured and harvested alongside MG132 treated cells. Cells were harvested at 4 and 6 h post-treatment; feeder cells were removed and cells pelleted by trypsinisation, as described previously. Cell pellets were washed by resuspension in ice-cold saline and centrifuged for 10000 × g for 5 minutes at 4°C, before being lysed in UTB buffer and subject to downstream western blot analysis (Section 2.2.4.2).

2.1.5.3 Protein kinase A activation and inhibition

As for MG132 treatment, in preparation for treatment with protein kinase A (PKA) inhibitors or activators, untransfected or HPV genome-containing keratinocytes were
seeded into a 6 cm dishes containing a feeder layer of $2 \times 10^5$ irradiated J2-3T3 fibroblasts. HeLa cells were seeded at a density of $2 \times 10^5$ cells/dish. Cells were allowed to attach for 24 h before the addition of specific PKA activating or inhibiting compounds. The PKA activators forskolin (FK) and 3-isobutyl-1-methylxanthine (IBMX, both Sigma-Aldrich) were dissolved in DMSO and used in combination at the final concentrations of 50 $\mu$M and 1 mM in culture medium, respectively. The more specific PKA activator PKI and cyclic AMP (cAMP) analogue, dibutyl cAMP (dbcAMP, EMD Millipore) was dissolved in Hank’s balanced salt solution (HBSS) and diluted in culture medium to a final concentration of 1 mM. The PKA inhibitor H89 (Sigma-Aldrich, UK) was dissolved in DMSO and added to the culture medium at a final concentration of 10 $\mu$M. Exogenous expression of the PKI peptide was induced by transfection with mPKI plasmid (Section 2.1.5.4, gift from M. Diskar). Vehicle treated cells were cultured and harvested alongside those treated with activators or inhibitors. Cells were harvested at various time points up to 96 hr; feeder cells were removed and washed off before cells were harvested by trypsinisation, washed in saline and lysed in UTB for western blot analysis as described in Section 2.2.4.2.

2.1.5.4 Transfection of cells

HeLa, C33A and HFKs were cultured overnight in serum free medium prior to transfection. Cells were transfected at 80% confluency with 3-5 $\mu$g DNA in Xtremegene\textsuperscript{TM} transfection reagent (Roche), at a ratio of 2 $\mu$l per 1 $\mu$g of DNA, in 500 $\mu$l OptiMem (Invitrogen). Reagent, DNA and OptiMem were incubated together for 30 minutes at room temperature before dropwise addition to cells. Cells were then incubated at 37°C with 5% CO$_2$ for 48 hours before harvesting. Transfection with plasmid expressing the peptide PKA inhibitor mPKI was confirmed by observation of GFP expression, using a
Zeiss Axiovert 100 microscope. Cells were washed with ice-cold saline before downstream protein analysis (Section 2.2.4).

2.2 Molecular biology

2.2.1 RNA analysis

2.2.1.1 RNA isolation from cultured cells

Total RNA was extracted from 1×10⁶ cells using an RNeasy Mini kit (QIAgen, UK) and treated with DNase (Promega), according to manufacturer’s instructions. Cultured cells were centrifuged and collected as a cell pellet at the time of passage. The pellet was then washed by resuspension in 500 μl PBS and centrifuged for 5 minutes at 16000 × g. The supernatant was discarded and cells resuspended in 750 μl lysis buffer containing guanidinium thiocyanate, a strong denaturant which lyses cell membranes and rapidly inactivates ribonucleases (Chomczynski and Sacchi, 1987). This was mixed thoroughly by pipetting; any vortex mixing was avoided throughout the preparation in order to prevent shearing of RNA. One volume (750 μl) of 70% ethanol was added to the pellet lysate, and the mixture passed at least 5 times through a 0.9 mm (20 gauge) needle attached to a sterile plastic syringe (BD Biosciences). Lysate was then transferred to a silica filter cartridge inserted into a 1.5 ml collection tube. Centrifugation for 30 seconds at 16000 × g in a microcentrifuge 5415 R, refrigerated benchtop centrifuge (Eppendorf, UK), allowed the mix to be drawn through the filter and the RNA to bind the silica matrix. The bound RNA was then washed by the addition of 700 μl wash buffer RW1 and centrifuged for 30 seconds at 16000 × g. The flow through was discarded. Two sequential washes with 500 μl wash buffer RPE followed; first for 30 seconds at 16000 × g and the second for 2 minutes at 16000 × g, to remove any traces of salts remaining from previous buffer washes. The
flow through was discarded each time. The filter cartridge was then transferred to a clean collection tube and centrifuged for 2 minutes at 16000 × g, in order to remove any last traces of ethanol from the wash solutions. The cartridge was placed into a sterile 1.5 ml collection tube and RNA eluted by the application of 30 µl RNase-free water to the filter, incubation at room temperature for 1 minute, followed by centrifugation for 2 minutes at 16000 × g. To ensure maximum RNA yield, a second application of 30 µl RNase-free water and a final centrifugation for 1 minute at 16000 × g followed.

2.2.1.2 RNA extraction from tissue

In order to extract RNA, tissue samples were first homogenised in 1 mL RNA stat-60 (AMS Biotechnology) per mg of tissue, using a glass Dounce homogeniser. This was incubated for 5 minutes at room temperature in the RNA Stat-60, and 200 µl of chloroform added per 1 ml of RNA stat-60. The sample was shaken vigorously for 15 seconds and incubated for 5 minutes at room temperature. After centrifugation at 12,000 × g for 15 minutes at 4°C, the total RNA in the aqueous phase was transferred to a pre-cooled collection tube, leaving any DNA and proteins in the organic phase. To this, 500 µl isopropanol was added, and incubated at room temperature for 5 minutes to allow the precipitation of the RNA from the aqueous phase. This was centrifuged at 12000 × g for 10 minutes at 4°C and the pelleted washed by resuspension in 1 ml 75% (v/v) ethanol and centrifugation at 7500 × g for 5 minutes at 4°C. Once the pellet had been dried briefly, it was resuspended in 40µl of nuclease-free water.

2.2.1.3 DNase treatment of RNA

RNA was treated with DNase in order to degrade any contaminating DNA. One microgram of eluted RNA was mixed with 1 µl of 10× DNase buffer (Promega, UK; 100 mM Tris-HCl, 25 mM MgCl₂, 1 mM CaCl₂), 1 µl DNase I (2000 units/ml, Promega) and
 MATERIALS AND METHODS

2.2.1.4 Spectrophotometry of nucleic acids

RNA and DNA samples were quantified using a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific). The sample pedestal was first cleaned with 2 µl nuclease-free water and then calibrated with 1 µl elution solution, to provide a reference measurement of background light absorbance against which sample light absorbance could be compared. One µl of each sample was pipetted onto the sample pedestal and it’s absorbance measured at 280 nm. The ND-1000 Spectrophotometer v 3.2 software (Thermo Fisher Scientific) calculated the concentration (ng/µl) using a modified form of the Beer-Lambert equation and the wavelength-dependant molar absorptivity (or extinction) coefficient to correlate measured light with sample concentration:

\[ c = \frac{(A \times e)}{b} \]

Where: 
- \( c \) = nucleic acid concentration (ng/µl) 
- \( A \) = Absorbance (absorbance units, AU) 
- \( e \) = extinction coefficient (ng-cm/ul) 
- \( b \) = light path length (cm)

2.2.1.5 Reverse transcription

RNA from cultured cells was reverse transcribed using the Promega MLV reverse transcription (RT) system to produce complementary DNA (cDNA). One microgram of RNA was preincubated with random hexodeoxynucleotide (100 mM) or oligodT (100 mM) primers at 72°C for 10 minutes in a reaction volume made up to 12 µl with sterile distilled water to a final volume of 10 µl. This was incubated at 37°C for 30 min, after which the reaction was inactivated by the addition of 1 µl stop solution and incubation at 65°C for 10 min. This was stored at -20°C for short term, or at -80°C for longer term.
water. This was then cooled on ice, to denature any secondary RNA structures that would interfere with primer annealing. To this mixture, 5 µl of 5× reaction buffer, deoxynucleotides (10 mM, Promega), 1 µl Moloney murine leukemia virus reverse transcriptase (MLV RT, 20 units/µl, Promega), 1 µl RNase inhibitor (40 units/µl, Promega) and nuclease-free water, were added to a final volume of 25 µl (Thermo Fisher Scientific). Reactions were incubated in a heat block at 37°C for 1 hour, after which the enzyme was denatured by incubation at 70°C for 10 min. To confirm the absence of genomic DNA during downstream applications, concurrent reactions were performed in the absence of reverse transcriptase.

2.2.1.6 Oligonucleotide primer design

Oligonucleotide primers were designed using primer 3 software (www.primer3.com) and previously published genomic and mRNA transcript sequences available from the NCBI Genbank database (http://www.ncbi.nlm.nih.gov/genbank/). The following formula was used to calculate the melting temperature ($T_m$) of designed primers:

$$T_m = \%GC \times 0.41 + 64.9 - \left(\frac{600}{R}\right)$$

where: $T_m$ = melting temperature of primers

\%GC = percentage of G and C bases in primer

$R$ = total number of bases in primer

Primer sets (forward and reverse pairs) were designed to have similar GC content and total number of base pairs (bp) to produce complementary primers with similar annealing temperatures, at around 60°C (Table 2.2). Basic local alignment search tool (BLAST) searches (http://www.ncbi.nlm.nih.gov/BLAST/) were performed in all cases to ensure that primers were not complementary to other regions of the genome and bound specifically to
the target sequences. Following oligonucleotide synthesis (Alta Bioscience, Birmingham UK), optimum conditions for each primer set were determined by performing parallel reverse transcription polymerase chain reactions (RT-PCR) at different temperatures based around the calculated Tm (Table 2.2).

Table 2.2: qRT-PCR primers

<table>
<thead>
<tr>
<th>Target protein</th>
<th>Forward primer sequence (5’-3’)</th>
<th>Reverse primer sequence (5’-3’)</th>
<th>Amplified Fragment (bp)</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCRIB</td>
<td>CCTGCCAGCTCCAGCACCAC</td>
<td>CGGCTGGGGTTGGGGCAGTTA</td>
<td>147</td>
<td>59</td>
</tr>
<tr>
<td>INADL</td>
<td>AGATCGGAGTTTTGCAAGGT</td>
<td>TCGAAGCAATGTCTTGGACCCA</td>
<td>100</td>
<td>59</td>
</tr>
<tr>
<td>NHERF1</td>
<td>TCAACCAATGGGAGATACAG</td>
<td>GTCTTGGGAATTCAAGCTCC</td>
<td>126</td>
<td>59</td>
</tr>
<tr>
<td>PTPN13</td>
<td>ATGAAACCTGTGCTAGGGAT</td>
<td>GACCTTCAGGCTCAAGTAGT</td>
<td>286</td>
<td>60</td>
</tr>
<tr>
<td>MAGI-1</td>
<td>GACTGTGAAGGATTCTTGGA</td>
<td>TCGTGAATCTTTCCACTTC</td>
<td>125</td>
<td>59</td>
</tr>
<tr>
<td>MAGI-2</td>
<td>TTCAAGGGACAGACCAACC</td>
<td>GGAAGAAGGACGACTTTGCG</td>
<td>146</td>
<td>59</td>
</tr>
<tr>
<td>MAGI-3</td>
<td>CCACAAACTGCAGCAAGTGA</td>
<td>AATTGCTTCATTCTTTCCAGC</td>
<td>150</td>
<td>59</td>
</tr>
<tr>
<td>B2M</td>
<td>GCCTGCGGTGTCAACCAGT</td>
<td>GCAGCCTCCTAGAGCTACC</td>
<td>296</td>
<td>59</td>
</tr>
<tr>
<td>DLG1α</td>
<td>CAGTGGAGAGAAGGCACAGTCA</td>
<td>AGCGTGAAGGTCTGCTGCA</td>
<td>103</td>
<td>59</td>
</tr>
<tr>
<td>DLG1β</td>
<td>CGGAAGCAGATACCAGAG</td>
<td>AATTGGTTCAGAGCGCTTTG</td>
<td>113</td>
<td>59</td>
</tr>
<tr>
<td>DLG1 HOOK region</td>
<td>GCCCATTTCCCTAGGCATGAA</td>
<td>GCCATTTCGGTTGGATGCAGAG</td>
<td>98</td>
<td>59</td>
</tr>
</tbody>
</table>

Tm indicates annealing temperature used in PCR reactions, having calculated predicted Tm and carried out reactions at several temperatures to obtain an optimal Tm. B2M – β2-microglobulin.

2.2.1.7 Reaction conditions for RT-PCR

RT-PCR reactions were set up on ice using filter pipette tips (StarLab, UK). Each 25 μl RT-PCR reaction contained; 5 μl cDNA, 12.5 μl hotstart GoTaq mastermix (Promega) and 0.1 nM of forward and reverse primers, with nuclease-free water made up to the final volume. Reactions were performed in a 2720 Thermocycler (Applied Biosystems, UK),
beginning with 10 minutes at 95°C, in order to activate the polymerase. This was followed by 40 cycles of denaturation at 95°C for 30 seconds, annealing for 50 seconds) and an extension step at 72°C for 1 minute.

Table 2.2 details the primer sequences and optimal annealing temperatures. Amplified DNA fragments of expected length were visualized by 1% (w/v) agarose gel electrophoresis and ethidium bromide staining.

2.2.1.8 Agarose gel electrophoresis

DNA from PCR, as well as plasmid DNA, was visualised using agarose gel electrophoresis. Fragments were resolved on a 1-2% (v/w) agarose gel, prepared by heating 1.5-3 g of molecular grade agarose (Bioline) in 150 ml 0.5× TBE buffer (0.1 mM Tris-HCl, 0.1 mM Boric acid, 100 mM EDTA). Ethidium bromide was added to a final concentration of 0.5 µg/ml prior to cooling and the entire solution poured into a Mini Sub Gel GT Electrophoresis Tank (BioRad). Once set, this was submerged in 0.5× TBE buffer. DNA samples were combined with 5 µl loading buffer (30% v/v glycerol, 0.3% w/v bromophenol blue, 0.1% xylene cyanol FF, 10× TBE) and loaded onto the gel, alongside a lane containing 2 µl DNA marker (1kb plus, Invitrogen). A potential of 100 V was applied across the gel for approximately 1 hr. Ethidium bromide is an intercalating agent that fluoresces when exposed to ultraviolet (UV) light, allowing the visualization of DNA, using a Gene Flash UV light box (Syngene Bio Imaging, UK).

For the resolution of genomic DNA prior to Southern blot analysis, 0.8% (w/v) agarose gels were prepared as described and cast in Fisherband horizontal gel electrophoresis tanks (Thermo Fisher Scientific), run at 50 V overnight.
2.2.1.9 Purification of cDNA from agarose gels

cDNA was isolated from agarose gels and purified using the QIAquick Gel Extraction Kit (QIAgen), as per the manufacturer’s instructions. Each DNA band (approximately 100 mg of gel) was excised from the gel using a sterile scalpel and placed in a sterile microcentrifuge tube. Once weighed, three volumes (typically 300 µl) of QG buffer was added and the gel dissolved by incubation at 50°C with occasional agitation. Once the gel had completely dissolved (approximately 15 min), 1 volume (typically 100 µl) isopropanol was added and mixed by inversion of the tube several times, aiming to precipitate the DNA. The solution was transferred to a QIAquick Spin column placed in a fresh microcentrifuge tube and drawn through the column by centrifugation at 16000 × g for 1 minute, allowing cDNA to bind to the filter cartridge. The flow-through was discarded, and any traces of agarose were washed away by the addition of 500 µl QG buffer and centrifugation at 16000 × g for 1 minute. To remove salts and ethidium bromide, a second wash was performed by the addition of 750 µl PE buffer and centrifugation at 16000 × g for 1 minute. The flow-through was discarded after each wash stage. The filter column was again centrifuged at 16000 × g for 1 minute to remove any remaining PE buffer before being transferred into a sterile microcentrifuge tube. Thirty microlitres of sterile, nuclease-free water was added to the centre of the filter and the column incubated for 1 minutes at room temperature. The column was then centrifuged at 16000 × g for 1 minutes to elute the DNA. This addition of water was repeated for a second time, and DNA concentration measured using a NanoDrop spectrophotometer before storage at -20°C.

2.2.1.10 DNA sequencing

RTPCR products were sequenced in order to confirm that the intended genes of interest had been amplified. Sequencing was performed by the Functional Genomics facility.
(Department of Biosciences, University of Birmingham, UK). In preparation, 10 μg of RTPCR product and 5 picomoles forward or reverse primer was made up to a final volume of 10 μl with nuclease-free water. Returned chromatograms were viewed using FinchTV software (GeoSpiza Inc.) and the EMBL-EBI Clustal-Omega facility (http://www.ebi.ac.uk/Tools/msa/clustalo/) used to align sequencing results to published sequences and confirm homology.

**Table 2.3: DLG1 primers**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5'-3')</th>
<th>Region of DLG1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLG1 F1</td>
<td>CCCAGAGACATGGCACCC</td>
<td>α N terminus</td>
</tr>
<tr>
<td>DLG1 F2</td>
<td>CAGAGACCAGGAGAATGCTGCC</td>
<td>Nts 428-446</td>
</tr>
<tr>
<td>DLG1 F3</td>
<td>CTGACACGATGATCATATTG</td>
<td>Nts 428-446</td>
</tr>
<tr>
<td>DLG1 F4</td>
<td>CAGAAGACGGGTTAATGCTG</td>
<td>Nts 2140-2160</td>
</tr>
<tr>
<td>DLG1 F5</td>
<td>CTCGACCAGTGGATCATATTG</td>
<td>Nts 2140-2160</td>
</tr>
<tr>
<td>DLG1 F6</td>
<td>CTGACACGAGGTTAATGCTG</td>
<td>GUK domain</td>
</tr>
<tr>
<td>DLG1 F7</td>
<td>GAGATCTCTGACGCGATGGG</td>
<td>Insertion 2</td>
</tr>
<tr>
<td>DLG1 F8</td>
<td>TTCACTATCGCTGGCATTAGAA</td>
<td>Insertion 5</td>
</tr>
<tr>
<td>DLG1 F9</td>
<td>CCTGTCTTTCATAGGTTGCA</td>
<td>GUK domain</td>
</tr>
</tbody>
</table>

### 2.2.1.11 Quantitative RTPCR (qRTPCR)

qRTPCR reactions were performed to determine the level of expression of particular genes of interest, using SYBR green chemistry. SYBR Green I is an assymetrical cyanine dye which intercalates into the minor groove of double-stranded DNA (dsDNA) and produces a complex that emits a green fluorescent signal (524 nm) upon excitation with blue light (485 nm) (Zipper et al., 2004). The intensity of the fluorescent signal is proportional to the amount of SYBR-bound DNA present and so measurement of the fluorescent signal can be used to quantify DNA. During the qPCR reaction, the fluorescent signal is measured at every cycle of PCR and immediately following the product extension phase. During the reaction, the number of copies of the gene increases exponentially as described by the following equation (Livak and Schmittgen, 2001):

\[
X_n = X_0 \times (1 + E_x)^n
\]
Where: 
\[ X_n = \text{number of molecules of gene of interest at cycle } n \]
\[ X_0 = \text{initial number of molecules of gene of interest} \]
\[ E_x = \text{efficiency of amplification of gene of interest} \]
\[ n = \text{number of completed cycles} \]

The values obtained for the genes of interest were referenced to the values of expression of the housekeeping gene β2-microglobulin, as an internal control to account for difference in cell number and initial RNA load (Schmittgen and Zakrajsek, 2000). Prior to analysis of gene expression, several housekeeping genes of choice were compared by qRTPCR for consistency of expression across cDNA samples. The result of gene expression quantification was then given as a ratio of the levels of mRNA transcripts of the gene of interest in an experimental sample (HPV genome-containing keratinocytes) to a defined control sample (untransfected HFK).

Each qRTPCR reaction contained 40 ng cDNA template, 12.5 µl SYBR green mastermix (Applied Biosystems) and 0.1 µM forward and reverse primer, made up to a total volume of 25 µl with nuclease-free water. Reactions were set up in triplicate on ice in 96-well reaction plates (Applied Biosystems) and analysed using the ABI Prism 7700 Sequence Detection System (Applied Biosystems). Samples were initially subjected to a denaturation step at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 20 seconds, annealing for 45 seconds and an extension step at 72°C for 1 minute. Primers specific to the housekeeping gene β2-microglobulin were included on each qRTPCR plate, to account for loading variability.

2.2.1.12 Comparative Ct method (ddCt) for relative quantitation of gene expression

The delta-delta (dd)Ct method enables relative quantitation of transcripts without the need for standard curves when looking at expression levels of a target gene relative to an endogenous control (e.g. β2-microglobulin, Livak and Schmittgen, 2001). This method was
used to measure the amount of target gene transcripts in different samples. Firstly, the difference (dCt) between the Ct values of the target gene and the endogenous gene was calculated for each sample studied (dCt = target Ct - endogenous Ct). The sample selected as the baseline sample for expression of the target gene is referred to as the reference sample. Here, untransfected, non-HPV containing keratinocytes were used as the reference sample, to which gene expression in the HPV-containing keratinocytes were normalised, unless otherwise stated. The difference between the dCt values of each sample and its reference sample was calculated, generating a ddCt value (ddCt = reference dCt - target dCt). The ddCt for each sample was then converted to an absolute value using the following equation:

\[ \text{fold change in expression level} = 2^{-\text{ddCt}} \]

Where: \( \text{dCt} = \text{target Ct} - \text{endogenous Ct} \)
\( \text{ddCt} = \text{reference dCt} - \text{target dCt} \)

2.2.2 DNA analysis

2.2.2.1 DNA extraction

DNA was extracted from cells using a standard phenol-chloroform method: \(2 \times 10^6\) cells were resuspended in lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM EDTA; 400 mM NaCl; all Sigma-Aldrich), treated with RNAse A at a final concentration of 50 µg/ml for 30 minute at 37°C, and digested overnight by proteinase K (50 µg/ml final concentration, Sigma-Aldrich) at 37°C with 0.2% (w/v) SDS. The resulting digestion was passed through a blunt 0.8 gauge needle 10 times to shear the DNA. The DNA was extracted firstly by washing twice with 2 volumes of phenol:chloroform:isoamylalcohol (25:24:1, Sigma-Aldrich), mixing thoroughly and centrifuging at 16000 \(\times\) g and washing once with 2
volumes of chloroform:isoamyl alcohol (24:1). Following each wash, the liquid phase was transferred to a fresh 15 ml conical tube. DNA was precipitated by incubation in two volumes of 100% ethanol, with 0.3 M sodium acetate (pH 5.2), overnight at -20°C. Tubes were centrifuged for 30 minutes at 4°C, the pellet was washed with ice-cold 70% ethanol and centrifuged for 15 minutes at 4°C. The resulting pellet was dried and resuspended in 100 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). DNA concentration was measured, as previously described, using a NanoDrop ND-1000 Spectrophotometer.

2.2.2.2 DNA digestion

Restriction enzyme digestions were typically performed upon 2-10 µg of plasmid DNA, using enzymes purchased from New England Biolabs (NEB, UK). Enzymes, DNA and the appropriate buffer were incubated in a total volume of 20-50 µl at 37°C overnight. After incubation, enzymes were heat-inactivated by incubation at 65°C for 10 minutes followed by cooling on ice for 2 minutes. Digestion with multiple restriction enzymes were carried out by selecting the appropriate buffer for the activity of both restriction enzymes, according to the manufacturer’s guidelines.

2.2.2.3 Preparation of DNA for Southern analysis

For the analysis of HPV genomes, 5 µg of genomic DNA extracted from cells was subjected to restriction enzyme digest (Table 2.4). All digestions also included the enzyme DpnI (NEB) to remove residual input DNA. Digested genomic DNA was analysed by electrophoresis using a 0.8% (w/v) agarose gel. To enable quantification of genome copy numbers, copy number standards were generated by digestion of pGEMII-HPV18 with EcoRI (NEB) to release the HPV18 genome from the vector and the equivalent 5 and 50 genome copies/cell were run alongside samples as standards. In the case of HPV16 copy number standards were generated by digestion of pUC-HPV16 with BamHI (NEB).
Table 2.4: Restriction enzymes used to assess HPV genome status

<table>
<thead>
<tr>
<th>HPV genome</th>
<th>Restriction Enzyme</th>
<th>Linearises genome</th>
<th>Non-cutter</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV16</td>
<td>HindIII</td>
<td></td>
<td>BamHI</td>
</tr>
<tr>
<td>HPV18</td>
<td>EcoRI</td>
<td></td>
<td>BglII</td>
</tr>
</tbody>
</table>

2.2.2.4 Transfer of DNA from agarose gel to nylon membrane

Prior to blotting, the agarose gel was washed twice in 250 mM HCl for 20 minutes (the depurination stage), and twice in 400 mM NaOH for 30 minutes (denaturation and neutralization stages), all at room temperature. To set up the transfer of DNA from the agarose gel to a Gene Screen Plus nylon membrane (Perkin Elmer, UK), a tray was filled with 1 L of 400 mM NaOH and a glass plate rested on top. Three sheets of 24 × 33 cm Whatman™ 3MM paper were soaked in 400 mM NaOH, laid across the glass plate with both ends submerged in the NaOH reservoir, and all the bubbles removed. The washed agarose gel was placed on top of this paper wick, with the loading wells facing downwards. A 20 × 22.5 cm sheet of Gene Screen Plus nylon membrane (Perkin Elmer) pre-soaked in NaOH was placed on top. Four layers of 21 × 23.5 cm Whatman™ soaked in 0.4 M NaOH were placed onto the membrane and the bubbles removed. On top of this, two stacks of absorbent paper towels were placed, to reach approximately 10 cm in height and covering the filter paper layers, with no gaps between the stacks. A glass plate was placed in top of the towels, and a weight centered on top. The filter paper layer was separated from the wick via plastic wrap placed between the top and bottom filter paper layers. This set up was left overnight for DNA to transfer by capillary action.
Upon disassembly, the position of the loading wells was marked onto the membrane with a pencil, and the DNA crosslinked to the membrane using a UV crosslinker 90 (Stratalinker, Stratagene, UK) on auto-crosslink mode. Following rinsing in 2× saline-sodium citrate (SSC; 300 mM NaCl, 30 mM sodium citrate, pH 7.0), the membrane was stored at -20°C in plastic wrap.

### 2.2.2.5 Preparation of radiolabelled DNA probe

In preparation for generation of a HPV18 DNA probe, the pGEMII-HPV18 vector was linearised by EcoRI digestion. To generate a HPV16 DNA probe, the pUC-HPV16 vector was linearised by digestion with BamHI. Probe DNA was separated from the vector backbone by agarose gel electrophoresis, and purified as previously described. Fifty nanograms of DNA was subsequently diluted into 45 µl of TE buffer and the DNA denatured by heating to 95°C for 5 minutes, followed by incubation on ice slurry for 2 minutes. The denatured DNA was then used as a template for radiolabelled probe generation using the Ready To Go DNA labelling beads-dCTP kit (Amersham Biosciences, UK), according to manufacturer’s instructions. Following re-suspension of the labelling beads with the denatured linear DNA, 50 µCi of [α-³²P] dCTP (Perkin Elmer) was added and the mixture left to incubate at 37°C for 30 minutes. The labelled probe was purified using an Illustra Probe Quant G-50 microcolumn (Amersham Biosciences), a radiolabelled probe purification kit, following manufacturer’s instructions.

### 2.2.2.6 Hybridisation of radiolabelled probe to DNA

To prepare the hybridisation buffer, a 2× hybridisation buffer (5× SSC; 750 mM NaCl, 1.5 M sodium citrate, pH 7.0), 10× Denhart’s (0.2% w/v ficoll 400, 0.2% w/v polyvinylpyrolilone, 0.2% w/v BSA fraction V, and 20% w/v dextran sulphate, all Sigma-Aldrich), was diluted with deionized formamide (Sigma-Aldrich) at a 1:1 ratio, and SDS
added to a final concentration of 0.1% (w/v). A 200 µl aliquot of 10 mg/ml salmon sperm DNA (Invitrogen) was denatured by heating to 95°C for 5 minutes, cooled on ice slurry for 2 minutes, and subsequently diluted into 10 ml of 1× hybridisation buffer; this constituted the pre-hybridisation buffer. The nylon membrane onto which the DNA was immobilised was then defrosted and placed onto a damp square of gauze. Together, the gauze and membrane were rolled and carefully placed into a glass hybridisation canister (Hybaid, UK). The pre-hybridisation buffer was added and the canister incubated in a hybridisation oven (Hybaid), rotating at 42°C for 1 hr, during which hybridisation buffer containing the radiolabelled DNA probe was prepared. A 200 µl aliquot of (10 mg/ml) salmon sperm DNA was added to the radiolabelled probe, followed by denaturation 95°C for 5 minutes and cooling on ice slurry for 2 minutes. This was added to 10 ml 1× hybridisation buffer. Once the pre-hybridisation buffer was removed, the hybridisation buffer containing the probe was added to the canister and this was incubated overnight, rotating at 42°C.

2.2.2.7 Stringency washes

Following removal from the hybridisation canister, the nylon membrane was rinsed in Buffer I (2× SSC, 0.1% w/v SDS), using a sponge to firmly wipe the surface of the membrane to remove excess, unbound probe. The membrane was then washed twice for 15 minutes in the same buffer, with gentle agitation. Two successive 15 minute washes in Buffer II (0.5× SSC, 0.1% w/v SDS) followed, and then two 15 minute washes in Buffer III (0.1× SSC, 0.1% w/) SDS), both with gentle agitation at room temperature. Following this, the membrane was then washed in Buffer IV (0.1× SSC, 1% w/v SDS) at 65°C for 30 min. The membrane was then wrapped in plastic wrap and exposed to autoradiography film (GE Healthcare), in a cassette containing an intensifying screen. In order to slow the β-particles, this was stored at -20°C. Films were typically developed after 24 h and another
after 1 week exposure times, using a Compact ×4 automatic processor (Xograph Healthcare, UK).

2.2.3 Bacterial culture

2.2.3.1 Bacterial growth

The DH5α strain of *Escherichia coli* (*E. coli*) was used as a bacterial host for the amplification of plasmid vectors. Bacteria were grown on Agar plates consisting of 5% w/v agar in Luria-Bertani (LB) medium (1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 1% w/v NaCl, all Sigma-Aldrich), which was sterilized by autoclaving. To make up ampicillin plates, the agar medium was melted, cooled to approximately 50°C and ampicillin (Sigma-Aldrich) added to a final concentration of 100 µg/ml. This was then dispensed into petri dishes in 25 ml aliquots and left in a sterile environment to dry. Excess plates were sealed and stored at 4°C prior to use.

2.2.3.2 Bacterial transformation

The chemically competent DH5α strain was used to amplify plasmids. To do so, *E. coli* was first thawed on ice and 30-50 µl of cells pipetted into a pre-cooled tube. To this, 1-2 µg of plasmid DNA containing ampicillin resistance gene, was added directly into the thawed cells and incubated on ice for 30 min. The *E. coli* were heat shocked at 42°C for 45 seconds, followed by a 2 minute incubation on ice and the addition of 500 µl LB media (1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 1% w/v NaCl). This was incubated at 37°C shaking at 200 rpm for 1 hr. From this, 50-200 µl of transformed cells were plated onto ampicillin agar plates and incubated overnight at 37°C.

A starter culture of transformed cells was prepared by inoculating 6 ml LB media supplemented with 100 µg/ml antibiotic, with a single bacterial colony and incubating
overnight in a shaker at 37°C. Glycerol stocks were also prepared from the starter culture by the addition of 500 µl of sterile 80% (v/v) glycerol solution to a 500 µl aliquot of bacteria. To amplify bacteria, the 6 ml preparation was added to 200 ml of ampicillin-supplemented LB and grown overnight in a 2 L conical flask, shaking at 200 rpm at 37°C.

2.2.3.3 Storage of E.coli

Bacteria were grown from glycerol stocks by streaking out the bacteria across the surface of an agar plate with a sterile fine loop, which was then incubated overnight at 37°C. Transformed bacteria were stored on agar plates for short periods of time at 4°C, or for longer periods of time, stored as a glycerol stock at -80°C. Untransformed stocks of E.coli were stored at -80°C.

2.2.3.4 Purification of plasmid DNA

Following incubation overnight in 200 ml LB supplemented with antibiotic, the overnight culture was pelleted by centrifugation at 13,000 rpm in a Sorvall GS3 rotor at 4°C for 20 min. Amplified plasmid DNA was extracted using the QIAprep Spin MaxiPrep Kit (Qiagen), following manufacturer’s instructions. The purification procedure involved lysis of bacterial cells using alkaline lysis buffer, adsorption of plasmid DNA onto a silica filter in the presence of high salt buffer, followed by washing and elution of plasmid DNA from that filter.

Plasmid DNA was precipitated by the addition of 0.7 volumes of isopropanol (Thermo Fisher Scientific) and the sample was then centrifuged at 3000 × g in a Sorvall SS-34 rotor for 40 minutes at 4°C. The pellet was then washed with 5 ml 70% (v/v) ethanol and centrifuged for a further 15 minutes at 4°C. DNA was air-dried and resuspended in 50 µl of 1× Tris-EDTA (TE) buffer for storage at 4°C (short term) or -20°C (long term). The concentration of DNA was determined by spectrophotometry (Nanodrop).
2.2.3.5 Plasmid vectors

The pGEMII-HPV18 plasmid contains the total HPV18 genome (accession number: NC 001357) cloned into the pGEMII vector at the EcoRI restriction site at residue 2440. The pUC-HPV16 plasmid contains the total HPV16 genome cloned into the pUC vector at the BamHI restriction site. The HPV18 genome containing plasmid was a gift from Frank Stubenrauch, University of Tubingen, and the HPV16 114/K (HPV16K) and 114/B (HPV16B) containing plasmids were both provided by Ethel de Villiers, the German Cancer Research Centre, Heidelberg.

2.2.3.6 Preparation of HPV genomes for transfection

To extract the HPV18 genome from the bacterial vector, 10 µg of pGEMII-HPV18 wild type or mutant genomes were digested with EcoRI. Similarly, the HPV16 genome was extracted from the pUC-HPV16 genome by digestion with BamHI; both digestions occurred in a total volume of 50 µl and were incubated overnight at 37°C. Digested HPV genomes were recircularised in 900 µl total volume ligation reaction mix with T4 DNA ligase containing 400 units/ml (NEB), to encourage self-ligation of the HPV genomes. This was incubated overnight at 16°C. HPV DNA was then precipitated by the addition of 2 volumes of isopropyl alcohol and 1/5 volume of 5 M NaCl, followed by vortexing and subsequent incubation at -20°C overnight. DNA was pelleted by centrifugation at 16,100 × g for 30 minutes at 4°C, the supernatant removed and the pellet washed with 70% ethanol (pre-cooled to -20°C), prior to centrifugation at 16,100 × g for a further 15 minutes at 4°C. The supernatant was subsequently removed and the DNA pellet resuspended in 12 µl of 1× TE buffer. The DNA concentrations were then estimated by running a 1 µl aliquot on an agarose gel alongside 1 µl of DNA marker (1 µg/ml, 1kb plus ladder, Invitrogen).
2.2.4 Protein analysis

2.2.4.1 Protein extraction and preparation

Cells were harvested by trypsinisation, washed twice with PBS and finally with ice-cold saline. Pelleted cells were lysed in RIPA (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% v/v SDS, 2% sodium deoxycholate, 1% v/v Triton-x-100), UTB or NP40 (200 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% v/v NP40, Sigma-Aldrich) lysis buffers, as appropriate and incubated on ice for 20 minutes and then sonicated for ten seconds at 2 watts. This was then centrifuged at 4°C for 20 minutes at 16000 × g, to remove the insoluble fraction. The amount of protein was quantified by Bradford assay. Bradford reagent (Biorad) was diluted 1:4 in distilled water, and standards of bovine serum albumin (BSA) at 4, 8, 12, 16 and 20 µg/mL set up. The standard curve was set up and absorbance measured using an Eppendorf Biophotometer (Eppendorf).

Prior to loading, the protein of samples were heated to 95°C in an equal volume of 2× Laemmli buffer (Sigma-Aldrich) supplemented with 5% (v/v) β-mercaptoethanol, for 5 min, cooled on ice and spun quickly to collect the contents of the tube.

2.2.4.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels were set up in two phases using the mini-Protean 3 Bio-rad apparatus. A 10 or 12% resolving gel containing 30:1 acrylamide:bisacrylamide (Bio-rad), 390 mM Tris-HCl pH 8.8, 0.1% w/v SDS (Sigma-Aldrich), 0.06% w/v N,N,N’,N’- tetramethylethylenediamine (TEMED, Sigma-Aldrich) and 0.1% w/v ammonium persulphate (APS, Sigma-Aldrich) was made. A layer of isopropanol was applied to the top to give a uniform surface. Once this had set, the isopropанol was removed and a stacking gel containing 3% acrylamide, 125 mM Tris-HCl pH 6.8, 0.1% w/v SDS, 0.1% w/v TEMED and 0.1% w/v APS was poured on top of
the resolving gel, a comb was placed in the stacking gel and allowed to set for 30 min. Protein lysates were added to equal volumes of loading buffer (5% v/v β-mercaptoethanol in Laemmli buffer), heated at 95°C for 5 min, cooled on ice and centrifuged quickly to collect. Samples were then loaded into each SDS-PAGE gel well alongside a lane containing 7 µl of ColorPlus Prestained protein ladder (NEB), and separated by electrophoresis at 100V for 1 hr. Gels were run using Mini Protean® 3 Cell (BioRad) tanks, containing 1× running buffer consisting of 25 mM Tris-HCl pH 7.5, 2 mM glycine (Thermo Fisher Scientific) and 10% SDS.

2.2.4.3 Electrophoretic protein transfer

Two pieces of sponge, two pieces of Whatman™ chromatography paper (Sigma-Aldrich) and one piece of BioTrace NT membrane (VWR International, UK), cut to gel size, were soaked in transfer buffer (25 mM Tris-HCl pH 7.5, 2 mM glycine, 20% v/v methanol in distilled water). The ‘transfer sandwich’ was set up in (BioRad), ensuring air bubbles were pushed out between each layer, in the following order: sponge, filter paper, gel, BioTrace NT membrane, filter paper, sponge. The transfer module was then placed into the transfer tank (BioRad) and filled with transfer buffer. Protein was then transferred from the gel to the BioTrace NT membrane at 350 mAmp for 3 hr.

2.2.4.4 Protein visualisation

Successful transfer was assessed by Ponceau S stain (0.1% w/v Ponceau S, 5% acetic acid), which was then rinsed off in phosphate- or Tris-HCl-buffered saline (PBS or TBS) as appropriate. Non-specific protein binding was blocked by incubating the membrane for 1 h at room temperature in 2.5% (w/v) non-fat milk (Marvel, Premier Foods) or 2% (w/v) BSA (Sigma-Aldrich) dissolved in PBS or TBS. The membrane was incubated overnight at 4°C in primary antibody, diluted in the appropriate milk/BSA PBS/TBS solution (see
To assess for equal loading, the levels of housekeeping genes β-actin or GAPDH were determined using a mouse anti-β-actin monoclonal antibody, or mouse anti-GAPDH antibody, respectively. After three 15 minute washes in TBS or PBS with 0.1% (v/v) Tween-20, the membrane was incubated for 1 h in HRP-conjugated secondary IgG (Table 2.6) and then washed as before, with the last wash in TBS or PBS without Tween. Antibody-protein complexes were detected using the enhanced chemiluminescence (ECL) kit (Amersham Biosciences). 1 ml of each ECL reagent was mixed in a universal tube immediately prior to use. The membrane was incubated in the ECL mixture for 1 minutes before being wrapped carefully in plastic wrap and placed into an autoradiography cassette, containing X-ray film (Hyperfilm™, Amersham Biosciences) and developed for between 30 seconds to 1 hr, using a Compact ×4 automatic processor (Xograph Healthcare) at various exposure times (typically 30 seconds, 5 minutes and 30 minutes).

2.2.4.5 Generation of DLG1α specific antibody

An immunogenic peptide was selected from the N-terminus sequence unique to DLG1α. This peptide (KGSSDELQAEPEPSR) was inoculated into rabbits by a commercial company (Generon, UK). The subsequent antibody (DLG1α 1524) recognised the peptide in ELISA based assay (performed by Generon). Rabbit serum extracted before exposure to immunogen (pre-immune) was taken for use as a control for non-specific binding.

2.3 Immunocytochemical techniques

2.3.1 Immunofluorescence

A heat-induced epitope retrieval (HIER) technique using an EDTA buffer was applied to paraffin-embedded and formalin-fixed samples, in order to unmask the antigens in the
formalin-fixed sections and thus enhance the staining intensity of antibodies (Watson et al., 2002). Firstly, 4 nm raft sections were incubated for 5 minutes in Histo-Clear™ (National Diagnostics, UK), to clear the paraffin. This was followed by hydration for 15 minutes in 100% ethanol (IMS). Slides were then rinsed three times in water, incubated for 10 minutes in 0.3% (v/v) hydrogen peroxide, and washed three times in water again. The sections were immersed overnight in EDTA buffer (1 mM EDTA NaOH pH 8, 0.1% v/v Tween) heated to 65°C and continually agitated on a stirrer.

Retrieved raft sections were blocked in 5% v/v heat-inactivated goat serum (HINGS) with 0.1% w/v BSA in PBS, for 1 h at room temperature. During the staining procedure, incubations took place in a humidified chamber, in order to prevent the evaporation of solutions. Primary antibody was made up in blocking buffer, applied to raft sections and incubated overnight at 4°C. Excess antibody was removed by washing slides in agitated PBS for 10 min, three times. The appropriate secondary fluorophore-conjugated antibody (Molecular probes®, Life Technologies), made up in the same blocking buffer, was applied to each slide, incubated for an hour at 37°C and washed in agitated PBS as before. The final wash contained 1 µl of 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich), a blue (461 nm) fluorescent stain that binds strongly to A-T rich regions of DNA, highlighting the nuclei of cells. Slides were mounted onto coverslips in 80% (v/v) glycerol in PBS containing 2% 4-Diazabicyclo-2,2,2-octane (DABCO, Sigma-Aldrich) and visualized on a Nikon Eclipse E600 microscope (Nikon, USA). Images were captured using a Leica DC200 camera and software. Slides were stored at -20°C.
2.4 Computational methods

2.4.1 Protein and nucleotide sequence alignment
Protein and nucleotide sequences were obtained from NCBI and Ensembl databases (http://www.ncbi.nlm.nih.gov, Altschul et al., 1990 and http://www.ensembl.org/index.html, Flicek et al., 2014), and sequence alignments were performed using ClustalW software (http://www.ebi.ac.uk/Tools/msa/clustalw2/, Larkin et al., 2007). Nucleotide sequences resulting from sequencing reactions were aligned using the NCBI basic local alignment search tool (BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi, Altschul et al., 1990), which finds regions of similarity between sequences, and compares sequences to sequence databases and calculates the statistical significance of matches (Altschul et al., 1990; Johnson et al., 2008).

2.4.2 Protein motif prediction
Protein sequences were analysed for hypothetical secondary structure motifs using the programs Scansite (www.scansite3.mit.edu, Obenauer et al., 2003), and Prosite (http://prosite.expasy.org, Sigrist et al., 2010). These sites scan an input amino acid sequence and, using a database of consensus sequences, provide a prediction and probability score for the presence of secondary structure motifs found within that sequence.

2.4.3 Nucleotide motif analysis
Splice site analysis was performed using nucleotide sequences from DLG1 variants, and from sequencing data. Analysis was performed on more than one platform; the Maximum Entropy (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html, Yeo and Burge, 2004) program and the Human Splicing Finder (HSF, http://www.umd.be/HSF/, Desmet et al., 2009), which are designed to predict the effects of mutations on splicing
signals or to identify splicing motifs in any human sequence. Results were also compared to known splice site recognition motifs (Harris and Senapathy, 1990; Srebrow and Kornblihtt, 2006; Ghigna et al., 2008).

2.4.4 Densitometry analysis

Western blot films were scanned in and .tif images analysed using ImageJ software (Schneider et al., 2012). Once quantified, the area values for absorbance peaks were transferred to Microsoft Excel, where normalising and comparison calculations took place and graphs created.

Table 2.5: Primary antibodies used in western blotting and immunofluorescence

<table>
<thead>
<tr>
<th>Category</th>
<th>Protein detected</th>
<th>Size (kDa)</th>
<th>Supplier/No.</th>
<th>Animal</th>
<th>Dilution for WB</th>
<th>Dilution for IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDZ Proteins</td>
<td>DLG1</td>
<td>135</td>
<td>Santa Cruz 2D11</td>
<td>Mouse</td>
<td>1/1000</td>
<td>1/30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Santa Cruz H-60</td>
<td>Rabbit</td>
<td>1/500</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NT- Gift from L. Banks</td>
<td>Rabbit</td>
<td>1/500</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I2 specific, Roberts Lab</td>
<td>Rabbit</td>
<td>1/200</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>I3 specific</td>
<td>Rabbit</td>
<td>1/200</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Gift from R. Javier</td>
<td>Rabbit</td>
<td>1/200</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NHERF1</td>
<td>50</td>
<td>BD Biosciences</td>
<td>Mouse</td>
<td>1/500</td>
<td>1/50</td>
</tr>
<tr>
<td></td>
<td>PTPN13</td>
<td>250</td>
<td>Santa Cruz</td>
<td>Rabbit</td>
<td>1/300</td>
<td>1/50</td>
</tr>
<tr>
<td></td>
<td>SCRIB</td>
<td>210</td>
<td>Santa Cruz</td>
<td>Goat</td>
<td>1/2000</td>
<td>1/100</td>
</tr>
<tr>
<td></td>
<td>MAGI-1</td>
<td>120,130, 170</td>
<td>Sigma</td>
<td>Rabbit</td>
<td>1/500</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>MUPP-1</td>
<td>219</td>
<td>Gift from R. Javier</td>
<td>Rabbit</td>
<td>1/500</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>PARD3</td>
<td>151</td>
<td>Sigma</td>
<td>Rabbit</td>
<td>1/500</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>CASK</td>
<td>105</td>
<td>Cell Signaling Technology</td>
<td>Rabbit</td>
<td>1/400</td>
<td>1/50</td>
</tr>
<tr>
<td></td>
<td>ZO-2</td>
<td>134</td>
<td>Cell Signaling Technology</td>
<td>Rabbit</td>
<td>1/500</td>
<td>1/100</td>
</tr>
</tbody>
</table>
**Table 2.6: Secondary antibodies used in western blotting and immunofluorescence**

<table>
<thead>
<tr>
<th>Category</th>
<th>Protein detected</th>
<th>Isotype</th>
<th>Supplier</th>
<th>Animal</th>
<th>Dilution for WB</th>
<th>Dilution for IF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Secondary Antibodies</strong> (HRP conjugated)</td>
<td>Anti-mouse HRP</td>
<td>-</td>
<td>Sigma</td>
<td>Goat</td>
<td>1/3000</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Swine anti-rabbit HRP</td>
<td>-</td>
<td>Sigma</td>
<td>Swine</td>
<td>1/2000</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Rabbit anti-Goat HRP</td>
<td>-</td>
<td>Sigma</td>
<td>Rabbit</td>
<td>1/2000</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>AlexaFluor® 488 anti-mouse</td>
<td>IgG</td>
<td>Invitrogen</td>
<td>Goat</td>
<td>-</td>
<td>1/3000</td>
</tr>
<tr>
<td><strong>Secondary Antibodies</strong> (Fluorophore conjugated)</td>
<td>AlexaFluor® 594 anti-goat</td>
<td>IgG2a</td>
<td>Invitrogen</td>
<td>Rabbit</td>
<td>-</td>
<td>1/3000</td>
</tr>
<tr>
<td></td>
<td>AlexaFluor® 594 anti-rabbit</td>
<td>IgG</td>
<td>Invitrogen</td>
<td>Rabbit</td>
<td>-</td>
<td>1/3000</td>
</tr>
</tbody>
</table>

Antibodies were diluted in 2% w/v Milk (Marvel) PBS, * indicates requirement for 2% w/v BSA PBS.
2.5 List of suppliers

Abcam, 330 Cambridge Science Park, Cambridge, CB4 0FL, UK
Agilent Technologies, 5301 Stevens Creek Blvd, Santa Clara, CA 95051, USA
Alta Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
Amersham Pharmacia Biotech UK Ltd, Amersham Place, Little Chalfont, Bucks, HP7 9NA, UK
AMS Biotechnology Limited, 184 Park Drive, Milton Park, Abingdon, OX14 4SE, UK
Applied Biosystems, Lingley House, 120 Birchwood Boulevard, Birchwood, WA3 7QH, UK
BD Biosciences, Edmund Halley Road, Oxford Science Park, OX4 4DQ, UK
BioRad Laboratories Ltd., Maxted Road, Hemel Hempstead, Hertfordshire, HP2 7DX, UK
Cell Signaling Technology Inc., 3 Trask Lane, Danvers, MA 01923, USA
Clonetics Biowhittaker, 8830 Biggs Ford Rd, Walkersville, MD 21793-8415, USA
DAKO UK Ltd, Cambridge House, St Thomas Place, Ely, CB7 4EX, UK
EMD Millipore Headquarters, 290 Concord Road, Billerica, MA 01821, USA
Eppendorf UK Ltd, Endurance House, Chivers Way, Histon, Cambridge, CB24 9ZR, UK
GE Healthcare, Pollards Wood, Nightingales Lane, Chalfont St Giles, Bucks, HP8 4SP, UK
Genron Ltd., 12 Rawcliffe House, Howarth Road, Maidenhead, Berkshire, SL6 1AP
Gibco-Invitrogen, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK
Hycor Biomedical Ltd., Douglas House, Bush Loan, Penicuik, Midlothian, EH26 0PL, UK
Hybaid Ltd, Solaar House, 19 Mercers Row, Cambridge, CB5 8BZ, UK
Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, USA
Chapter 2: Materials and Methods

Life Technologies, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK
MP Biomedicals Inc., 1263 S. Chillicothe Road, Aurora, OH 44202, USA
National Diagnostics, Unit 4 Fleet Business Park, Itlings Lane, Hessle, Yorkshire HU13 9LX
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, BA22 8YG, UK
Perkin Elmer Ltd, Post Office Lane, Beaconsfield, Buckinghamshire, HP9 1QA, UK
Promega, Delta House, Chilworth Research Centre, Southampton, SO16 7NS, UK
Propath, Willow Court, Netherwood Road, Hereford, HR2 6JU, 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
Sigma-Aldrich Company Ltd., Fancy Road, Poole, Dorset, BH12 4QH, UK
Starlab (UK) Ltd., 5 Tanners Drive, Blakelands, Milton Keynes, MK14 5BU, UK
Stratagene, 11011 N. Torrey Pines Road, La Jolla, CA 92037, USA
Syngene (UK), Beacon House, Nuffield Road, Cambridge, CB4 1TF, UK
Takara Bio Europe (Clontech), 2 Avenue du President Kennedy, 78100 Saint-Germain-en-Laye, France
Thermo Fisher Scientifc Ltd, Bishop Meadow Road, Loughborough, LE11 5RG, UK
Vector Laboratories Ltd., 3 Accent Park, Bakewell Road, Orton Southgate, PE2 6XS, UK
VWR International Ltd., Hunter Boulevard, Magna Park, Lutterworth, LE17 4XN, UK
Xograph Healthcare, Xograph House, Ebley Road, Stonehouse, Gloucestershire, GL10 2LU, UK
Zeiss, 15 - 20 Woodfield Road, Welwyn Garden City, Hertfordshire, AL7 1JQ, UK
CHAPTER 3: THE INVESTIGATION OF PDZ PROTEIN EXPRESSION DURING THE HPV LIFE CYCLE

3.1 Introduction

The E6 PBM is a highly conserved feature of high-risk α-papillomaviruses types, which is a strong indication that a PDZ targeting function is important in the life cycle of this group of viruses (Figure 3.1). This concept was reinforced by the discovery of a PDZ binding motif in the E7 protein of fellow papillomavirus RhPV and the identification of a mode of PDZ targeting by HPV8 (Tomaic et al., 2009; Lazic et al., 2012). The human genome encodes over 300 PDZ domain-containing proteins and to date seventeen of these have been identified as targets of high-risk E6 proteins (Table 3.1), with several more suggested from high-throughput screening experiments (Belotti et al., 2013). The identified PDZ substrates of E6 comprise a number of proteins involved in a range of functions, including defining cell polarity, cell-cell attachment, and organising cell signalling pathways associated with these functions (Figure 1.10, Table 3.1).

The first proteins shown to interact with the E6 PBM were the Scribble complex components DLG1 and SCRIB (Kiyono et al., 1997; Nakagawa and Huibregtse, 2000). The second PDZ domain of DLG1 is required for HPV16 E6 binding, whereas HPV18 E6 is able to bind all three DLG1 PDZ domains independently (Gardiol et al., 1999; Kiyono et al., 1997). Moreover, of the four PDZ domains encoded by SCRIB, HPV binds to only one (PDZ3); indicating a specificity in HPV PDZ targeting, even within a protein (Nakagawa and Huibregtse, 2000). It was further discovered that the co-expression of HPV16 or HPV18 E6 alongside the PDZ proteins induced their degradation by the proteasome, and that E6*, a spliced isoform of E6 which lacks a PBM, is able to target endogenous DLG1, and to a lesser extent SCRIB, for degradation (Gardiol et al., 1999; Nakagawa and
Figure 3.1: C-terminal alignment of high- and low-risk α-HPV E6 proteins

To date, twelve HPV types have been defined by the WHO as high-risk with respect to causing cancer - these are designated as IARC Group 1- and a further ten types are recognised as ‘probably’ (Group 2A) and ‘possibly’ cancer-causing (Group 2B). HPV6 and 11 are included as examples of low risk virus. HPV8 is a β-papillomavirus associated with cutaneous SCC, and has been shown to effect the transcription of PDZ protein Syntenin-2.

The canonical PBM sequence ($X^3\alpha X^1\nu$) is shown above the highlighted region, which indicates the E6 PBM sequence.
Huibregtse, 2000; Pim et al., 2009). It has been suggested that SCRIB degradation by E6 is dependent upon the oncoprotein binding to the cellular ubiquitin ligase E6AP, and that this is not the case for DLG1; E6-mediated degradation of DLG1 still occurs in the absence of E6AP (Pim et al., 2000; Nakagawa and Huibregtse, 2000; Grm and Banks, 2004; Massimi et al., 2008). However, studies also demonstrated only a modest rescue of DLG1 and SCRIB in HPV-positive cancer cell lines where E6AP is silenced, suggesting that other cellular ligases may be recruited, the identity of which remain elusive (Kranjec and Banks, 2011; Massimi et al., 2008). Furthermore, E6AP was initially identified as important in the E6-mediated degradation of DLG4 in co-expression studies, but this was not observed in HPV-positive cancer cell lines, indicating that the different cellular conditions may contribute to the mechanism of E6-mediated PDZ protein degradation (Handa et al., 2007; Kranjec and Banks, 2011). Although co-expression studies have demonstrated E6-mediated DLG1 degradation, endogenous levels of DLG1 and SCRIB protein remain unaltered in primary human keratinocytes expressing HPV16 E6 and E7 (Choi et al., 2013; Lee et al., 2000; Lee and Laimins, 2004). In addition, no change in abundance or subcellular localisation of these PDZ proteins was observed in HPV31 genome containing HFKs (Lee and Laimins, 2004). Furthermore, in mice expressing HPV16 E6 under the K14 promoter, no change was observed in DLG1 or SCRIB protein levels or distribution compared to the equivalent skin of non-transgenic mouse (Simonson et al., 2005). These discrepancies between in vitro and in vivo experimental results highlight differences between the experimental models and the requirement to validate these targets in a relevant model of the HPV life cycle.

The Par complex includes PDZ proteins PARD3 and PAR6, and the complex associates with MUPP1 and the MAGUK family members MAGI-1, -2 and -3 (Figure 1.10). MUPP1 has been identified as a target of HPV18 E6, leading to degradation of the PDZ protein
In addition, MAGI-1, -2 and -3 have all been shown to be targeted for degradation by HPV16 and HPV18 in vitro (Glaunsinger et al., 2000; Thomas et al., 2002). Moreover, E6 mediated degradation was demonstrated in the HPV16 and 18 positive cancer cell lines Caski and HeLa, where the oncoprotein was shown to target membrane bound and nuclear pools of MAGI-1 (Kranjec and Banks, 2011).

Conversely, the recently shown interaction between HPV18 E6 and PARD3 does not lead to degradation of the PDZ protein (Facciuto et al., 2014). Instead, it was observed that HPV18 E6 redistributes the protein; a loss of PARD3 from the apical domain cell junctions and an increase in cytoplasmic and nuclear protein was observed when the viral protein was expressed (Facciuto et al., 2014). In addition, cellular protein TIP-1 was identified as interacting with HPV16 E6, but no degradation of the protein was seen during co-expression of the two proteins (Hampson et al., 2004). It was also observed that knockdown of endogenous E6 in Caski and HeLa did not affect TIP-1 protein levels (Kranjec and Banks, 2011), reiterating that not all known E6 PDZ interactions promote PDZ protein degradation.

The Crumbs complex protein PATJ associates with both HPV16 and HPV18 E6, via its fourth and fifth PDZ domains (Storrs and Silverstein, 2007). The resulting PATJ degradation does not require E6AP, and was the first example of a PDZ substrate degraded by E6* and thus revealed an alternative method of E6 association with PDZ proteins (Storrs and Silverstein, 2007).

The signalling proteins PTPN13, PTPN3 and NHERF1 have also been identified as targets of high-risk HPV E6 proteins. Jing et al. and Töpffer et al. showed that both HPV16 and HPV18 target the protein tyrosine phosphatase PTPN3, and that resulting degradation of the phosphatase is dependent upon both E6AP and proteasome. However, silencing of E6 in HeLa and Caski cervical cancer cell lines does not rescue PTPN3 protein levels and, as
PTPN13 decreases in these conditions, it was suggested that E6 and E7 stabilises PTPN13 in the cancer cell lines (Kranjec et al., 2014). In primary keratinocytes, expression of HPV16 E6 using retroviral vectors induced the PBM-dependent degradation of PTPN13, and expression of both HPV16 and HPV18 E6 led to the proteasomal degradation of NHERF1 (Spanos et al., 2008b; Accardi et al., 2011). In the latter experiments, HPV16 E7 activation of CDKs promoted NHERF1 phosphorylation and E6 preferentially targeted these phosphorylated forms for proteasome-mediated degradation in a PBM-dependent manner (Accardi et al., 2011). These findings suggest a cooperative mechanism of action on the PDZ protein NHERF1.

More recently, immunoprecipitation and yeast two hybrid screening experiments have identified interactions between the HPV16 E6 PBM and the known interacting proteins DLG1, SCRIIB, MAGI-1, MAGI-3, MUPP1 and PTPN3 as well as highlighting potential novel targets SNTA1, SNTB1, SNX27, CASK and LIN7C (Belotti et al., 2013). Similar methods were used to identify TIP-2, a known target of HTLV-1 Tax, as an interacting protein of HPV18 E6 (Favre-Bonvin et al., 2005; Rousset et al., 1998), and GOPC as able to interact with both HPV16 and HPV18 E6 PBM sequences (Jeong et al., 2007). Most recently, PDZRN3 has been identified as a target for proteasomal degradation by both HPV16 and HPV18 E6 (Thomas and Banks, 2014). Silencing of E6 in the HPV18 positive HeLa cells rescued PDZRN3 protein levels (Thomas and Banks, 2014).

High-risk HPV E6 proteins exhibit differing specificities for cellular PDZ proteins. (Kranjec and Banks, 2011). For example, both HPV16 and HPV18 E6 proteins are able to target DLG1 and SCRIIB, but exhibit different preferences in binding the PDZ proteins; HPV16 E6 preferentially binds SCRIIB over DLG1 and *vice versa* for HPV18 (Thomas et al., 2005, ). The PBM of these two oncoproteins only differs by a single amino acid and, notably, the specificity for DLG1 and SCRIIB can be switched within the context of the
whole E6 protein by switching the last residue (Thomas et al., 2005). In the same study, HPV18 exhibited a stronger affinity for MAGI-1 than HPV16 (Thomas et al., 2005). Similarly, GOPC was found to bind more strongly to HPV16 E6 than HPV18 and DLG4 has been shown to be a stronger target for HPV18 than for HPV16 (Handa et al., 2007; Kranjec and Banks, 2011). PDZ targeting by E6 is also specific to subcellular localisation; the cytoplasmic and nuclear pools of both DLG1 and MAGI-1 are preferentially targeted by HPV18 for degradation over protein at cell-cell contacts (Massimi et al., 2004; Kranjec and Banks, 2011; Narayan et al., 2009).

Interestingly, Lazic et al. revealed that expression of the skin-cancer associated HPV8 E6 significantly downregulated syntenin-2 transcription, with HPV16 E6 also producing a decrease in syntenin-2 mRNA, but to a lesser extent (Lazic et al., 2012). This study suggests alternative mechanisms by which HPVs modify PDZ protein expression, a field that requires further investigation.

While the E6 PDZ interaction is highly specific, it has been suggested that, at high enough concentrations, the E6 PBM may recognize any class I PDZ domain (Zhang et al., 2007). Thus, analyses in overexpression conditions are more analogous to the interactions during HPV-driven oncogenesis (Zhang et al., 2007). In terms of interactions with cellular targets, it is unclear which are the important PDZ targets in driving HPV propagation and, indeed, they may vary at different stages of the HPV life cycle.
Table 3.1: PDZ proteins associated with high-risk HPV E6 proteins

<table>
<thead>
<tr>
<th>PDZ protein (Gene Symbol)</th>
<th>Alternative Name(s)</th>
<th>Cellular Functions</th>
<th>Effect of Interaction with E6</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLG1</td>
<td>SAP97</td>
<td>Cell polarity, proliferation, migration and adhesion</td>
<td>Degradation, Mislocalisation</td>
<td>Kiyono et al., 1997; Gardiol et al., 1999</td>
</tr>
<tr>
<td>SCRIB</td>
<td>Scribble, Vartul, LAP4</td>
<td>Cell polarity and integrity, proliferation, apoptotic signalling</td>
<td>Degradation</td>
<td>Nakagawa and Huibregtse, 2000</td>
</tr>
<tr>
<td>INADL</td>
<td>PATJ</td>
<td>Tight junction formation, cell polarity</td>
<td>Degradation</td>
<td>Latorre et al., 2005; Storrs and Silverstein, 2007</td>
</tr>
<tr>
<td>SLC9A3R1</td>
<td>NHERF1, EBP50</td>
<td>GPCR regulation, cytoskeleton anchoring</td>
<td>Degradation</td>
<td>Accardi et al., 2011</td>
</tr>
<tr>
<td>PTPN13</td>
<td>FAP-1, PTPL-1, PTP-BAS, PTPBL</td>
<td>Fas-associated phosphatase, cell growth signalling, apoptotic signalling</td>
<td>Degradation</td>
<td>Spanos et al., 2008b</td>
</tr>
<tr>
<td>MPDZ</td>
<td>MUPP1</td>
<td>Receptor clustering, scaffold protein</td>
<td>Degradation</td>
<td>Lee et al., 2000</td>
</tr>
<tr>
<td>MAGI-1</td>
<td>AIP3, WWP3</td>
<td>Cell polarity, invasiveness, synapse formation, apoptosis</td>
<td>Degradation</td>
<td>Glaunsinger et al., 2000</td>
</tr>
<tr>
<td>MAGI-2</td>
<td>AIP1</td>
<td>Cell polarity, invasiveness, synapse formation, apoptosis</td>
<td>Degradation</td>
<td>Thomas et al., 2002</td>
</tr>
<tr>
<td>MAGI-3</td>
<td>-</td>
<td>Cell polarity, invasiveness, synapse formation, apoptosis</td>
<td>Degradation</td>
<td>Thomas et al., 2002</td>
</tr>
<tr>
<td>DLG4</td>
<td>PSD95, SAP90, TIP15</td>
<td>Receptor clustering</td>
<td>Degradation</td>
<td>Handa et al., 2007</td>
</tr>
<tr>
<td>PTPN3</td>
<td>PTPH1</td>
<td>Phosphatase, cell growth signalling, apoptotic signalling</td>
<td>Degradation</td>
<td>Jing et al., 2007; Töpffer et al., 2007</td>
</tr>
<tr>
<td>TAX1BP3</td>
<td>TIP-1</td>
<td>Wnt pathway inhibitor</td>
<td>-</td>
<td>Hampson et al., 2004</td>
</tr>
<tr>
<td>GIPC1</td>
<td>Synectin, TIP-2</td>
<td>Cell surface receptor expression and trafficking, TGFβ signalling</td>
<td>Degradation</td>
<td>Favre-Bonvin et al., 2005</td>
</tr>
<tr>
<td>GOPC</td>
<td>CAL</td>
<td>Intracellular trafficking</td>
<td>Degradation</td>
<td>Jeong et al., 2007</td>
</tr>
<tr>
<td>PARD3</td>
<td>PAR3</td>
<td>Cell polarity, assymetrical cell division</td>
<td>Mislocalisation</td>
<td>Facciuto et al., 2014</td>
</tr>
<tr>
<td>SDCBP2</td>
<td>Syntenin-2</td>
<td>Cell adhesion, growth factor signalling</td>
<td>Decreased transcription</td>
<td>Lazic et al., 2012</td>
</tr>
<tr>
<td>PDZRN3</td>
<td>LNX3</td>
<td>E3 ubiquitin ligase</td>
<td>Degradation</td>
<td>Thomas and Banks, 2014</td>
</tr>
</tbody>
</table>
3.2 Hypothesis and Aims

Studies to date have demonstrated that the high-risk HPV E6 PBM interacts with PDZ targets, yet the function of the E6 PBM in the context of the viral life cycle, where E6 is expressed at physiological levels, remains to be investigated. This study aims to investigate within the context of a physiologically relevant model of the HPV16 and HPV18 life cycle to identify relevant PDZ protein targets of the E6 PBM.

Thus, using a primary human foreskin keratinocyte (HFK) model of the HPV life cycle, the experiments described in this chapter aimed to:

i. Investigate the effect of HPV16 or HPV18 replication on transcription of E6 PDZ targets.

ii. Determine the abundance and distribution of E6 PDZ targets in HPV genome containing cells.
3.3 Results

3.3.1 Generation of cell-based models of the HPV16 and HPV18 life cycle

In order to study the life cycle of high-risk HPVs, primary human keratinocyte cell lines containing HPV16 and 18 genomes were generated in two donor backgrounds. Two separate HPV16 genome variants were utilised, 114/B (HPV16B) and 114/K (HPV16K). Variants are defined by having a less than 2% difference in the L1 ORF nucleotide sequence compared to the reference genome (Chen et al., 2011; Chan et al., 1995). Both HPV16B and HPV16K were originally derived from productive cervical lesions, and differ by a single amino acid; L1 asparagine 379 in HPV16K is replaced by a histidine residue in HPV16B (Kirnbauer et al., 1993). Early passage (p2) primary foreskin keratinocytes were co-transfected with the recircularised viral genomes and a neomycin resistance plasmid (see Section 2.1.4.4). Following drug selection and growth of cells, the presence of viral DNA was confirmed by Southern blot analysis (Section 2.2.2). Analysis of early passages of HPV16 and HPV18 genome containing keratinocytes confirmed the presence of HPV genomes in the two donors (Figure 3.2). These genomes migrate as supercoiled (SC) and open circle (OC) forms when digested with non-cutting restriction enzymes (HPV16, HindIII; HPV18, BglII), indicating that the HPV genome has established as episomes in these cells (Figure 3.2).

In a previous study by the Roberts group, HPV18 genomes were generated in which the E6 gene sequence was mutated to encode an E6 protein lacking the four amino acid C-terminal PDZ binding motif (∆PDZ). At the same time, HPV18 genomes were generated with a mutation which abrogates PKA recognition of a sequence (RRXS) that overlaps with the PBM (∆PKA) (Figure 3.3, Delury et al., 2013). Since loss of the E6 PBM function is known to affect the episomal genome maintenance upon increased population doubling
Figure 3.2: Generation of HPV16 and HPV18 genome containing primary keratinocyte cell lines

Southern blot analysis of total DNA isolated from two donor lines transfected with HPV16B, HPV16K (A) or HPV18 (B) genomes. DNA was digested with HindIII (HPV16B and HPV16K) or BglII (HPV18), which do not cut within the HPV genome. The addition of DpnI digests residual input DNA. Radiolabelled genomic HPV16 (A) or HPV18 (B) DNA was used as a probe. Episomal genomes migrate as open circle (OC) and supercoiled (SC) forms.
Figure 3.3: Confirmation of episomal genomes in HPV18 genomes with modified PDZ binding activity

(A) Southern blot of keratinocytes containing HPV18 wild-type, ΔPKA and ΔPDZ genomes. Total DNA was extracted and digested with EcoRI or BgII, alongside DpnI. HPV18 genomes are linearised by digestion with EcoRI, but remain uncut by BgII. Radiolabelled genomic HPV18 DNA was used as a probe. Episomal genomes migrate as open circle (OC) and supercoiled (SC) forms.

(B) Schematic of mutations within the HPV18 mutation taken from Delury et al. 2013. The ΔPDZ mutation encodes a deletion of the E6 PBM (ETQV), whereas the ΔPKA mutation abrogates the PKA recognition site which overlaps with the PBM (RRXST).
(Delury et al., 2013), the presence of viral episomes in early passages of the cell lines containing these genomes was also confirmed by Southern blot analysis (Figure 3.3). In the cell lines used, both mutant HPV genomes migrated as open circle and supercoiled forms, indicating replication of unintegrated episomes (Figure 3.3).

When grown in monolayer culture, these HPV16 and HPV18 genome containing cell lines, recapitulate the non-virus producing stages of the virus life cycle.

3.3.2 Transcription of E6 PDZ targets in keratinocytes containing high-risk α-papillomavirus genomes

The primary cell lines described above formed the basis for the following systematic analysis of PDZ protein expression. Although many studies have reported alterations in PDZ protein levels where high-risk HPV E6 is expressed, few have reported upon the transcriptional activity of the genes encoding this group of cellular proteins in the presence of HPV. It has been shown that the high-risk β-papillomavirus HPV8 E6 protein inhibits the transcription of the PDZ protein syntenin-2, leading to the downregulation of protein expression (Lazic et al., 2012). Although the same study identified a low level of transcriptional inhibition of syntenin-2 by HPV16, there has been no in-depth analysis of PDZ protein gene expression in cells replicating HPV genomes. Therefore, this study aimed to investigate the transcription of known E6 PDZ targets using the cell-based models of the HPV16 or HPV18 life cycle.

Prior to this study, a microarray gene analysis (Affymetrix) had been performed on monolayer cultures of HPV16 and HPV18 containing keratinocytes prepared in a different donor keratinocyte background to those used in this thesis (Leonard et al., unpublished data). Scrutiny of the array data for the expression of known E6 PDZ targets showed an
increase in the transcription of selected targets in the presence of both high-risk types when compared to transcription in untransfected HFKs of the same donor (Table 3.2).

Table 3.2: Affymetrix microarray analysis of E6 PDZ target expression in HPV16 and HPV18 genome containing keratinocytes

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene Symbol</th>
<th>Probe Set ID</th>
<th>Chromosomal Location</th>
<th>Fold change in expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>HPV16B</td>
<td>HPV16K</td>
</tr>
<tr>
<td>DLG1</td>
<td>DLG1</td>
<td>230229_at</td>
<td>3q29</td>
<td>11.67</td>
</tr>
<tr>
<td>SCRIB</td>
<td>SCRIB</td>
<td>212556_at</td>
<td>8q24.3</td>
<td>1.57</td>
</tr>
<tr>
<td>PATJ</td>
<td>INADL</td>
<td>214705_at</td>
<td>1p31.3</td>
<td>4.47</td>
</tr>
<tr>
<td>PTPN13</td>
<td>PTPN13</td>
<td>243792_x_at</td>
<td>4q21.3</td>
<td>4.22</td>
</tr>
<tr>
<td>MAGI-2</td>
<td>MAGI2</td>
<td>207702_s_at</td>
<td>7q21</td>
<td>3.66</td>
</tr>
<tr>
<td>MAGI-3</td>
<td>MAGI3</td>
<td>226770_at</td>
<td>1p12-p11.2</td>
<td>2.13</td>
</tr>
<tr>
<td>PTPN3</td>
<td>PTPN3</td>
<td>203997_at</td>
<td>9q31</td>
<td>1.24</td>
</tr>
<tr>
<td>TIP-1</td>
<td>TAX1BP3</td>
<td>209154_at</td>
<td>17p13</td>
<td>-1.75</td>
</tr>
<tr>
<td>TIP-2</td>
<td>GIPC1</td>
<td>207525_s_at</td>
<td>19p13.1</td>
<td>1.23</td>
</tr>
<tr>
<td>NHERF1</td>
<td>SLC9A3R1</td>
<td>201349_at</td>
<td>17q25.1</td>
<td>-1.23</td>
</tr>
<tr>
<td>p53</td>
<td>TP53</td>
<td>210609_s_at</td>
<td>17p13</td>
<td>-1.67</td>
</tr>
<tr>
<td>Syntenin-2</td>
<td>SDCBP2</td>
<td>233565_s_at</td>
<td>20p13</td>
<td>1.18</td>
</tr>
</tbody>
</table>

In the case of genes encoding the E6 PDZ targets MAGI-1, MUPP1, DLG4, PARD3 and GOPC, cellular transcript levels were below the detectable limit of the microarray probes, and these genes were consequently omitted from further analysis. Of the remaining E6 targets, transcripts were present. The level of SCRIB, MAGI-2, MAGI-3, MPDZ (MUPP1) PTPN3, TAX1BP3 (TIP-1) and GIPC1 (TIP-2) transcripts were not influenced (by more than ~2.5 fold) by HPV16 or HPV18 genome replication. In contrast, transcript levels of DLG1, PATJ and PTPN13 showed an upregulation of greater than 2.5-fold in the cells containing both high-risk genotypes (Table 3.2). The level of increase of both DLG1 and PATJ transcripts was much greater in the presence of HPV16 than HPV18. DLG1 transcription increased by 11.67-fold (HPV16) and 7.5-fold (HPV18) and PATJ
transcription was upregulated by 4.47-fold (HPV16) and 2.58-fold (HPV18). PTPN13 transcription increased equally by just over 4-fold between the two HPV genotypes (Table 3.2). The level of transcription of TP53, another protein target of high-risk α-HPV types was unaffected in keratinocytes containing HPV16 or 18 genomes (Table 3.2).

A similar profile of gene transcription was seen in a second cell line containing the HPV16K variant, indicating that transcriptional changes are not variant specific (Table 3.2). Next, qRTPCR analysis was used for verification of the microarray data. Because of the changes in transcription seen in the microarray analysis, the E6 PDZ targets DLG1, PATJ and PTPN13 were selected for further validation, alongside the known target SCRIB, whose transcription remained unaffected. As well as the analysis of DLG1, SCRIB, PATJ and PTPN13, a primer set was included for the PDZ protein NHERF1, gene symbol SLCR9A3R1, which had recently been identified as a novel target of HPV16 and HPV18 E6 proteins (Accardi et al., 2011). By including NHERF1 in qRTPCR validation, the expression of the PDZ protein could be investigated in a model system where whole HPV genomes were present. The microarray analysis indicated that high-risk HPV replication does not affect NHERF1 transcription; the fold change detected in the microarray was below 2.5 in both HPV16 and HPV18 containing keratinocytes compared to untransfected primary cells (Table 3.2). This data is in agreement with the findings of Accardi et al., who showed that there was no change in NHERF1 transcription following co-expression of E6 and E7 in primary keratinocytes.

Oligonucleotide primer sets for each PDZ target were designed to correspond to the region of the gene recognized by the probe used in the microarray analysis (see Section 2.2.1.2, Table 2.2). Each primer set was initially optimized by performing PCR at a range of different annealing temperatures, using DNase treated cDNA transcribed from the same RNA utilized in the microarray (Appendix II). Gene expression quantification by qRTPCR
was carried out on RNA isolated from HPV16 and HPV18 genome containing keratinocytes in the two donor backgrounds, each harvested at passage 3 (P3) and transcribed using oligodT primers (see Section 2.2.1.1). Changes in transcript levels were calculated relative to untransfected cells of the same donor. Figure 3.4 shows levels in individual donors, each repeated three times, and the data is summarized in Table 3.3 where the fold change in transcription is given as the mean change averaged across the two donors.

In this analysis, DLG1 transcripts were elevated in both HPV16 and HPV18 genome containing cells in both donor backgrounds (Figure 3.4 A). Where HPV16 was present, DLG1 expression was increased by ~18-fold (HPV16B) over untransfected, and in HPV18 genome containing keratinocytes, DLG1 transcription was increased by more than 5-fold (Figure 3.4 A). By contrast, transcription of the fellow Scribble complex protein SCRIB remained largely unaffected in cells containing the high-risk HPV genomes (Figure 3.4 B).

However, one donor exhibited a significant 3-fold increase of SCRIB transcripts in cells containing the HPV16B variant, but this was not consistently upheld in the other donor background, or in the presence of the HPV16K variant (Figure 3.4 B). Other than this, analysis of keratinocytes transfected with the two different variants HPV16 genomes (HPV16B and HPV16K) showed consistent effects upon the gene expression of the selected PDZ proteins; where there was an increase in cells containing the variant 16B (DLG1, PATJ, NHERF1 and PTPN13), the same trend occurred in the presence of HPV16K (Figure 3.4, summarized in Table 3.3). This analysis indicates that transcriptional changes are not variant specific.

Levels of PATJ transcripts were elevated in HPV16 and HPV18 genome containing cells, exhibiting an 11 to 14-fold and 6 to 9-fold increase in transcripts respectively (Figure 3.4
Transcription analysis of *DLG1* (A), *SCRIB* (B), *PATJ* (C), *NHERF1* (D) and *PTPN13* (E) genes in two donor keratinocyte backgrounds. mRNA was extracted from primary foreskin keratinocytes following establishment of stable lines containing HPV16 or HPV18 genomes, alongside untransfected keratinocytes at passage 3 (P3). This was transcribed to cDNA using oligodT primers and levels were measured by qRTPCR. ΔCt values were calculated relative to the housekeeping gene β2-microglobulin, which was normalised to transcript levels in untransfected keratinocytes (ΔΔCt) derived from the same donor background (indicated by dotted line). Error bars show ±SEM of three experimental repeats, statistical significance indicated by * (p<0.05) and ** (p<0.01).
Like the *DLG1* gene, transcription of *PATJ* was increased to a greater extent by HPV16 genome replication than by replication of HPV18 in monolayer keratinocytes. Transcription of the signalling protein *NHERF1* was increased by 3 to 4-fold in keratinocytes containing the HPV16 variants (Figure 3.4 D). The level of upregulation of NHERF1 transcription in HPV18 positive cells was only significant in one donor, where a more than 7-fold increase was observed (Figure 3.4 D). This analysis differs from published work, which demonstrated no change in *NHERF1* mRNA expression in HFKs expressing HPV16 E6 and/or E7 (Accardi et al., 2011).

Like *DLG1* and *PATJ*, *PTPN13* transcription was significantly upregulated in the presence of both HPV16 and HPV18; there was a ~4-fold increase in PTPN13 mRNA in HPV16 containing cells, and an ~3- (Donor 1) and ~3- (Donor 2) fold increase in cells replicating HPV18 (Figure 3.4 E).

Transcriptional analysis was extended to the E6 PBM targets *MAGI-1, -2* and -3, which encode junctional complex components. *MAGI-1* transcription was unchanged in HPV16 genome containing cells, although transcript levels were decreased in HPV18 genome containing cells to approximately half the levels of untransfected, in both donor keratinocytes (Figure 3.5 A). By contrast, a more than 3-fold increase in *MAGI-2* transcription was observed in the presence of both HPV types analysed and the transcription of *MAGI-3* remains unchanged by high-risk HPV replication (Figure 3.5 B and C).

Cumulatively, when averaged across donor keratinocyte backgrounds, there is a significant increase in the transcription of selected E6 targets in both HPV16 and HPV18 genome containing cells (Figure 3.6). The targets *DLG1, PATJ, MAGI-2, NHERF1* and *PTPN13* are increased transcriptionally in the presence of both high-risk α-HPV genotypes, although to different extents. The largest increase in transcription was seen for the *DLG1*
Figure 3.5: Quantification of MAGI-1, MAGI-2 and MAGI-3 transcription in HPV16 and HPV18 genome-containing keratinocytes

Transcription analysis of MAGI-1 (A), MAGI-2 (B) and MAGI-3 (C) genes. mRNA was extracted from primary foreskin keratinocytes following establishment of stable lines containing HPV16 or HPV18 genomes, alongside untransfected keratinocytes at passage 3 (P3). The mRNA was transcribed to cDNA and levels were measured by qRTPCR. ΔCt values were first were calculated relative to the housekeeping gene β2-microglobulin, and then to untransfected keratinocytes (ΔΔCt) derived from the same donor background (indicated by dotted line). Error bars show ±SEM of three experimental repeats, statistical significance indicated by * (p<0.05) and ** (p<0.01).
Figure 3.6: Cumulative qRTPCR analysis PDZ targets of E6 expression in HPV genome containing HFKs

Transcription of E6 PDZ targets in HPV16 and HPV18 genome containing keratinocytes as measured by qRTPCR, in two donor backgrounds. Transcription of DLG1, SCRIB, PATJ, NHERF1, PTPN13, MAGI-1, MAGI-2 and MAGI-3 were averaged following normalization to untransfected HFKs (dotted line). Error bars show ± SEM of three repeats in two donors. **indicates p<0.01, * indicates p<0.05.
gene in HPV16 positive cells (20-fold), with a similarly large (16-fold) increase in \textit{PATJ} transcription in the same cells. In both HPV16 and HPV18 positive keratinocytes, the transcription of these targets, and of NHERF1 and PTPN13, was increased in HPV18 containing keratinocytes. Notably, the transcription of both \textit{SCRIB} and \textit{MAGI-3} remained unaffected by both HPV genotypes (Figure 3.6). Interestingly, \textit{MAGI-1} transcription decreases in the presence of HPV18, whilst remaining unaffected by HPV16 (Figure 3.6).

To establish whether the increase in transcription of selected PDZ targets is maintained upon cell propagation, qRTPCR was performed on RNA extracted from the cells after a further five population doublings (P4). The profile of change in the level of transcripts in comparison to the untransfected cells remained the same as found in cells grown to P3 (Figure 3.7, summarized in Table 3.4). The expression of \textit{PATJ}, \textit{NHERF1} and \textit{PTPN13} remained elevated across the population doublings analysed, with \textit{PATJ} showing a trend towards a further increase in mRNA transcripts in the later passage.

**Table 3.3: Summary of PDZ target qRTPCR analysis**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fold change in expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HPV16B</td>
</tr>
<tr>
<td></td>
<td>p3</td>
</tr>
<tr>
<td>DLG1</td>
<td>17.56</td>
</tr>
<tr>
<td>SCRIB</td>
<td>2.39</td>
</tr>
<tr>
<td>PATJ</td>
<td>12.73</td>
</tr>
<tr>
<td>NHERF1</td>
<td>4.74</td>
</tr>
<tr>
<td>PTPN13</td>
<td>3.70</td>
</tr>
<tr>
<td>MAGI-1</td>
<td>-</td>
</tr>
<tr>
<td>MAGI-2</td>
<td>-</td>
</tr>
<tr>
<td>MAGI-3</td>
<td>-</td>
</tr>
</tbody>
</table>

Average of two donors. Fold change calculated relative to levels of transcription in HFKs derived from the same donor.

To investigate the contribution of the E6 PBM binding function with respect to PDZ target gene expression, qRTPCR gene expression analysis was carried out on monolayer
Figure 3.7: Quantification of E6-PDZ target transcription in HPV16 and HPV18 genome containing keratinocytes following cell propagation

Transcription analysis of DLG1 (A), SCRIB (B), PATJ (C), NHERF1 (D) and PTPN13 (E) genes. mRNA was extracted from primary foreskin keratinocytes following passage (p4) of stable lines containing HPV16 or HPV18 genomes in two donor backgrounds. mRNA levels were measured in HPV16 and HPV18 positive keratinocytes, alongside untransfected cells, by qRTPCR. ΔCt values were first were calculated relative to the housekeeping gene β2-microglobulin, and then to untransfected keratinocytes (ΔΔCt) derived from the same donor background (indicated by dotted line). Error bars show ±SEM of three experimental repeats, statistical significance indicated by * (p<0.05) and ** (p<0.01).
keratinocytes containing wild-type HPV18 genomes (WT), or the genomes containing a
PBM deletion (ΔPDZ) or the PKA unresponsive genome ΔPKA. In this analysis, gene
expression in two donors was carried out and calculated relative to transcript levels in wild-
type genome containing keratinocytes of the same donor background (Figure 3.8). There
were no significant changes in the transcription of DLG1, SCRIB, PATJ, NHERF1, 
PTPN13, MAGI-1 MAGI-2 or MAGI-3 in both mutant lines (Figure 3.8). These data
indicate that changes in PBM activity are not linked to the transcriptional change observed
for selected PDZ proteins.

In summary, the upregulation of selected PDZ targets in HPV16 and HPV18 observed in
microarray analysis was confirmed by qRTPCR (Table 3.4). The two donor cell lines
exhibited a consistent trend towards increased gene expression in the presence of both
high-risk HPV genomes, which was calculated as significant compared to HFK gene
expression. Notably, in the presence of both viral genotypes, transcripts of DLG1 were
substantially increased and SCRIB transcript levels remained largely unchanged.

Table 3.4: Comparison of microarray and qRTPCR analyses

<table>
<thead>
<tr>
<th>Gene</th>
<th>Average fold change in expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HPV16B</td>
</tr>
<tr>
<td></td>
<td>Microarray</td>
</tr>
<tr>
<td>DLG1</td>
<td>11.67</td>
</tr>
<tr>
<td>SCRIB</td>
<td>1.57</td>
</tr>
<tr>
<td>PATJ</td>
<td>4.47</td>
</tr>
<tr>
<td>NHERF1</td>
<td>1.54</td>
</tr>
<tr>
<td>PTPN13</td>
<td>4.22</td>
</tr>
<tr>
<td>MAGI-1</td>
<td>2.82</td>
</tr>
<tr>
<td>MAGI-2</td>
<td>3.66</td>
</tr>
<tr>
<td>MAGI-3</td>
<td>2.13</td>
</tr>
</tbody>
</table>

MAGI-1 transcripts below detectable level of microarray probe. qRTPCR shows average of 2 donors.
Figure 3.8: Quantification of E6-PDZ target transcription in HPV18 genome containing keratinocytes

Transcription of DLG1 (A), SCRIB (B), PATJ (C), NHERF1 (D) and PTPN13 (E) genes in primary keratinocytes replicating HPV18 genomes. mRNA extracted from HPV18 encoding wild-type (WT), ΔPDZ or ΔPKA E6 proteins. Expression levels in keratinocytes containing mutant (ΔPDZ and ΔPKA) genomes were calculated relative to mRNA levels in WT genome containing keratinocytes (dotted line) derived from the same donor background. Transcription analysis of DLG1 (A), SCRIB (B), PATJ (C), NHERF1 (D) and PTPN13 (E) genes. Graphs show results from three technical repeats.
3.3.3 Analysis of PDZ protein expression during the HPV life cycle

Given that the transcription of known E6 PDZ targets were altered in the presence of high-risk α-HPV genotypes, the next step was to investigate protein expression in the HPV life cycle model. In this study, seven identified targets of high-risk HPV E6 proteins (DLG1, SCRIB, PATJ, MAGI-1, MUPP1, NHERF1 and PTPN13) were investigated. This was performed alongside analysis of putative targets, suggested by recent screening experiments (Belotti et al., 2013, CASK, ZO-1 and ZO-2). These proteins are representative of proteins involved in the Scribble, Crumbs and Par epithelial junctional complexes, as well as in cellular signalling and scaffold organisation (Figure 1.10).

For the initial PDZ protein analysis, HPV16 or HPV18 genome containing keratinocytes were cultured until they reached 70% confluency (to ensure they remained in an undifferentiated state) and were then harvested and solubilised in a lysis buffer containing 8 M urea, and the protein extracts subjected to western blot analysis (see Section 2.1.4.4). PDZ protein expression was also examined in the HPV18 mutant genome containing keratinocytes, ΔPDZ and ΔPKA, and compared to wild-type HPV18 positive cells in order to assess whether any changes in protein expression were dependent upon the activity of the E6 PBM.

Since it has been shown that specific cellular pools of some PDZ proteins are preferentially degraded by E6 (Massimi et al., 2004; Simonson et al., 2005; Kühne et al., 2000; Narayan et al., 2009; Kranjec and Banks, 2011; Latorre et al., 2005), the subcellular localisation of PDZ proteins was examined in the life cycle model by fractionation of the monolayer-grown cells. This process used a series of lysis buffers to sequentially extract the following cellular compartments; the cytoplasm, the membrane, the nucleus and the cytoskeleton. The resulting protein extracts was subjected to western blot analysis, alongside analysis of host cell marker proteins in order to show enrichment of each of cellular fractions; growth
factor receptor-bound protein 2 (GRB2) in the cytoplasm, E-cadherin in the membrane, p53 in the nucleus and keratin 18 (K18) in the cytoskeleton (Figure 3.9). As can be seen in Figure 3.9 A, the fractionation process successfully enriched for the various cellular compartments of monolayer-grown keratinocytes. Each marker was found predominantly in the expected fraction; both p53 and K18 are exclusively detected in the nucleus and cytoskeleton respectively. GRB2 was enriched in the cytoplasmic and membrane fractions and E-cadherin was predominantly detected in the membrane fraction with some protein present in the nuclear and cytoskeletal fractions (Figure 3.9). The enrichment of these proteins was consistent across the cell lines; the presence of HPV16 or 18 genomes does not affect their localisation, and thus the distribution of these proteins is a suitable indicator for successful subcellular fractionation.

To investigate changes in PDZ protein expression during the productive phase of the virus life cycle, the HPV18 genome containing cells (WT and ΔPDZ) were stratified by growth in organotypic raft culture for 13 days (see Section 2.1). The differentiated structures were solubilized in urea lysis buffer and the protein extracts subjected to western blot analysis. Further analysis of sections of raft structures allowed the analysis of PDZ protein localisation during the HPV life cycle by immunofluorescent staining for PDZ targets. Thus, the following analysis examined known and indicated targets of high-risk HPV E6 proteins, proteins involved in the epithelial junction complexes, Scribble, Par and Crumbs, each of which contain known targets of E6, as well as PDZ-proteins involved in cellular signalling.
<table>
<thead>
<tr>
<th></th>
<th>HFK</th>
<th>HPV16</th>
<th>HPV18</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>GRB2</th>
<th>E cadherin</th>
<th>p53</th>
<th>K18</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV16</td>
<td></td>
<td></td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>HPV18</td>
<td></td>
<td></td>
<td>50%</td>
<td></td>
</tr>
</tbody>
</table>

**Proportion of protein**

- F1, Cytoplasm
- F2, Membrane
- F3, Nucleus
- F4, Cytoskeleton
Subcellular fractionation was carried out on monolayer-grown keratinocytes at 70% confluency. Protein from subcellular fractions was extracted by treatment with sequential lysis buffers. Successful enrichment was confirmed by western blot analysis; detection of the marker proteins GRB2, E-cadherin, p53 and K18 indicate enrichment of the cytoplasmic, membrane, nuclear and cytoskeletal fractions, respectively. The cell lines containing HPV16 and HPV18 (A) and HPV18 PBM mutants (B) were subject to fractionation in parallel. Western blots were quantified by densitometry using ImageJ software. Figures representative of three experimental repeats carried out on cells derived from two donor backgrounds.

Figure 3.9: Confirmation of successful enrichment of protein from subcellular fractions

Subcellular fractionation was carried out on monolayer-grown keratinocytes at 70% confluency. Protein from subcellular fractions was extracted by treatment with sequential lysis buffers. Successful enrichment was confirmed by western blot analysis; detection of the marker proteins GRB2, E-cadherin, p53 and K18 indicate enrichment of the cytoplasmic, membrane, nuclear and cytoskeletal fractions, respectively. The cell lines containing HPV16 and HPV18 (A) and HPV18 PBM mutants (B) were subject to fractionation in parallel. Western blots were quantified by densitometry using ImageJ software. Figures representative of three experimental repeats carried out on cells derived from two donor backgrounds.
3.3.3.1 The Scribble complex

3.3.3.1.1 DLG1

Despite the increase in transcription of the DLG1 gene, there was no change in DLG1 protein levels in HPV18 genome containing keratinocytes (Figure 3.10 A). Moreover, deletion of the E6 PDZ binding function (ΔPDZ) did not impact upon DLG1 protein levels (Figure 3.10 B). Furthermore, altering the E6 sequence to encode a constitutively active PBM, via removal of the PKA recognition site (ΔPKA), also had no affect upon the level of DLG1 protein (Figure 3.10 B). In the HPV18-positive cancer cell line HeLa, there was a 70% decrease in DLG1 protein levels, compared to normal keratinocytes (Figure 3.10 A), indicating differences between the life cycle model and a late stage carcinogenesis cell line. In HPV16 genome containing cells, there was a subtle trend towards a decrease in DLG1 protein; in two donors DLG1 protein was 0.7 and 0.8 times the level of untransfected cells (Figure 3.10 A). Taken together, no extensive change in protein levels was observed in these cells.

Subcellular fractionation of monolayer grown cells detected the majority of DLG1 in the nucleus of primary keratinocytes (Roberts et al., 2007), although the protein was also strongly detected in the membrane and weakly detected in the cytoskeleton and cytoplasm (Figure 3.11 A). Subcellular distribution of DLG1 protein remained unaffected by the presence of HPV16 or HPV18 genomes (Figure 3.11 A), and changes in the activity of the E6 PBM did not affect the distribution of DLG1 between subcellular pools (Figure 3.11 B). The distribution of DLG1 was verified using two antibodies specific for DLG1, a mouse monoclonal antibody (2D11) and a polyclonal antibody (NAG, Watson et al., 2002), which were raised to different N-terminal regions of DLG1.

During epithelial differentiation, the subcellular localisation of DLG1 changes (Roberts et al., 2007), and so analysis of DLG1 was applied to differentiating keratinocytes. To
Figure 3.10: DLG1 protein expression in keratinocytes replicating HPV genomes

DLG1 protein levels in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B). Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay and separated by SDS-PAGE. Membranes were probed with mouse monoclonal antibody raised against DLG1 (2D11), and β-actin as a loading control. Figures representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to β-actin and then to HFK (A) or WT (B) levels. Representative of two experimental repeats. Markers show molecular weight in kDa.
Figure 3.11: DLG1 protein distribution in keratinocytes
Distribution of DLG1 in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic, membrane, nuclear and cytoskeletal fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE. Membranes were probed with mouse monoclonal (2D11) or rabbit polyclonal (NAG) antibody raised against DLG1, Western blots were quantified by densitometry using ImageJ software. Figures representative of three experimental repeats carried out on cells derived from two donor backgrounds.
determine if HPV affects DLG1 during stratification, HPV18 genome containing keratinocytes were grown in organotypic raft culture and protein was extracted in urea lysis buffer (Section 2.1). Detection of DLG1 with the mouse monoclonal antibody 2D11 indicated a decrease in DLG1 protein where the E6 PDZ binding function was removed, compared to cells containing the wild-type HPV18 genome (Figure 3.12). However, detection with the NAG antibody indicated no change in DLG1 protein levels in differentiated keratinocytes irrespective of E6 PBM function (Figure 3.12).

Both the 2D11 and NAG antibodies were utilized in analysis of DLG1 localization in stratified keratinocytes by immunofluorescence (Figure 3.13). Using both antibodies, immunofluorescent staining of raft cultures showed DLG1 located predominantly at the cell junctions in the upper layers of HFKs, with a more cytoplasmic form detected in the basal layer (Figure 3.13). In HPV18 containing keratinocytes, the cell-cell location of DLG1 occurred earlier in the stratification of HPV18 genome containing cells, although this was more obvious when using 2D11 to detect DLG1 compared to NAG (Figure 3.13).

3.3.3.1.2 SCRIB

While levels of SCRIB protein in keratinocytes containing HPV16 were comparable to HFKs, in HPV18 positive cells SCRIB protein increased by approximately 40%, and this was consistent between both donors (Figure 3.14 A). In contrast, the polarity protein is depleted by 20% in the HPV18 DNA-positive cervical cancer cell line HeLa, compared to primary foreskin keratinocytes (Figure 3.14). Interestingly, the observed change on SCRIB expression is unrelated to the HPV18 PDZ binding function, as removal of the PBM (ΔPDZ) or altering the PKA recognition site (ΔPKA) does not restore the SCRIB protein levels to the level found in primary keratinocytes (Figure 3.14 B).

The localisation of SCRIB, in HPV16 and HPV18 positive keratinocytes as well as in normal cells, was predominantly cytoplasmic, nuclear and cytoskeletal (Figure 3.15 A).
Keratinocytes containing wild-type (WT) HPV18 genomes or HPV18 encoding an E6 without a PBM (ΔPDZ) were grown in organotypic raft culture. After 13 days, the raft was harvested, homogenised and lysed in order to extract protein. Protein was separated by SDS-PAGE and membranes were probed with mouse monoclonal (2D11) or rabbit polyclonal (NAG) antibody raised against DLG1, alongside GAPDH as a loading control. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to GAPDH and then to WT levels. Figures representative of two technical repeats.

Figure 3.12: DLG1 protein in differentiated keratinocytes containing HPV18 genomes
Figure 3.13: Immunofluorescence of DLG1 in differentiated keratinocytes

Immunofluorescent staining of keratinocytes grown in organotypic raft culture. Raft cultures were grown for 13 days, harvested and paraformaldehyde fixed. Individual sections were cut and stained for DLG1 using 2D11 (A, green) or NAG (B, red) antibodies. The nucleus stain DAPI (blue) was also used. Images were taken at 40x magnification.
Figure 3.14: SCRIB protein levels in HPV genome containing keratinocytes

SCRIB protein levels in keratinocytes containing WT HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay assay and separated by SDS-PAGE. Membranes were probed with goat polyclonal antibody raised against SCRIB, and β-actin as a loading control. Representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to β-actin and then to HFK (A) or WT (B) levels.
subtle change in protein levels was observed in HPV16 genome containing cells, indicating an increase in SCRIB at the cytoskeleton and a decrease in the nuclear fraction, but this was not consistently observed. This distribution is maintained in the presence of HPV18 genomes encoding E6 proteins with modified PBM binding activity (Figure 3.15 B). Following differentiation of the HPV18 genome containing cells in organotypic raft culture, SCRIB protein levels decrease upon abrogation of PDZ binding (Figure 3.16). This was trend seen in two donors; in one case there was an almost 50% decrease in protein in the presence of ΔPDZ genomes compared to wild-type HPV18 genomes (Figure 3.16). Like DLG1, immunofluorescent staining of differentiated keratinocytes indicated that SCRIB protein becomes increasingly at the cell periphery as keratinocytes migrate upwards, with basal cells exhibiting a predominantly nuclear pattern of SCRIB expression (Figure 3.17). Compared to untransfected keratinocytes, stratified cells containing HPV18 genomes exhibited a more obvious junctional location of SCRIB protein, which occurred in the mid layers of the epithelia (Figure 3.17).

In summary, DLG1 and SCRIB exhibited no marked alterations in protein level or subcellular distribution in the presence of HPV16 or HPV18 genomes in monolayer grown primary human keratinocytes. An increase in SCRIB protein in the presence of HPV18 was not restored by removal of the PBM, although in differentiating cells, the levels of both proteins decreased in cells containing HPV18 genomes expressing E6 without a PBM. In differentiated culture, the location of SCRIB and DLG1 at the periphery of cells in the upper layers is altered in the presence of HPV18 and expression is increased in the mid-layers.

3.3.3.1.3 CASK

CASK is known to interact with DLG1 in epithelia (Nix et al., 2000), and has been observed by the Roberts lab to co-immunoprecipitate with DLG1 in human keratinocytes
Figure 3.15: Distribution of SCRIB in monolayer keratinocytes

Distribution of SCRIB in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic, membrane, nuclear and cytoskeletal fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE. Membranes were probed with goat polyclonal antibody raised against SCRIB, Western blots were quantified by densitometry using ImageJ software. Representative of two donors, each repeated three times.
Figure 3.16: SCRIB protein in differentiated keratinocytes containing HPV18 genomes

Keratinocytes containing wild-type (WT) HPV18 genomes or HPV18 encoding an E6 without a PBM (ΔPDZ) were grown in organotypic raft culture. After 13 days, the raft was harvested, homogenised and lysed in order to extract protein. Protein was separated by SDS-PAGE and membranes were probed with goat polyclonal antibody raised against SCRIB, alongside GAPDH as a loading control. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to GAPDH and then to WT levels. Representative of two technical repeats.
Figure 3.17: SCRIB distribution in differentiated keratinocytes
Immunofluorescent staining of keratinocytes grown in organotypic raft culture. Raft cultures were grown for 13 days, harvested and paraformaldehyde fixed. Individual sections were cut and stained for SCRIB (green) and the nucleus stained with DAPI (blue). Images were taken at 40x magnification.
(S Roberts, personal communication). Furthermore, the protein sequence exhibits 25% sequence homology to DLG1 and is suggested to associate with the HPV16 PBM in yeast-two hybrid screening experiments (Belotti et al., 2013). To assess whether CASK protein expression was affected by HPV16 or HPV18, lysates isolated from monolayer cultures were western blotted for CASK protein. There was no alteration in the CASK protein levels in keratinocytes containing HPV18 (Figure 3.18 A). Furthermore, removal or alteration of the HPV18 PDZ binding ability did not produce any effect upon CASK protein (Figure 3.18 A). Next, the distribution of CASK was analysed (Figure 3.18 B and C). CASK was detected predominantly in the membrane and the nucleus in untransfected keratinocytes, with a small amount present in the cytoplasmic and cytoskeleton fractions of the cell (Figure 3.18 B and C). This subcellular distribution was not perturbed by the presence of HPV16 or HPV18, nor by alteration in the activity of the E6 PBM (Figure 3.18).

3.3.3.2 Crumbs complex

3.3.3.2.1 PATJ

In this analysis, the levels of PATJ protein remained constant in HPV16 and HPV18 positive keratinocytes, compared to untransfected cells. Protein levels remained the same in the cancer cell line HeLa as primary keratinocytes (Figure 3.19 A). Furthermore, perturbing the PDZ binding function of E6 had no effect on protein levels compared to wild-type HPV18 genome containing cells (Figure 3.19 B). (Storrs and Silverstein, 2007) demonstrated that the truncated spliced form of E6, E6*, was able to induce PATJ degradation, indicating that the effects of E6 on this oncoprotein is PBM independent (Storrs and Silverstein, 2007). However, this analysis did not find any overwhelming change in PATJ in monolayer grown keratinocytes (Figure 3.20). The expression of adenovirus E4-
Figure 3.18: CASK protein level and distribution in monolayer grown keratinocytes

(A) Total protein was extracted from monolayer grown keratinocytes containing WT, ΔPDZ or ΔPKA HPV18 genomes, alongside HFKs and seperated by SDS PAGE. Representative of two repeats.

(B) Distribution of CASK in keratinocytes containing HPV16 and 18 genomes

(C) Distribution of CASK protein in keratinocytes containing HPV18 WT and mutant genomes

Cells for B and C were grown in monolayer culture. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic, membrane, nuclear and cytoskeletal fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE.

In all cases, membranes were probed with rabbit polyclonal antibody raised against CASK. Western blots were quantified by densitometry using ImageJ software. Markers show molecular weight in kDa. Representative of two donors, each repeated three times.
Figure 3.19: PATJ protein levels in HPV genome containing keratinocytes

PATJ protein levels in keratinocytes containing WT HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay and separated by SDS-PAGE. Membranes were probed with rabbit polyclonal antibody raised against PATJ, and β-actin as a loading control. Representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Markers show molecular weight in kDa.
Figure 3.20: Distribution of PATJ in monolayer keratinocytes

Distribution of PATJ in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Representative of two donors and three repeats. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic, membrane, nuclear and cytoskeletal fractions fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE. Membranes were probed with rabbit polyclonal antibody raised against PATJ. Markers show molecular weight in kDa.
CHAPTER 3: THE INVESTIGATION OF PDZ PROTEIN EXPRESSION DURING THE HPV LIFE CYCLE

ORF1 has been shown to cause mislocalisation of PATJ (Latorre et al., 2005), but it is not known whether expression of E6 similarly affects the localisation of PATJ protein. Thus, the distribution of PATJ was investigated in monolayer keratinocytes containing HPV genomes (Figure 3.20). During this analysis multiple bands were detected in all the fractions isolated (Figure 3.20). Five isoforms of PATJ have been identified, with molecular weights ranging between 125 to 196 kDa (Roh et al., 2002; Philipp and Flockerzi, 1997; Soejima et al., 2001). It may be that the bands detected by this western blot analysis correspond to PATJ variants; however, the results were inconsistent and therefore inconclusive.

3.3.3.3 The Par complex and other tight junction components

3.3.3.3.1 PARD3

Western blot detection of PARD3 revealed an increase in the level of protein, in both monolayer and differentiated keratinocytes replicating HPV18 (Figure 3.21 A and D). In organotypic raft culture, a decrease in PARD3 was observed in the cells containing the HPV18 ΔPDZ mutant genomes, suggesting a role for the E6 PBM in stability of the PDZ protein, although this was not the case in cells grown monolayer culture (Figure 3.21 A and D). In HeLa cells, the level of PARD3 was similar to HFK (Figure 3.21 A). A recent study has shown that HPV18 E6 causes a PBM-dependent change in PARD3 localisation, but not proteins levels (Facciuto et al., 2014). However, analysis of subcellular distribution showed no alteration in distribution of PARD3 in the viral genome containing cells; PARD3 remained predominantly nuclear and cytoskeletal, with some cytoplasmic forms (Figure 3.21 B and C). PARD3 exists as several isoforms, ranging from 100 to 200 kDa (Jan et al., 2013). Interestingly, PARD3 was detected as a doublet in protein from both HFK and HeLa cells, whereas only a single form of PARD3 was detected in HPV genome.
Figure 3.21: PARD3 protein levels and distribution in primary keratinocytes

(A) Total protein in monolayer grown keratinocytes, representative of two experimental repeats.

(B) Subcellular distribution in untransfected, HPV16- and HPV18-containing keratinocytes grown in monolayer.

(C) Subcellular distribution in HPV18 mutant genome containing cells grown in monolayer.

F1, cytoplasmic; F2, membrane; F3, nuclear; F4, cytoskeletal. Blots representative of three repeats in two donors.

(D) Total protein levels in keratinocytes grown in organotypic raft culture, representative of two technical repeats. Densitometry shows quantification of protein in WT and ΔPDZ genome containing cells.

All blots probed with rabbit polyclonal antibody raised against PARD3. Markers show molecular weight in kDa.
replicating cells (Figure 3.21 A). Similarly, PARD3 was detected as multiple bands in fractionation analysis, which made quantification difficult but indicated that more PARD3 variants were present in the nucleus and cytoskeleton in all keratinocytes investigated.

3.3.3.2.2 MAGI-1

Although MAGI-1 was difficult to detect by western blotting, there is a slight increase in protein levels in undifferentiated HPV18 genome containing keratinocytes, but this was not reinforced in analysis of differentiated cells (Figure 3.22). Upon subcellular differentiation fractionation, the protein exhibited a membrane, nuclear and cytoskeleton distribution in all the primary cells analysed and this trend was largely unaffected by replication of the two high-risk genomes (Figure 3.22).

3.3.3.3.3 MUPP1

MUPP1 protein levels were increased in HPV18 genome containing cells, and this increase was independent of the activity of the E6 PBM (Figure 3.23 A). MUPP1 protein was detected mainly in nuclear and cytoskeletal fractions, with minor expression in the cytoplasm and membrane (Figure 3.23 B and C). The distribution was largely unaffected by the viral genomes (Figure 3.23 D).

3.3.3.4 ZO-1

Although no change in ZO-1 protein levels was observed in the presence of HPV16 genomes, levels of ZO-1 protein increased in HPV18 genome containing cells (Figure 3.24 A). This increase may be related to E6 PBM activity; in monolayer cells, a subtle decrease was observed when ZO-1 protein was quantified in cells containing genomes where the E6 PDZ binding function is deleted (ΔPDZ or ΔPKA, Figure 3.24 B). A subtle effect of the E6 PBM on ZO-1 levels is also observed in raft cultures of E6ΔPDZ cells (Figure 3.25). However, the observed increase in protein does reflect a change in the distribution of ZO-1 (Figure 3.26). Fractionation indicated that ZO-1 resides mainly in the nucleus and
Figure 3.22: MAGI-1 protein levels and distribution in primary keratinocytes
(A) Total MAGI-1 protein in monolayer grown keratinocytes. Representative of two repeats.
(B) Subcellular distribution of MAGI-1 in untransfected, HPV16- and HPV18-containing keratinocytes.
(C) Subcellular distribution of MAGI-1 in HPV18 mutant genome containing cells. F1, cytoplasmic; F2, membrane; F3, Nuclear; F4, cytoskeletal.
B and C representative of three repeats in two donors. Cells grown to 70% confluency before protein extraction, and blots probed with rabbit polyclonal antibody raised against MAGI-1. Markers show molecular weight in kDa.
Figure 3.23: MUPP1 protein levels and distribution in primary keratinocytes

(A) MUPP-1 protein in monolayer grown keratinocytes. Representative of two repeats.

(B) MUPP1 subcellular distribution in untransfected, HPV16- and HPV18-containing keratinocytes

(C) Subcellular distribution of MUPP1 in HPV18 mutant genome containing cells. F1, cytoplasmic; F2, membrane; F3, Nuclear; F4, cytoskeletal. Markers show molecular weight in kDa.

B and C representative of three repeats in two donors. Cells grown to 70% confluency before protein extraction, and blots probed with rabbit polyclonal antibody raised against MAGI-1. Markers show molecular weight in kDa.
Figure 3.24: ZO-1 protein levels in HPV genome containing keratinocytes

Total ZO-1 protein in keratinocytes containing HPV16 and 18 genomes (A) and in keratinocytes containing mutant HPV18 genomes (B). Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay and separated by SDS-PAGE. Membranes were probed with mouse monoclonal antibody raised against ZO-1, and β-actin as a loading control. Representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to β-actin and then to HFK (A) or WT (B) levels.
Figure 3.25: ZO-1 protein in differentiated keratinocyte culture
Total ZO-1 protein in keratinocytes grown in organotypic raft culture. Cells containing wild-type HPV18 (WT) genomes or genomes encoding E6 where the PBM is deleted (ΔPDZ) were stratified for 13 days and total protein extracted in urea buffer. Separated protein was probed for ZO-1, alongside GAPDH as a loading control. Representative of two technical repeats. Western blots was quantified using ImageJ and fold change calculated relative to to WT levels.
associated with the cytoskeleton of keratinocytes, and that this situation is unperturbed by mutation of the HPV18 E6 PBM (Figure 3.26 B). A number of ZO-1 variants become apparent in subcellular fractionation analysis, and the distribution of these may be altered by the presence of HPV genomes, for example, the lower band detected in the cytoskeletal fraction is absent where HPV18 genomes are able to constitutively bind to PDZ proteins (Figure 3.26 B). However, as with PATJ and MAGI-1, this makes a conclusion difficult.

The distribution of ZO-1 was assessed in sections of differentiated keratinocytes, where expression was most apparent at the periphery of cells residing in the upper layers (Figure 3.27). However, in the presence of HPV18 genomes, expression of ZO-1 was more evenly spread across the layers of cells and the protein concentrates at regions of the membrane (Figure 3.27). Alteration of PBM activity restored the HFK phenotype in part; protein was located at the cell membrane in the upper layers of differentiated cells in the presence of ΔPDZ genomes, suggesting that the effect of HPV18 upon ZO-1 is PBM-dependent (Figure 3.27).

3.3.3.3.5 ZO-2

Similarly to ZO-1, ZO-2 protein levels are slightly increased in keratinocytes in the presence of HPV18 genomes compared to untransfected cells, but the change is not PBM dependent (Figure 3.28). In raft cultures, the level of ZO-2 in ΔPDZ genome containing cells was less than cells containing wild-type genomes (Figure 3.28 B). More apparent is the decrease in ZO-2 in the presence of HPV16 genomes (Figure 3.28 A). In cultured epithelial cells, HPV16 E6 expression has been shown to cause ZO-2 relocalisation to the nucleus and cytoplasm from cell borders (Hernandez-Monge et al., 2013). In this study of ZO-2 cellular distribution, in viral genome containing cells, ZO-2 was detected consistently in the membrane, nuclear and cytoskeletal fraction (Figure 3.29). An increase in ZO-2 protein was observed in the nuclear fraction of cells containing PBM-deletion mutants, and in
Figure 3.26: Distribution of ZO-1 in monolayer keratinocytes

Distribution of ZO-1 in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Representative of two donors and three repeats. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic (F1), membrane (F2), nuclear (F3) and cytoskeletal (F4) fractions fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE. Membranes were probed with mouse monoclonal antibody raised against ZO-1. Representative of three repeats of two donors.
Figure 3.27: ZO-1 distribution in differentiated keratinocytes
Immunofluorescent staining of keratinocytes grown in organotypic raft culture. Raft cultures were grown for 13 days, harvested and paraformaldehyde fixed. Individual sections were cut and stained for ZO-1 (green) and the nucleus stained with DAPI (blue). Images were taken at 40x magnification.
Figure 3.28: ZO-2 protein levels in HPV genome containing keratinocytes
ZO-2 protein levels in keratinocytes containing WT HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay and separated by SDS-PAGE. Membranes were probed with rabbit polyclonal antibody raised against ZO-2, and β-actin as a loading control. Representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to β-actin and then to HFK (A) or WT (B) levels.
Figure 3.29: Distribution of ZO-2 protein in monolayer keratinocytes
Distribution of ZO-2 in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Representative of two donors and three repeats. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic, membrane, nuclear and cytoskeletal fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE. Membranes were probed with rabbit polyclonal antibody raised against ZO-2. Western blots were quantified by densitometry using ImageJ software.
HFKs, indicating a PBM-dependent effect upon ZO-2 distribution by HPV. However, analysis of differentiated cells revealed no change in ZO-2 protein levels in the presence of HPV18 (Figure 3.30).

3.3.3.4 Signalling proteins

3.3.3.4.1 NHERF1

The expression of NHERF1 protein is unaffected by the presence of HPV16 or HPV18 episomes. Similarly, levels of the cellular PDZ protein are not affected by the activity of the E6 PBM (Figure 3.31). Investigation of the subcellular distribution of NHERF1 revealed that the pattern of NHERF1 expression is unaltered by viral genome replication (Figure 3.32). During Western blot analysis, several forms of NHERF1 migrating at different molecular weights were identified (Figure 3.32). These are likely to correspond to hypo and hyper-phosphorylated forms of the protein (He et al., 2001). NHERF1 protein was predominantly detected in the cytoplasm, but was also present at the membrane and nuclear fractions (Figure 3.32). A potential hypophosphorylated form of NHERF1 is predominantly detected in the cytoskeletal compartment, detected as a faster migrating polypeptide (Figure 3.32 He et al., 2001). Although high-risk viruses had no effect to report upon NHERF1 protein levels in monolayer culture or differentiated cells compared to untransfected cells, levels of NHERF1 protein increased (by 2-fold) in cells upon loss of E6 PBM activity (Figure 3.33). Further investigation by immunofluorescent staining of differentiated culture sections revealed that, whereas low levels of NHERF1 protein was detected throughout differentiated HFKs, in the presence if HPV18 NHERF1 levels remained low in the basal layer but expression was increased in the mid and upper layers (Figure 3.34). A similar phenotype was observed in the absence of the E6 PBM, although the protein is more obviously not at the cell membranes in these sections (Figure 3.34).
**Figure 3.30: ZO-2 protein in differentiated keratinocytes**

Total ZO-2 protein in keratinocytes grown in organotypic raft culture. Cells containing wild-type HPV18 (WT) genomes or genomes encoding E6 where the PBM is deleted (ΔPDZ) were stratified for 13 days and total protein extracted in urea buffer. Separated protein was probed for ZO-2, alongside GAPDH as a loading control. Representative of two technical repeats. Western blots was quantified using ImageJ and fold change calculated relative to to WT levels.
Figure 3.31: NHERF1 protein levels in HPV genome containing keratinocytes

NHERF1 protein levels in keratinocytes containing WT HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay and separated by SDS-PAGE. Membranes were probed with mouse polyclonal antibody raised against NHERF1, and β-actin as a loading control. Representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to β-actin and then to HFK (A) or WT (B) levels.
Figure 3.32: Distribution of NHERF1 protein in monolayer keratinocytes

Distribution of NHERF1 in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Representative of two donors and three repeats. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic (F1), membrane (F2), nuclear (F3) and cytoskeletal (F4) fractions were extracted by treatment with sequential lysis buffers, an equal amounts separated by SDS-PAGE. Membranes were probed with mouse monoclonal antibody raised against NHERF1. Western blots were quantified by densitometry using ImageJ software, quantifying the top band only.
Figure 3.33: NHERF1 protein in differentiated keratinocytes
Total NHERF1 protein in keratinocytes grown in organotypic raft culture. Cells containing wild-type HPV18 (WT) genomes or genomes encoding E6 where the PBM is deleted (ΔPDZ) were stratified for 13 days and total protein extracted in urea buffer. NHERF1 was detected alongside GAPDH as a loading control. Representative of two technical repeats. Western blots was quantified using ImageJ and fold change calculated relative to to WT levels.
Figure 3.34: NHERF1 distribution in differentiated keratinocytes
Immunofluorescent staining of keratinocytes grown in organotypic raft culture. Raft cultures were grown for 13 days, harvested and paraformaldehyde fixed. Individual sections were cut and stained for NHERF1 (green) and the nucleus stained with DAPI (blue). Images were taken at 40x magnification.
3.3.4.2 PTPN13

There was no change in PTPN13 protein levels in undifferentiated keratinocytes containing HPV16 or HPV18 genomes, compared to levels in the untransfected HFKs (Figure 3.35 A). Furthermore, the protein is not perturbed in HeLa compared to normal primary keratinocytes (Figure 3.35 A). Abrogation of the HPV18 E6 PBM did not affect cellular levels of PTPN13 protein; equivalent levels of PTPN13 were observed in cells containing WT, ΔPDZ and ΔPKA genomes (Figure 3.35 B). Although an increase in PTPN13 in the membrane fraction can be observed in HPV16 genome containing keratinocytes, the number of potential variants detected made the analysis in genome containing cells unclear (Figure 3.36). In monolayer, alteration of the E6 PBM led to a higher proportion of PTPN13 detected in the nucleus and, in differentiated keratinocytes removal of the PBM caused an increase in PTPN13 protein compared to wild-type HPV18 genome containing cells (Figure 3.37).

3.3.4 Proteasomal inhibition of primary keratinocytes

In the context of cells containing the complete HPV genome, where E6 is expressed from its natural promoter, no decrease in PDZ protein levels has been observed, compared to control keratinocytes. A number of PDZ proteins, including DLG1, SCRIB, PTPN13 and NHERF1, have been identified as targets of high-risk α-HPVs for proteasome-mediated degradation, in a PBM dependent manner. Therefore, to assess the expression of these proteins in response to inhibition of the proteasome in the viral genome containing cells, cells were treated with MG132 for 4 or 6 hours, alongside cells treated with the vehicle DMSO. Protein was harvested in urea lysis buffer and subjected to western blot analysis as described in Section 2.2.4.
Figure 3.35: PTPN13 protein levels in HPV genome containing keratinocytes
PTPN13 protein levels in keratinocytes containing WT HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Protein was extracted from cells grown in monolayer culture to 70% confluency, quantified by Bradford assay and separated by SDS-PAGE. Membranes were probed with rabbit polyclonal antibody raised against PTPN13, and β-actin as a loading control. Representative of two repeats. Western blots were quantified by densitometry using ImageJ software. Graphs show relative levels, first normalised to β-actin and then to HFK (A) or WT (B) levels.
Figure 3.36: Distribution of PTPN13 in monolayer keratinocytes

Distribution of PTPN13 in keratinocytes containing HPV16 and 18 genomes (A) and mutant HPV18 genomes (B) grown in monolayer culture. Representative of two donors and three repeats. Subcellular fractionation was carried out on keratinocytes at 70% confluency. The cytoplasmic, membrane, nuclear and cytoskeletal fractions were extracted by treatment with sequential lysis buffers, equal amounts separated by SDS-PAGE. Membranes were probed with rabbit polyclonal antibody raised against PTPN13 and Western blots were quantified by densitometry using ImageJ software.
**Figure 3.37: PTPN13 protein in differentiated keratinocytes**

Total PTPN13 protein in keratinocytes grown in organotypic raft culture. Cells containing wild-type HPV18 (WT) genomes or genomes encoding E6 where the PBM is deleted (ΔPDZ) were stratified for 13 days and total protein extracted in urea buffer. Separated protein was probed for PTPN13, alongside GAPDH to indicate equal loading. Representative of two technical repeats. Western blots were quantified using ImageJ and fold change calculated relative to WT levels.
Since p53 is a known target of proteasomal degradation, both in the presence and absence of the E6 oncoprotein, the lysates were immunoblotted for p53. As expected, an increase in p53 was observed in both HFK and HPV positive cells upon addition of MG132, confirming that proteasomal inhibition was successful (Figure 3.38 A). A decrease in p53 protein was observed when comparing HPV positive cells to untransfected control cells (Figure 3.38 B).

In HFKs under conditions of proteasomal inhibition, DLG1 protein accumulated following 4 and 6 hours of inhibition, confirming that the protein is targeted for proteasomal degradation (Mantovani et al., 2001, Figure 3.39). By contrast, there was a marked decrease in the level of DLG1 protein in viral genome containing untransfected keratinocytes treated with proteasomal inhibitor, compared to equivalent cells treated with vehicle (Figure 3.39 A and B). As a junction complex protein, it was expected that DLG1 levels would be increased in cells harvested at a higher density. However, a similar pattern was observed in cells harvested at a higher confluency (90% vs. 70%); levels of DLG1 decreased following proteasomal inhibition (Figure 3.39 C and D). Intriguingly, a similar profile was observed for SCRIB, albeit the effect was less marked and was not specific to keratinocytes containing viral genomes (Figure 3.40). This response to proteasomal inhibition was not consistent between all PDZ proteins tested, because both NHERF1 and PTPN13 levels increased upon proteasomal inhibition, irrespective of the presence of HPV genomes (Figure 3.41). Moreover, the change in PTPN13 levels in response to MG132 in cells containing HPV16 genomes is not striking (Figure 3.41).

The interaction between high-risk HPV E6 with NHERF1, SCRIB and DLG1 has been shown to result in proteasomal degradation of the PDZ proteins (Gardiol et al., 1999; Massimi et al., 2008; Accardi et al., 2011; Nakagawa and Huibregtse, 2000). However, treatment of primary keratinocytes with proteasomal inhibitor revealed that the presence of
Figure 3.38: PDZ protein levels in keratinocytes under proteasomal inhibition

(A) Monolayer-grown keratinocytes containing HPV16 or HPV18 genomes were treated alongside HFK with vehicle (DMSO, -) or the proteasomal inhibitor MG132 (+), for 4 or 6 hours before protein was extracted using a urea buffer. Equal amounts of protein were separated by SDS-PAGE and p53 levels detected alongside β-actin loading control.

(B) Comparison of untreated controls run on the same gels. Equal amounts of protein were separated by SDS-PAGE and p53 levels detected alongside β-actin loading control. Representative of three repeats in cells derived from one donor background.
<table>
<thead>
<tr>
<th></th>
<th>HFK</th>
<th>HPV16</th>
<th>HPV18</th>
<th>HPV16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>MG132:</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
Figure 3.39: DLG1 protein levels in keratinocytes under proteasomal inhibition

DLG1 protein analysis of keratinocytes treated with vehicle (DMSO, -) or proteasomal inhibitor MG132 (+) for four or six hours before harvesting. Keratinocytes containing HPV16 (A and C) or HPV18 (B and D) genomes were grown alongside untransfected cells to 70% (A and B) or 90% (C and D) confluency before treatment. DLG1 was detected using a mouse monoclonal antibody (2D11). Representative of three repeats in cells derived from one donor background.
Figure 3.40: SCRIB protein in keratinocytes under conditions of proteasomal inhibition
SCRIB protein analysis of keratinocytes treated with vehicle (DMSO, -) or proteasomal inhibitor MG132 (+) for four or six hours before harvesting. Keratinocytes containing HPV16 or HPV18 genomes were grown alongside untransfected cells to 70% confluency before treatment. SCRIB was detected using a goat polyclonal antibody. Representative of three repeats in cells derived from one donor background.
Figure 3.41: NHERF1 and PTPN13 protein levels in keratinocytes under proteasomal inhibition

NHERF1 and PTPN13 Protein analysis of keratinocytes treated with vehicle (DMSO, -) or proteasomal inhibitor MG132 (+) for four or six hours before harvesting. HFKs and HPV16 or 18 genome containing keratinocytes were grown to 70% confluency and harvested in urea buffer. Detection of β-actin was included as a loading control. Representative of three repeats in cells derived form one donor background.
HPV genomes does not affect the proteasome-mediated degradation of E6 PDZ targets SCRIB, PTPN13 and NHERF1. Moreover, proteasome inhibition of primary keratinocytes containing high-risk HPVs had an unexpected effect upon DLG1.

Table 3.5: Summary of PDZ protein analysis

<table>
<thead>
<tr>
<th>PDZ target</th>
<th>Compared to untransfected keratinocytes</th>
<th>Compared to HPV18 wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HPV16</td>
<td>HPV18</td>
</tr>
<tr>
<td>DLG1</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>SCRIB</td>
<td>NC</td>
<td>+</td>
</tr>
<tr>
<td>CASK</td>
<td>NT</td>
<td>NT</td>
</tr>
<tr>
<td>MAGI-1</td>
<td>NT</td>
<td>(+)</td>
</tr>
<tr>
<td>MUPP1</td>
<td>NT</td>
<td>+</td>
</tr>
<tr>
<td>PARD3</td>
<td>NT</td>
<td>+</td>
</tr>
<tr>
<td>PATJ</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>ZO-1</td>
<td>NC</td>
<td>+</td>
</tr>
<tr>
<td>ZO-2</td>
<td>-</td>
<td>NC</td>
</tr>
<tr>
<td>NHERF1</td>
<td>NC</td>
<td>NC</td>
</tr>
<tr>
<td>PTPN13</td>
<td>NC</td>
<td>NC</td>
</tr>
</tbody>
</table>

+, increase; -, decrease; NC, No change; NT, not tested. Bracketed values indicate subtle changes.
3.4 Discussion

3.4.1 Transcriptional upregulation in the presence of HPV genomes

Primary keratinocyte cell lines containing HPV16 or HPV18 were established and subjected to analysis of PDZ protein expression. The viral genomes in these lines exist as episomes and in monolayer the cell model is comparable to basal infected cells. Both microarray and qRTPCR analyses revealed that the polarity proteins DLG1 (Scribble complex), MAGI-2 (tight junctions) and PATJ (Crumbs complex), as well as the signalling proteins NHERF1 and PTPN13, were upregulated at a transcriptional level in HPV16 and HPV18 genome containing cells. This was in contrast to the other PDZ proteins investigated, SCRIB (Scribble complex) and MAGI-3 (tight-junctions), whose transcription was unaffected. The values of fold change were consistently higher than microarray values when quantified by qRTPCR, but the trend in transcriptional change was comparable. Differences in values between qRTPCR and microarray quantification may reflect the different donor backgrounds utilized, as well as being an indication of the sensitivity of each technique (Git et al., 2010). In the case of NHERF1 expression, only qRTPCR showed a significant increase in transcription. The difference in NHERF1 gene expression compared to work by Accardi et al. may be attributable to the presence of whole HPV16 genomes in this study; the presence of other viral factors may, directly or via cellular processes, affect the transcription of E6 PDZ targets. Similarly, in identifying PTPN13 as a HPV16 E6 target, Spanos et al. demonstrated that the presence of HPV16 E6 does not affect PTPN13 mRNA levels in primary keratinocytes, whereas this thesis presents an upregulation of the PDZ protein in high-risk HPV genome containing keratinocytes. Together with our data, this suggests that a protein other than E6 acts to increase the transcription of PTPN13 during viral genome replication.
A previous study by Lazic et al. demonstrated that HPV8 and HPV16 E6 protein expression induced a decrease in syntenin-2 transcript levels, but the transcription of other known PDZ targets, including DLG1, SCRIB, MAGI-1, MAGI-3, DLG4 and MUPP1 were unaffected, indicating that the effect of E6 upon PDZ protein transcription is specific. Such specificity was also observed in this study; the targets SCRIB and MAGI-3 remain transcriptionally unaffected, whilst MAGI-1 transcription was decreased by HPV18 only.

By using a PBM deletion mutant, this study ruled out any effects of the motif upon transcriptional upregulation; altering the PDZ binding activity of HPV18 revealed upregulation consistent results with WT HPV18 containing cells. Thus, the study in this thesis is the first in-depth study to indicate that high-risk α-HPVs are able to modulate PDZ protein expression, independently of the E6 PBM.

A potential viral candidate in causing the observed PDZ gene upregulation is the HPV protein E2. The N-terminal domain of E2 can regulate transcription from adjacent promoters and HPV16 has been shown to increase or decrease transcription of cellular genes depending upon the level of E2 protein (Bouvard et al., 1994a). A study by Gauson et al. identified more than 40 genes, that were downregulated and more than 30 genes, including the PDZ domain containing protein SNTB1, that were upregulated in cells expressing HPV16 E2. Thus, it is conceivable that E2 acts to upregulate selected PDZ targets of E6 in primary keratinocytes. Recently, E6 has been shown to be influential in inducing upregulation of transcription of the cytidine deaminase APOBEC3B (Vieira et al., 2014). Both the transfection of high-risk HPV genomes and the expression of E6 alone induced increased transcription and further experiments showed that this was a characteristic of high-risk E6 proteins (Vieira et al., 2014). The authors propose that derepression of gene transcription occurs via the inactivation of p53 by E6. Similarly, HPV16
E6 has been shown to transcriptionally upregulate components of the NFκB pathway in cervical cells and induce upregulation of involucrin gene transcription (Gyongyosi et al., 2012; Nees et al., 2001). Work by Vliet-Gregg et al. indicated that E6 can induce the transcriptional upregulation of NOTCH1 when expressed alongside the transcription factor NFX1-123. Thus, there is a precedent for E6 alone causing a change in gene transcription.

3.4.2 Protein distribution in HPV genome containing cells

3.4.2.1 Comparison to HeLa

This study shows a downregulation of the junctional proteins DLG1, SCRIB, ZO-1 and PARD3 in the HPV18 positive cancer cell line HeLa, compared to keratinocytes, which corresponds with data from the Banks lab (Kühne et al., 2000). HeLa express relatively high levels of HPV18 E6, and are thus more able to deregulate PDZ proteins than keratinocytes. However, the changes cannot solely be attributed to HPV18 E6, there are likely many consequences of continuous cell culture, including the accumulation of somatic mutations. The levels of NHERF1 protein in HeLa were equivalent to the levels in keratinocytes, which differs from the analysis by Accardi et al., who described an increase in NHERF1 protein in HeLa. This may be due to the different primary keratinocytes used as a comparison, but also highlights the variations seen in extensively cultured cells (Castillo et al., 2014).

3.4.2.2 Tight junction proteins

The presence of high-risk HPV genomes in monolayer grown keratinocytes had no or an undetectable effect upon the levels or distribution of PTPN13 or PATJ proteins, or fellow PDZ domain containing proteins CASK PARD3, MAGI-1 and MUPP1. This latter group (CASK, PARD3, MAGI-1 and MUPP1), along with PATJ, registered no change in protein
levels in differentiated keratinocytes replicating high-risk HPV genomes, despite studies which have shown protein interaction (CASK, MAGI-1, MUPP1), redistribution (PARD3) and E6AP independent degradation (PATJ. Facciuto et al., 2014; Glaunsinger et al., 2000; Belotti et al., 2013; Lee et al., 2000; Storrs and Silverstein, 2007). Additionally, it has been shown that E6 can preferentially target PDZ proteins at particular cellular locations, and also that E6 is able induce cellular redistribution of PDZ proteins, something that this study does not replicate (Narayan et al., 2009; Massimi et al., 2004; Facciuto et al., 2014). In monolayer, the distribution of known and putative E6 targets remained the same in HPV16 and HPV18 positive and negative keratinocytes. These disparities reflect the differences between high throughput and overexpression assays and our primary cell model of HPV replication. It is likely that the E6 interactions that promote cancerous growth differ from those that are formed during the HPV life cycle, where E6 expression is tightly regulated.

This study also encountered difficulties in detecting MAGI-1, PATJ and MUPP1 proteins, preventing the drawing of concrete conclusions. Optimisation of these reagents would enable further studies to investigate the abundance and distribution of these proteins in a cell-based model of HPV replication.

HPV16 expression leads to an upregulation of ZO-2 in mouse epithelium and a change in distribution in MDCK cells (Hernandez-Monge et al., 2013). Here, we show a decrease in ZO-2 protein in the presence of HPV16 genomes, whilst there was no effect in HPV18 genome containing cells. The levels of fellow tight-junction protein ZO-1 were increased in HPV18 genome containing cells. Initially, it may seem unexpected for the interaction between E6 and ZO-1 results in an increase in the PDZ protein, but it may be that PBM binding has a stabilising effect, much like that observed by Nicolaides et al. The stability of E6 is aided by interaction with SCRIB, and it is possible that such a stabilising effect can be
reciprocated (Nicolaides et al., 2011). It is interesting to note that HPV16 downregulated ZO-2 whilst HPV18 upregulated ZO-1, indicating that targeting of one of these tight-junction components alone is sufficient for viral purposes. Furthermore the interaction between ZO-1 and HPV18 E6 is likely to be dependent upon PBM activity, as removal of the PBM in monolayer relieved some of the increase in protein level seen in WT genome containing cells and analysis or raft cultures revealed that the change in protein distribution from the periphery of HFKs to more focused regions of the membrane of HPV18 genome containing cells was restored in part by deletion of the PBM.

3.4.2.3 The Scribble complex

The two components of the Scribble complex analysed are the main identified targets of high-risk α-papillomaviruses. These proteins are targeted for proteasomal degradation by HPV16 and HPV18 E6 oncoproteins, with HPV16 preferentially targeting SCRIB, and HPV18 targeting DLG1 (Thomas et al., 2005). It is interesting that, despite the different affinities of HPV16 and HPV18 E6 for DLG1 protein, both viruses genomes induced an increased DLG1 transcript levels, whereas neither had an effect upon SCRIB gene expression. Surprisingly, this upregulation did not correspond to an observed increase in DLG1 protein in the HPV16 and HPV18 genome containing keratinocytes, and this was validated using two detection antibodies. A more in depth analysis of DLG1 expression in HPV genome containing cells is provided in Chapter 4.

Numerous over and co-expression studies have demonstrated the degradation of DLG1 mediated by high-risk HPV E6 proteins (Lee and Laimins, 2004; Thomas et al., 2005; Gardiol et al., 1999). However, it has also been shown that endogenous DLG1 is not as readily targeted (Delury et al., 2013; Choi et al., 2013). Moreover, in primary tonsil keratinocytes, there was no change in DLG1 protein when overexpressing HPV16 E6
(Spanos et al., 2008b). The maintenance of the DLG1 protein distribution in HPV positive and negative primary cells observed in this study agrees with data from other high-risk HPV types; HPV31 genome containing keratinocytes demonstrate the same localisation of cellular DLG1 than untransfected keratinocytes (Lee and Laimins, 2004; Choi et al., 2013). Although both HPV types are known to target SCRIB, it is a preferential target of the HPV16 E6 PBM (Thomas and Banks, 2005). Thus, it was surprising that there was no change in SCRIB protein levels in HPV16 genome containing cells, but an increase in protein occurred in the presence of HPV18 genomes. Furthermore, transcription of this PDZ protein remained unchanged in HPV16 and HPV18 replicating cells. It was also unexpected that membrane protein SCRIB was not detected in membrane fraction in monolayer cells. However, the observation was reinforced by the subsequent detection of SCRIB in the nucleus of basal cells and at the cytoskeleton in the upper layers of differentiated keratinocytes. Simonson et al. suggest that any differences observed in the distribution of DLG1 and SCRIB are subtle, and observe no change in differentiated culture. Work form the same group suggest that HPV31 does not induce dramatic changes in the subcellular distribution of DLG1 when transfected into primary keratinocytes; only a subtle reduction in cytoplasmic DLG1 protein was detected (Simonson et al., 2005; Lee and Laimins, 2004).

In differentiated cells replicating HPV18 genomes, the levels of Scribble complex associated proteins DLG1, SCRIB and CASK, and the tight junction proteins ZO-1 and ZO-2 were influenced by the activity of the E6 PBM; a lower level of these proteins was detected in protein extracted from raft cultures containing genomes encoding the E6 PBM than those encoding a PBM deletion. However, examination of total protein from stratified cells does not indicate which pools of PDZ protein are affected. Furthermore, analysis of protein from differentiated culture can be problematic as the raft growth is inconsistent;
keratinocytes replicating HPV18 genomes that do not encode an E6 PBM grow more slowly than cells containing wild-type HPV18 genomes (Delury et al., 2013).

Observing sections of differentiated rafts to identify protein location during stratification of epithelial cells containing HPV genomes enables observation of protein localisation during the full viral life cycle. Perhaps the susceptibility of PDZ proteins to E6-mediated degradation depends upon the location of the PDZ target in question; PDZ proteins could be more resistant to E6-induced degradation when they are already incorporated into cytoskeletal structures around membrane regions of cell–cell contacts. In this study no HPV-related changes in DLG1 and SCRIB protein levels were observed in monolayer grown cells or in the basal layer of stratified cells. However, observation of organotypic raft culture showed an increased detection of these Scribble complex proteins in the mid layer of the cultured epithelia, indicating that interactions between HPV E6 and PDZ proteins occur specifically during stages of the HPV life cycle. Furthermore, although no changes in NHERF1 abundance or distribution were observed in monolayer culture, differentiated cells containing HPV18 exhibited an increase in the protein levels. However, as this does not correlate with any observed increase in protein extracted from stratified cells, this requires more thorough investigation, perhaps utilizing confocal techniques to more accurately detect changes in protein abundance and distribution.

Inhibition of the proteasome revealed that the observed increases in PDZ protein gene transcription, and any downstream increase in translation, was not countered by proteasomal degradation; no increase in DLG1 protein was observed in HPV genome containing cells treated with proteasome inhibitor. This is also true of PTPN13 and PATJ; the increase in transcript levels in HPV genome containing keratinocytes did not correspond to a change in protein levels in these cells. Counter-intuitively, in HPV16 and HPV18 genome containing cells, a decrease in DLG1 protein levels was seen in conditions
of proteasome inhibition. It is unclear what occurs to induce this, but the decrease in protein may be due to non-specific effects of MG132 on these primary cells.

Taken together, no extensive change in protein levels was observed in these cells. This was surprising as DLG1 is a well characterized target of high-risk HPV E6, for proteasome degradation (Gardiol et al., 1999; Thomas et al., 2005). However, the observed effects of HPV genomes on PDZ targets appears to concentrate on proteins involved in tight junction formation and signalling (Figure 3.42). Protein levels of the tight-junction proteins ZO-1 and ZO-2 were altered by HPV genomes and mRNA of the tight-junction associated MAGI-1 was decreased. Both NHERF1 and PTPN13, which were upregulated transcriptionally in the presence of HPV genomes, have roles in tight-junction organisation and signalling; NHERF1 has been shown to function in tight junction organization, specifically in ZO-1 and occluding localization and PTPN13 is suggested to function in junction formation via ephrin (Castellani et al., 2012; Sotelo et al., 2014; Vermeer et al., 2012). Disruption of epithelial polarity and tight junctions has been linked to activation of growth factor signalling and the release of pro-stimulatory transcription factors (Vermeer et al., 2003; Balda and Matter, 2003; Betanzos et al., 2004). Indeed, ZO-2 associates with the AP-1 components fos and jun at tight junctions, and our observed decrease of the PDZ protein in HPV16 positive cells may reveal a mechanism to release the pro-growth transcription factor (Betanzos et al., 2004). In addition, ZO-2 blocks cell cycle progression by acting upon cyclin D1; the PDZ protein downregulates protein synthesis, increases proteasomal degradation and downregulates cyclin D1 transcription via interaction with the TF Myc, (Tapia et al., 2009; Huerta et al., 2007; González-Mariscal et al., 2009). This culminates in the suppression of epithelial cell proliferation, which is counter to HPV
Figure 3.42: Summary of PDZ protein expression in the presence of high-risk α-HPV genomes
The effects of HPV genomes upon keratinocytes concentrates on the tight junctions.
Red outline and arrows indicate a change in transcription and a black outline indicates a change in protein level in the presence of HPV. The number on the line indicates HPV type.
proliferation. Thus it is conceivable that, by decreasing ZO-2, HPV16 E6 promotes progression of the cell cycle.

The TJ protein ZO-1 interacts with the transcription factor ZONAB (ZO-1 associated nucleic acid binding protein), which is involved in the regulation of epithelial cell proliferation (Balda and Matter, 2000; 2003). Perhaps an increase in this protein by HPV18 functions to promote the PDZ:TF interaction with the aim of promoting cell proliferation. Furthermore, it is conceivable that this release of TFs and propagation of growth signalling by deregulated E6, promotes an uncontrollable growth that transforms infected cells and promotes cancer development.

By interrupting the junctions, HPV propagation may be favoured by promoting cell growth signalling. However, too much disruption of junction complexes may lead to dissociation of the epithelia and a halt in differentiation, which would also interrupt viral replication. Thus targeting of junctional proteins is selective, and does not result in the complete degradation of proteins involved.

3.4.3 Future directions

Protein analysis of keratinocytes does not validate existing data that shows E6-mediated degradation of PDZ targets. This is not to say that interactions between E6 and PDZ proteins do not occur during the HPV life cycle, PARD3 is an example of an E6 target which is not degraded as a result of the interaction (Facciuto et al., 2014). However the effect of E6 upon PDZ proteins during the HPV life cycle is subtle, and it is still unclear which is the major target at a protein level. The PBM of high-risk HPV E6 proteins contributes to episomal maintenance, and it has been suggested that the stabilizing effect of PBM PDZ interactions contributes to this (Nicolaides et al., 2011). E6 mediated degradation may be an artifact of E6 overexpression and the deregulation of viral
oncoproteins that occurs in cancer, as such it would be of interest to investigate these proteins in cells where the HPV genome has integrated, which occurs in primary keratinocytes following serial culture.

This study identifies a number of targets of the HPV E6 PBM, which are upregulated at a transcriptional level in the presence of viral genomes. This does not, however, correspond to an increase in protein levels. However, this study was limited to investigating one method of protein degradation; it may be that E6 targets these proteins for degradation via the lysosomal pathway. It is possible that a viral factor decreases the mRNA stability, before translation occurs. Alternatively, the upregulated transcripts may encode less stable proteins or not translate as efficiently that other variants. As such, the pattern of DLG1 expression investigated in more depth in Chapter 4.

Further investigation may also identify viral factors and cellular pathways involved in the observed changes in PDZ target transcription. This could begin with expression analysis of cells containing single viral gene products (such as E2), before dissection of mechanisms of action.
CHAPTER 4: IDENTIFICATION AND CHARACTERISATION OF A NOVEL ISOFORM OF DLG1

4.1 Introduction

4.1.1 DLG1 protein structure

Like other members of the MAGUK family of proteins, DLG1 is composed of modular interaction domains that enable a variety of protein-protein interactions. The protein consists of 3 PDZ domains, an Src Homology 3 (SH3) domain and a guanylate kinase (GK) domain (Figure 4.1). Two isoforms of DLG1 have been identified, one containing an N-terminal L27 motif, referred to as DLG1β, and another without this L27 motif termed DLG1α. These isoforms correlate to Drosophila DLGS97 and DLGA, respectively (Figure 4.2, Albornoz et al., 2008; Mendoza et al., 2003). Further diversity in DLG1 is introduced by the inclusion of short amino acid sequences designated insertions (I1-I5), produced by splicing alternative exons in the DLG1 gene. This splicing occurs in two main regions of DLG1, one at the N-terminus and another in the C-terminal HOOK region that lies between the SH3 and GK domains (Figure 4.2). Five alternatively spliced insertions have been defined to date, I1A and I1B at the N-terminal region, and I2, I3, I4 and I5 within the HOOK domain, each of which results in the inclusion of a short (10-30) amino acid sequence (Table 4.1). These insertions influence DLG1 localisation as well as potentially altering binding partner preferences of DLG1 (McLaughlin et al., 2002).

The N-terminal L27 motif of DLG1 facilitates dimerization via interaction with the L27 domains of other proteins as well as other DLG1 molecules (Lin et al., 2013). The region remains largely unstructured until association with other L27 motifs, when the motifs together form an ordered helical structure and become a compact four-helix bundle.
Figure 4.1: Domain structure of DLG1 isoforms

Adapted from Roberts et al. 2012. Two human DLG1 isoforms (accession numbers U13896 and U13897) were originally identified, encoding proteins of 926 and 904 amino acids respectively. Both the DLG1α and DLG1β isoforms encode 3 PDZ domains, an SH3 and a C-terminal GUK domain. In addition, the DLG1β isoform encodes an N-terminal L27 region, which is involved in DLG1 homomultimerization and mediates interactions with other DLG1 binding proteins. Regions of alternative mRNA splicing at the N-terminus (I1A/I1B and in the HOOK region (I2, I3, I4 and I5) are shown. Splicing further diversifies the isoforms of DLG1 expressed. The five unstructured regions (U1-U5) are indicated. U5 is also known as the HOOK region.
Figure 4.2: Regions of alternative splicing of DLG1
Adapted from Roberts et al. 2012. Both the DLG1α and DLG1β isoforms encode 3 PDZ domains, an SH3 and a C-terminal GUK domain. In addition, the DLG1β isoform encodes an N-terminal L27 region, sequence indicated in red. Regions of alternative mRNA splicing at the N-terminus (I1A/I1B and in the HOOK region (I2, I3, I4 and I5) are shown.
(Doerks et al., 2000). In non-neuronal cells, the L27 domain of DLG1 facilitates interactions with the first L27 domain of fellow MAGUK protein, calcium/calmodulin-dependent serine protein kinase (CASK) and stabilises a more extended form of DLG1 (Jeyifous et al., 2009; Lee et al., 2002; Waites et al., 2009). Moreover, this interaction with CASK is important for localization of the β isoform of DLG1 to the baso-lateral membrane of epithelia (Lozovatsky et al., 2009). Similarly, this region in the β isoform has been shown to mediate membrane localization via association with DLG2 (Hanada et al., 2003). In neuronal cells, the presence of the α and β isoforms affects the trafficking of specific receptors at the synapse; the β isoform associates with AMPA receptors primarily at post-synaptic regions and the α form at perisynaptic regions (Waites et al., 2009). In the fly, both the DlgA and DlgS97 isoforms are present at larval neuromuscular junctions (Mendoza et al., 2003). Recently, DlgA and DlgS97 were shown to be differentially expressed in the fly; whereas DlgA is required for adult viability, specific loss of DlgS97 leads to perturbation of circadian activity and courtship (Mendoza-Topaz et al., 2008). It has also been observed that DlgA is predominant in epithelia whereas the DlgS97 is mainly expressed in the Drosophila neuronal system (Mendoza et al., 2003; Lee et al., 2002).

DLG1 encodes three PDZ domains, which influence DLG1 localisation and interact with a wide range of cellular proteins. For example, DLG1 binds to the tumour suppressor adenomatous polyposis coli (APC) protein, via PDZ1 and PDZ2, forming a complex which inhibits cell cycle progression in response to epithelial cell-cell contact, indicating a role for DLG1 in controlling progression of cells from G0/G1 to S phase (Ishidate et al., 2000; Zhang et al., 2011; Makino et al., 1997). The PDZ domains of DLG1 are known to interact with different transmembrane proteins; for example PDZ1 interacts with the AMPA receptor subunit GluR1, PDZ2 with the NR2B subunit of NMDA receptor and PDZ3
with neuroligin (Kornau et al., 1995; Leonard et al., 1998; Cai et al., 2002; Wang et al., 2005). Binding of PDZ domains 1 and 2 of DLG1 to the Rho-GEF NET1 (neuroepithelioma transforming gene 1) leads to the relocalisation of both proteins to the nucleus (Garcia-Mata et al., 2007; Carr et al., 2009).

Table 4.1: Insertions in DLG1 variants

<table>
<thead>
<tr>
<th>Insertion</th>
<th>Exon No.</th>
<th>Amino acid sequence</th>
<th>Present in variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>6</td>
<td>PTEAVLPSPPTVPVPIPVLPVPAENTVILTIPQ</td>
<td>1, 2, 6 and 7</td>
</tr>
<tr>
<td>1B</td>
<td>7</td>
<td>ANPPPVVLNTDSLETPTY</td>
<td>1, 2, 3, 4, 6, 7, 8 and 9</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>EIPDDMSKGLK</td>
<td>1, 3, 5, 6, 7, 8 and 9</td>
</tr>
<tr>
<td>3</td>
<td>20 A/B</td>
<td>(Q) SFNPKRKKNLFSRRFFPYKNKQSEQETSDADQ</td>
<td>2 and 4</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>LILITDEYGCSKG</td>
<td>7 and 9</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>HVTSNASDSESSY</td>
<td>1, 2, 3, 4, 7, 8 and 9</td>
</tr>
</tbody>
</table>

Brackets indicate alternative inclusion of residue at the beginning of I3

Mammalian SH3 domains bind to proline rich regions of target protein and DLG1 is no exception; DLG1 SH3 and GK domains form an interaction with the synaptic protein GKHolding (GukH), enabling proper localisation of fellow junctional protein SCRIB (Newman and Prehoda, 2009). DLG1 SH3 interactions have been shown to facilitate the organization of protein complexes. DLG1 protein colocalizes with E-cadherin at the sites of cell-cell contacts in mammary epithelial cells (Reuver and Garner, 1997; Wu et al., 1998; Reuver and Garner, 1998). In these cells, depletion of DLG1 altered the integrity of adherens junctions or tight junctions (Laprise et al., 2004; Stucke et al., 2007). Furthermore, in Drosophila Dlg, loss of the SH3 and GK domains leads to a dissociation of septate junctions, reinforcing the importance of DLG1 domains in facilitating the protein interactions that form cell-cell contacts (Woods et al., 1996; Hough et al., 1997).
In all MAGUK proteins, the GK domain is catalytically inactive, despite sharing a strong structural similarity with cellular soluble guanylate kinases (Kuhlendahl et al., 1998; Olsen and Bredt, 2003). Studies in C. elegans have shown that the GK domain is important for cell viability (Lockwood et al., 2008), suggesting that, though inactive, the domain contributes to DLG1 function. Investigation in human cells have revealed that the GK domain forms a number of interactions with cellular proteins, including guanylate kinase associated protein (GKAP), microtubule-associated protein 1A (MAP1A), guanylate kinase associated kinesin (GAKIN), kinesin superfamily protein 13B (KIF13B), brain-enriched guanylate kinase-associated protein (BEGAIN), SPA-1-like protein (SPAL), and leucine glycine asparagine repeat protein (LGN) (Zhu et al., 2011; Sans et al., 2005; Deguchi et al., 1998; Saadaoui et al., 2014; Brenman et al., 1998; Yamada et al., 2007; Kim et al., 1997). These protein-protein interactions facilitate the attachment of DLG1 to cytoskeletal factors, microtubules and actin-based transport machinery, as well as aiding the organization of signal transduction complexes. The GK domain has also been shown to associate with protein kinase C (PKC) and is required for the localization of DLG1 to the synapse (Thomas et al., 2000; O'Neill et al., 2011). Recently, the DLG1 GK domain has been shown to bind to phosphorylated threonine or serine residues of fellow Scribble complex protein LGL and regulator of asymmetric cell division LGN (Zhu et al., 2014; 2011). DLG1 binding to phosphorylated LGL has been indicated as a regulatory mechanism in Scribble complex formation and binding to phosphorylated LGN has been suggested to be important in localizing DLG1 to the midbody along with LGN during cytokinesis (Zhu et al., 2011; 2014). Furthermore, a nuclear export signal has been identified in the GK domain (KRKK), removal of which restricts DLG1 to the cytoplasm (Kohu et al., 2002). Since a number of DLG1 interacting proteins, including protein 4.1 and ErbB4, exist in the nucleus, it has been speculated that
CHAPTER 4: IDENTIFICATION AND CHARACTERISATION OF A NOVEL ISOFORM OF DLG1

DLG1 translocates with these proteins, to function in the nucleus as a transcription factor, splicing factor or scaffolding protein (Kohu et al., 2002; Lue et al., 1994; Huang et al., 2000). Post-translational modifications of DLG1 have been reported. Commonly, phosphorylation is related to the control of DLG1 subcellular localization. In D. melanogaster, phosphorylation of Dlg PDZ1 by CAMKII decreases the synaptic localisation of the protein (Koh et al., 1999). In human cells, phosphorylation of DLG1 by the MAPK p38γ promotes dissociation from GKAP, releasing the protein from the cytoskeleton (Sabio et al., 2005). Furthermore, stress conditions induce phosphorylation of DLG1 by JNK (Jun N-terminal kinase), which leads to the accumulation of DLG1 at cell-cell contacts (Massimi et al., 2006). DLG1 interacts with the mitotic kinase PBK (PDZ binding kinase) and both proteins are phosphorylated during mitosis (Gaudet et al., 2000). The differential phosphorylation of DLG1 during S phase and mitosis is likely to affect the interactions that DLG1 can form and thus function to control the protein interactions during the cell cycle (Gaudet et al., 2000). Furthermore, DLG1 is differentially phosphorylated during the cell cycle by CDK1 and 2, enhancing nuclear localisation of the protein (Narayan et al., 2009; 2008).

4.1.1.1 DLG1 Insertions

Extensive alternative processing of the DLG1 gene further diversifies DLG1 protein interactions and functions, and this occurs both in Drosophila and in humans (McLaughlin et al., 2002). Splice variants differ in the use of alternative exons, which have been identified in a region at the N-terminus and in the C-terminal HOOK region, between the SH3 and GK domains (Figure 4.2). These insertions have been found in different combinations in different variants of DLG1 protein and they affect DLG1 localisation as well as promoting interactions with cellular factors (Table 4.1).
The N-terminal alternatively spliced region of DLG1 comprises the proline rich insertions I1A and I1B, which are able to moderate oligomerisation, forming interactions with other DLG1 molecules and with other cellular proteins (Nix et al., 2000; Wu et al., 2000). Immunoprecipitation of DLG1 and the pro-proliferative tyrosine kinase Lck in T-cells demonstrated that the SH2 domain of the kinase interacts with DLG1 I1A and I1B (Hanada et al., 1997). More recently, it has been shown that the I1A-containing DLG1 isoform is phosphorylated in response to T-cell receptor (TCR) stimulation and that this event is mediated by Lck (Crocetti et al., 2014). It has been suggested that this interaction allows DLG1 to couple TCR activation to the release of cytotoxic granules via Lck (Round et al., 2007a; Crocetti et al., 2014; Hanada et al., 1997).

The amino acid sequences of insertions I1A and I1B are involved in DLG1 self-association (McLaughlin et al., 2002; Hanada et al., 2000; 1997). Furthermore, I1A has been shown to influence the activity and surface expression of the voltage-gated potassium channel Kv1.5; presence of the insertion enables DLG1 to stimulate Kv1.5-mediated current (Godreau et al., 2003; McLaughlin et al., 2002). The stimulatory function of I1A containing isoforms may be due to the ability of the protein to organise synaptic channels and facilitate interactions with accessory proteins (Hibino et al., 2000; DeMarco and Strehler, 2001).

The inclusion of insertions affects DLG1 function, for example McLaughlin et al. suggest that isoforms containing I1A and I1B are involved predominantly in DLG1 signalling functions, whereas isoforms without these insertions are predominantly involved in scaffolding. This is perhaps most clearly observed in T-cells where the two expressed forms of DLG1 differ by the inclusion or exclusion of I1B. Both forms facilitate cytotoxic granule release but the I1B-containing form is unable to coordinate p38 activation and proinflammatory cytokine production (Crocetti et al., 2014).
Nuclear functions of DLG1 are perhaps mediated by the I2 isoform. Inclusion of I2 in DLG1 is associated with relocation of the protein to the nucleus and it has been suggested that the I2-form is involved in chromatin reorganization (McLaughlin et al., 2002). During epithelial differentiation, I2-containing forms of DLG1 translocate from the nucleus to the cytoplasm, suggesting that the nuclear I2-mediated functions are not necessary for differentiated keratinocytes (Roberts et al., 2007).

The alternatively spliced I3 is associated with DLG1 localisation to the cell membrane and has been shown to interact with the membrane-bound proteins β1-Adrenergic GPCR, A-kinase anchor protein 5 (AKAP5) and protein 4.1 (Lue et al., 1996; Nooh et al., 2013; Gardner et al., 2007). The I3 containing isoform of DLG1 promotes the recycling of β1-Adrenergic receptor by DLG1 and provides a weak binding site for CAMKII; in the absence of I3, CAMKII is able to bind upstream of I2, indicating that this interaction is important for more than one DLG1 isoform (Paarmann et al., 2002). Phosphorylation of I3 by CAMKII disrupts binding of the PKA organizing protein AKAP5 to the insertion sequence (Nikandrova et al., 2010; Schlueter et al., 2006). Protein 4.1 binding to I3 facilitates DLG1 recruitment to the neuronal synapse, association with the actin cytoskeleton within dendritic spines and association to membranes in both polarized and non-polarized cells (Hanada et al., 2003; Lue et al., 1994; Rumbaugh et al., 2003; Nooh et al., 2013). The HOOK region also contains a calmodulin binding site, a region which is also able to compete for binding to DLG1 I3 in homodimerisation (Paarmann et al., 2002).

The combinations of insertions included in DLG1 are also dependent upon the tissue in which they are expressed. I3 is predominantly expressed in cardiac tissue whereas the I2-containing form appears to be restricted to neurons and liver (Mori et al., 1998; McLaughlin et al., 2002; Godreau et al., 2003). However, both insertions have been detected in epithelia,
although the expression of each changes during epithelial differentiation (Godreau et al., 2003; McLaughlin et al., 2002; Roberts et al., 2007).

Although the first 65 amino acids of DLG1 direct the recruitment of DLG1 to sites of cell-cell contact, and the presence of I3 affects the efficiency of this localization, DLG1 mutants without PDZ domains show a nuclear localization. Insertions and domains all influence DLG1 localisation and function (Wu et al., 1998). This is in part due to the fact that the interactions formed by DLG1 involve more than one domain interface; for example, binding of the SH3 increases affinity of DLG1 to CAMKII and promotes calmodulin and protein 4.1 binding to the HOOK domain (Gardoni et al., 2003; Mauceri et al., 2007).

4.1.2 DLG1 expression

DLG1 mRNA is transcribed from chromosome 3 and consists of 24 described exons (Figure 4.3). Transcription of the DLG1 gene is repressed by the snail family of transcription factors and a number of transcription factor binding sites have been identified in the 5' untranslated region (UTR) of DLG1, including AP-1, AP-2 NFkB and SP-1 (Cavatorta et al., 2008). The DLG1 gene contains 50 distinct introns which could be transcribed to produce 33 different mRNAs (Thierry-Mieg and Thierry-Mieg, 2006). It has been suggested that there are 11 potential alternative promoters and 16 alternative polyadenylation sites (Cavatorta et al., 2008; Thierry-Mieg and Thierry-Mieg, 2006). Two different 5' UTRs have been identified by the Gardiol lab (Cavatorta et al., 2011). These UTRs are generated by alternative splicing within the 5' region of DLG1 mRNA and is a region which mediates DLG1 transcriptional control, influences mRNA stability, localization and translational efficiency (Cavatorta et al., 2011). It is therefore reasonable to
assume that the inclusion of different 5’ regions affects DLG1 protein abundance (Cavatorta et al., 2011).

Thus, DLG1 is a multi-functional protein, able to form a range of interactions and perform cell context and subcellular-specific roles. DLG1 was the first identified PDZ target of high-risk α-HPVs and the ability of the E6 oncoprotein to target the protein for ubiquitin mediated degradation is well documented (Kiyono et al., 1997; Gardiol et al., 2006; Narayan et al., 2009; Thomas et al., 2005). The virus has been shown to target specific subcellular pools of DLG1, exhibiting a preference for nuclear, phosphorylated forms of the protein (Narayan et al., 2009; Massimi et al., 2004).
4.3 Hypothesis and objectives

A number of forms of the DLG1 protein have been identified, with diversity introduced via short, alternatively spliced insertions and expression of alternative N-terminal sequences (Figure 4.1 and 4.2). The expression of DLG1 isoforms is tissue and cell context specific, and likely affects DLG1 localisation and function within a cell. Moreover, many studies of DLG1 have focussed on the β isoform and little is understood of DLG1α expression. DLG1 is a major target of high-risk HPV E6 oncoproteins and this study has identified an increase in DLG1 transcription in the presence of HPV genomes. Thus, this investigation aims to characterize DLG1 expression in epithelia under a number of conditions, in the presence of high-risk HPVs, in normal tissue and in cancer tissue.

The experiments outlined in this chapter aim to:

i. Characterise DLG1 expression during the HPV life cycle

ii. Investigate novel DLG1 isoform in the context of mRNA and protein

iii. Characterise DLG1 isoform expression in cancer cell lines and normal tissue
4.4 Results

4.4.1 Identification of a novel insertion in DLG1 in mRNA

Previously in this thesis, an investigation into the expression of DLG1 during the HPV life cycle was carried out (Chapter 3). During this, it was noted that mRNA expression of the DLG1 gene was increased in the presence of high-risk HPV genomes (Chapter 3, Section 3.3.2). Since DLG1 exists as two main isoforms and is further diversified by the inclusion of spliced insertions, it was deemed important to characterise the expression of DLG1 in these α-HPV genome containing cells and to determine which of the isoforms were being upregulated in the genome containing cells. Microarray and subsequent qRTPCR analysis had identified that the increased level of DLG1 expression occurred in transcripts containing a region in the HOOK domain. The primers used in qRTPCR analysis were designed according to the nucleotide sequence of the Affymetrix probe (230229_at), both of which covered the HOOK domain region of DLG1 (Figure 4.3, Table 4.2). It was surprising that, upon examination of the probe position and primer sequences (Table 4.2), the region under analysis in the Affymetrix array corresponded to was in fact a region of the DLG1 gene previously described as intronic, between known exons 21 and 22. Exon 21 encodes DLG1 insertion I2 and exon 22 encodes I5 (Figure 4.3, Table 4.2). Furthermore, Affymetrix included this probe because of the previous identification of two cDNA clones from hypothalamus- and brain-derived libraries, containing sequence overlapping with this region (Strausberg et al., 2002; Wiemann et al., 2003, Figure 4.3).

To further examine the sequence of this HPV-upregulated ‘intronic’ region of DLG1, PCR primers were designed to the flanking exons of I2 and I5 (I2_fwd and I5_rev, Figure 4.4) and used in combination with the HOOK region primers (HOOK_fwd and HOOK_rev, Figure 4.4), previously used in qRTPCR (Chapter 3), to amplify from cDNA generated from
Figure 4.3: *DLG1* HOOK domain corresponding primers and probes

(A) Schematic of the *DLG1* gene showing the 23 exons, and the exon start and finish nucleotide numbers, according to NCBI. The exons coding for insertions I1A, I1B and I2-I5 are shown. The region included in the dashed line is the region of focus in B.

(B) Exon 21 and 22 encode I2 and I5, respectively. It is between these exons where the microarray probe (blue line) and the HOOK domain primers (blue arrows) anneal. This region overlaps with the known cDNA IMAGE clone, derived from human hypothalamus, BC037916 (Strausberg *et al*. 2002), corresponding to nucleotides 231928 to 233005 of *DLG1* (orange line). Also in this region is an area of alignment with IMAGE clone DKFZp761P0818 derived from amygdala (Wiemann *et al*. 2003, green line). Further primers were designed to known exonic sequence in Exons 21 and 22, indicated by the pink and purple arrows.
CHAPTER 4: IDENTIFICATION AND CHARACTERISATION OF A NOVEL ISOFORM OF DLG1

DNAse-treated polyA selected RNA that had been isolated from HPV16 genome positive keratinocytes. A control reaction containing water was included. The PCR products were separated by ethidium bromide agarose gel electrophoresis and the lower band observed in both the cDNA and water control indicated amplification of primer dimers (Figure 4.4). The amplified bands were excised; the single band amplified by the HOOK primers and the single product amplified by the HOOKFwd in combination with the I5Rev (Figure 4.4). In the case of the HOOKFwd I5Rev and I2Fwd I5Rev combinations the brightest and the fainter bands was excised (a total of 2 and 3 bands cut out, respectively), but in both cases the fainter bands were unsuccessful in sequencing reactions (Figure 4.4).

Table 4.2: DLG1 primer and array probe positions

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5'-3')</th>
<th>Position in DLG1</th>
</tr>
</thead>
<tbody>
<tr>
<td>I5Rev</td>
<td>CCTGTCTTTTCATAGGTCCCA</td>
<td>230051-230071</td>
</tr>
<tr>
<td>I2Fwd</td>
<td>GAGATCCCTGACGACATGGG</td>
<td>223464-223484</td>
</tr>
<tr>
<td>HOOKFwd</td>
<td>GCCCATTCCTAGGGATGGCAATGC</td>
<td>228817-228841</td>
</tr>
<tr>
<td>HOOKRev</td>
<td>GCCATTCCTGGGATGGCAGAG</td>
<td>228908-228930</td>
</tr>
<tr>
<td>Affymetix probe 230229_at</td>
<td></td>
<td>227651-228736</td>
</tr>
</tbody>
</table>

Sequencing of excised bands from the I2Fwd I5Rev combination of primers revealed the established I2-I5 junction. Successful sequencing of excised bands from I2Fwd HOOKRev combination and the HOOKFwd I5Rev combination revealed the presence of 208 nucleotides between exon 21 and 22, which has not been shown previously to be transcribed (Figure 4.5, sequence data Appendix II). The 208 nucleotides corresponded to DLG1 gene sequence in a region of the gene between exons encoding I2 and I5, sequence previously thought to be intronic (Figure 4.4). Further sequencing reactions confirmed the presence of this intronic sequence in mRNA containing I2 and I5, with sequence data覆盖ing the splice junctions (Figure 4.5). Thus, the region amplified by the HOOK region primers, and
Figure 4.4: Amplification of the DLG1 HOOK region

(A) Schematic of the primers positioned with respect to the DLG1 gene (Boxes indicate exons and the line intronic regions). Primers indicated by arrows, with names above. Pink, I2_{Fwd}; Blue HOOK_{Fwd} and HOOK_{Rev} and purple I5_{Rev}.

(B) The PCR products amplified using primer combinations were separated by agarose gel electrophoresis and visualised using ethidium bromide. cDNA was transcribed using oligodT primers from DNase-treated RNA extracted from HPV16-genome containing keratinocytes. PCR containing water was included as a control. Bands indicated by the arrow were excised and sequenced using the above primers. Primer dimers indicated by *. Fainter bands were not successfully sequenced. Ladder shown is 1 µg of 1kb plus (Invitrogen). Expected molecular weights of amplified fragments (Left to right) are 98, 149, 222 and 274 bp.
**Figure 4.5: Sequencing across exon junctions**

(A) Purified amplified DLG1 fragments were subject to sequencing reactions to reveal sequence indicated. Pink corresponds to exon 21 and purple to exon 22. Exons are indicated in capitals and intronic sequence in lower case.

(B) Sequencing results corresponded to DLG1 sequence and covered the splice sites of exons 19, 21, the novel exon, 22 and 24. The schematic indicates DLG1 mRNA with the order of the exons and the nucleotide sequence of splice junctions and the corresponding image of sequencing peaks below. Red dotted line indicates splice junction.
covered by the Affymetrix microarray probe 229130_at, is a region that is expressed and is upregulated in the presence of HPV genomes. In keeping with DLG1 insertion terminology, this exon was termed I6.

As the data indicated a novel insertion in the HOOK region, flanked by I2 and I5, the next stage was to characterise the transcripts containing the I6 insertion in more detail. PCR amplification was performed using a combination of forward primers specific to the N-terminus of either DLG1α or DLG1β isoforms, alongside a reverse primer within the novel insertion I6 (Figure 4.6). Both combinations of primers amplified fragments and more than 6 bands were visible from each reaction. Successful sequencing was achieved of one excised band (~1000 bp), amplified by the DLG1αFwd and HOOKRev combination (Figure 4.6). Sequencing of this purified fragment revealed the presence DLG1α specific N-terminus along with the first two PDZ domains (Figure 4.6, Appendix II). Thus, this data indicates that the I6 insertion is present in α forms of DLG1.

However, it cannot be ruled out that that I6 is present in DLG1β transcripts; more than 6 amplified bands were visible, which were not successful in sequencing reactions. The bands may be a result of non-specific amplifications due to the increased PCR extension time (1.5 min) and number of amplification cycles (50); increased cycling increases the opportunity for non-specific amplification and errors (Bustin and Nolan, 2004; Bell and DeMarini, 1991). This is something that requires further investigation.

Thus, this study has identified a novel insertion in an established region of extensive DLG1 alternative splicing. It is likely that mRNAs containing this insertion (I6) also encode the α form of the DLG1 N-terminus. Furthermore, I6 is upregulated in the presence of both high-risk HPV16 and 18 genomes (Chapter 3, Figure 3.4). The rest of this chapter aims to investigate potential consequences of the inclusion of the I6 sequence in DLG1 mRNA.
Figure 4.6: Amplification of the DLG1 N-terminus

(A) Schematic of the primers positioned with respect to the DLG1 N-terminus (Boxes indicate exons and the line intronic regions). Primers indicated by arrows, with names above. Green, DLGβ_Fwd and Rev; Purple DLGα_Fwd and Rev.

(B) Agarose gel electrophoresis separated PCR products. cDNA was transcribed using oligodT primers upon DNAse-treated RNA extracted from from HPV16-genome containing keratinocytes, or water as a control. Bands were excised and sequenced. Star (*) indicates successfully sequenced band.
4.4.1.1 Splice site analysis of novel insertion

The analysis of DLG1 transcription in the genome containing keratinocytes identified an isoform of DLG1 that contained a previously unknown insertion. To examine the splice donor and acceptor regions of this novel exon and compare them to the consensus sequences, splice site analysis was carried out upon the DLG1 gene sequence between exons 21 and 22. Splice site analysis was also carried out on the known DLG1 spliced insertions that result in the α or β N-terminus and in the inclusion of I1A I1B, I2, I3, I4 and I5. A computational splice site analysis programme designed to predict the effects of mutations on splicing signals or to identify splicing motifs in any human sequence, was used alongside knowledge of known splice site recognition motifs (Table 4.3). The consensus sequences included the nucleotides at the 5' end of the donor exon (\(\wedge\)CAG) and 3' of the acceptor exon (\(\wedge\)TAG), as well as intronic sequence proximal to the exons, i.e. 3' of the donor and 5' of the acceptor exons, with the consensus nucleotides GU\(^\wedge\)/CAGU and CAG, respectively (Table 4.3).

<table>
<thead>
<tr>
<th>5’Exon-3’Exon</th>
<th>5’ Donor</th>
<th>Intron (5’-3’)</th>
<th>3’ Acceptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus</td>
<td>---------(\wedge)AG</td>
<td>GU(^\wedge)/CAG---CAG</td>
<td>(\wedge)T------------</td>
</tr>
<tr>
<td>α N-terminus-I1B</td>
<td>CTCCACCAG</td>
<td>GTAACA---AAG</td>
<td>GCAAATCTCTCC</td>
</tr>
<tr>
<td>β N-terminus-I1B</td>
<td>ACCAAATAAG</td>
<td>GTAGGA---AAG</td>
<td>GCAAATCTCTCC</td>
</tr>
<tr>
<td>I1A-I1B</td>
<td>CATACCACAG</td>
<td>GTTTTT---AAG</td>
<td>GCAAATCTCTCC</td>
</tr>
<tr>
<td>I2-I5</td>
<td>GGGCTGGAAG</td>
<td>GTAAGT---CAG</td>
<td>CATGTAACTTC</td>
</tr>
<tr>
<td>I3-I5</td>
<td>GATGCTGACC</td>
<td>GTAAGT---CAG</td>
<td>AGCATGTAACT</td>
</tr>
<tr>
<td>I2-I6</td>
<td>AAAGCCGTA</td>
<td>GTAAGT---CAG</td>
<td>TAGTTTTACCA</td>
</tr>
<tr>
<td>I6-I5</td>
<td>TTTATTTTTC</td>
<td>CTTAAA---CAG</td>
<td>CATGTAACCT</td>
</tr>
<tr>
<td>I5-GK</td>
<td>AGTAGTTACC</td>
<td>GTAAGT---TAG</td>
<td>CTTAAGCAT</td>
</tr>
<tr>
<td>I5-I4</td>
<td>AGTAGTTACC</td>
<td>GTAAGT---AAG</td>
<td>TAATCCTGATT</td>
</tr>
</tbody>
</table>

Consensus derived from Srebow and Kornbluth, 2006; Ghigna et al., 2008; Harris and Senapathy, 1990. Full sequences of exons are listed in Appendix II.
Examination of the known DLG1 exons encoding the insertions revealed that the exons corresponding to the α (exon 5) and β (exon 3) N-terminus, and to I1A (exon 6) and I2 (exon 21), all contain a consensus 5’ AG dinucleotide. However, the exons encoding I3, I5 and the new insertion I6 contain a 5’ cytosine that does not conform to the consensus sequence (Table 4.3). Similarly, the canonical 3’ C/T AG acceptor nucleotides were present in exons encoding I5 and I6 but not in the other exons analysed (Table 4.3).

Although there are consensus sequences, these are not strictly adhered to by all the exons of DLG1. The intronic sequence is as important as the donor and acceptor sequences in splicing, as it contains the branchpoint adenosine and regions of spliceosome recognition (Williams, 2002; Harris and Senapathy, 1990). Screening of the DLG1 gene sequence showed conservation of a 5’ GT in all the introns analysed, apart from between I6 and I5, where a cytosine replaced the guanosine (Table 4.3). The 3’ intronic AG dinucleotide was conserved in all the sequences included in the screening analysis (Table 4.3). The non-adherence to consensus sequence may indicate obscure or ‘weak’ splice sites, where the inclusion of exons requires additional specific factors, which may be cell or condition specific.

4.4.2 Investigation into protein expression of the novel DLG1 isoform

4.4.2.1 In silico screening suggests interacting proteins for novel insertion

In order to investigate the potential roles for the novel DLG1 sequence, the nucleotide sequence was translated in silico using the ExPasy resource and the resulting amino acid sequence of the insert I6 and upstream I2 was screened for potential protein-protein interactions using ScanSite software. An amino acid sequence of 20 residues was predicted, followed by a stop codon (Figure 4.7). In the context of the whole DLG1 protein, inclusion of this sequence would produce a truncated protein containing the three
**A**

Frame 1
STOP FYQFAHSTOP RMETNSHCCISLSMET LDSVLCFHFSHSGGWQRHSLWLSTOP FAFPSTOP QVINLNTFLYVYWLFDCLL

Frame 2
SFTSLPIPSASTOPTPIAAYHRCLTCLFCVFILAILVDGRDTLCGYNLHFPDKASTEPLTSTOP TLYMETFIGYLTVFF

Frame 3
VLPVCPLAHELPLLHII IDASTOPHVCFVFSFSTOP PFWMET AEJLVVIPICISLTSN STOP LEHFSICLVI STOP SLSS

**B**

REIPDDMSKGLIVLPVCPLAHELPLLHII IDASTOPHVCFVFSFSTOP PFWMET AEJLVVIPICISLTSN STOP LEHFSICL
I STOP LSSLACNFSTOP CQRSTOP STOP KSTOP L

**Figure 4.7 In silico translation of nucleotide sequence**

(A) Translation of the nucleotide sequence of the novel insertion into amino acid sequence. The sequence was translated in three frames.

(B) Translation of sequenced fragment including exons 21 (I2, pink) and 23 (I5, purple) at either side. The novel insertion is designated I6. Non-translated sequence is indicated in grey. The final lysine of I2 is excluded in the I2-I6 variant.
PDZ domains, followed by the SH3 domain, I2 and a novel sequence of 21 amino acids; the truncated protein however would lack the GK domain (Figure 4.7). Furthermore, the final lysine of I2 is not included, instead an isoleucine takes its place. Similar alternative inclusion of single amino acids in insertions has been observed with the glutamine residue at the beginning of DLG1 I3 (McLaughlin et al., 2002). The predicted molecular mass of the novel isoform is 60 kDa, compared to the full-length protein of 120 kDa.

Table 4.4: Potential protein interactions with I2-I6 sequence

<table>
<thead>
<tr>
<th>Motif Group</th>
<th>Motif</th>
<th>Gene</th>
<th>Sequence target</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinase binding site group</td>
<td>Erk D-domain</td>
<td>MAPK1</td>
<td>DMGSKGLIVLVPVF</td>
<td>0.868</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GSKGLIVLVPVFLLA</td>
<td>0.798</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KGLIVLVPVFLLAHE</td>
<td>0.726</td>
</tr>
<tr>
<td>Src homology group (SH3)</td>
<td>Cbl-associated protein</td>
<td>SORBS1</td>
<td>LIVLVPVFLLAHELP</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>Cortactin SH3</td>
<td>HCLS1</td>
<td>LIVLVPVFLLAHELP</td>
<td>0.760</td>
</tr>
<tr>
<td></td>
<td>PLCg SH3</td>
<td>PLCG1</td>
<td>LIVLVPVFLLAHELP</td>
<td>0.687</td>
</tr>
<tr>
<td></td>
<td>Src SH3</td>
<td>SRC</td>
<td>SKGLIVLVPVFLLAHE</td>
<td>0.658</td>
</tr>
<tr>
<td></td>
<td>P85 SH3 mode 1</td>
<td>PI3KR1</td>
<td>SKGLIVLVPVFLLAHE</td>
<td>0.665</td>
</tr>
</tbody>
</table>

Target residues underlined. The SH3 domain binds proline residues, the Erk D-domain consensus sequence is Φ-X-Φ-P, where Φ indicates a hydrophobic residue.

The translated protein sequence of I2-I6 was analysed for secondary structure motifs to predict protein interacting domains, with a score to indicate the probability of an interaction, where 0 indicates an unfavourable interaction and 1 a highly favourable interaction (Table 4.4). Two protein binding motifs were highlighted via the screen: the kinase binding site group and the SH3 group (Table 4.4). Five protein families were suggested to bind the central proline of the novel insertion via an SH3 domain, including regulatory subunits of the kinases PI3K and PKCγ (Table 4.4). The mitogen-associated kinase MAPK1 was also predicted to bind to three residues within the insertion, an isoleucine, leucine and a valine, all predicted to bind with a score of over 0.7 (Table 4.4).
These residues stand in the place of ‘X’ in the Erk D (docking)-domain binding consensus sequence (φ-X-φ-P, Table 4.4, Sheridan et al., 2008; Fernandes and Allbritton, 2009).

4.4.2.2 DLG1 isoform expression analysis using antibodies

In order to identify whether a truncated form of DLG1 is expressed, several antibodies against DLG1 were used to detect protein extracted from cellular fractions of HPV16 and HPV18 genome containing keratinocytes, as well as from the HPV18 positive cancer cell line HeLa. These antibodies were a selection of commercial and non-commercial antibodies raised to different regions of DLG1 (Table 4.5). Control blots incubated with either secondary anti-mouse or anti-rabbit IgG antibody were carried out to highlight any non-specific binding (Figure 4.8). In these control blots, a non-specific band can be seen at 30 kDa in both the nucleus and cytoskeleton, when incubated with the anti-rabbit IgG horseradish peroxidase (HRP) -conjugated secondary antibody whereas no bands were detected in blots incubated with the anti-mouse IgG secondary antibody, which was conjugated to HRP (Figure 4.8).

**Table 4.5: DLG1 Antibodies**

<table>
<thead>
<tr>
<th>Antibody name</th>
<th>Source</th>
<th>Animal origin</th>
<th>Region of DLG1</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D11</td>
<td>Santa Cruz</td>
<td>Mouse</td>
<td>1-229 DLG1β</td>
<td>-</td>
</tr>
<tr>
<td>H-60</td>
<td>Santa Cruz</td>
<td>Rabbit</td>
<td>49-108 DLG1β</td>
<td>-</td>
</tr>
<tr>
<td>NAG</td>
<td>S. Roberts</td>
<td>Rabbit</td>
<td>1-200 DLG1β</td>
<td>Roberts et al., 2007;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>McLaughlin et al., 2002</td>
</tr>
<tr>
<td>NT</td>
<td>L. Banks</td>
<td>Rabbit</td>
<td>DLG1α N-terminus:</td>
<td>Designed during this thesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>KGSSDELQAEPEPSR</td>
<td></td>
</tr>
</tbody>
</table>

The predicted molecular mass of the truncated form of DLG1 was 60 kDa. The mouse monoclonal antibody 2D11 and the rabbit polyclonal antibody H60 were both raised to the N-terminus of DLGβ. As expected, both antibodies detected DLG1 migrating at
Figure 4.8: Cross-reaction of subcellular fractions with HRP-conjugated secondary antibody

HeLa and HPV18 genome containing keratinocytes were cultured in monolayer to 70% confluency and subjected to subcellular fractionation. The cytoplasmic (F1), membrane (F2), nuclear (F3) and cytoskeletal (F4) fractions were then separated by SDS-PAGE and probed with anti-mouse (top panel) or anti-rabbit (top panel) HRP-conjugated secondary antibody. Molecular weight markers show kDa.
approximately 120 kDa in both HeLa cells and in the genome containing keratinocytes (Figure 4.9 A). As noted previously, the majority of DLG1 was located in the nucleus. (Figure 4.9 A). Detection with 2D11 revealed no smaller bands, but the H60 antibody identified a number of small bands in the cytoplasmic fraction including several between 60-80 kDa in both donors and to a lesser extent in HeLa.

Likewise, the NT antibody detected full-length DLG protein plus an array of faster migrating proteins, but the profile differed from the one recognized by H60. NT detects several bands between 40 and 80 kDa in the membrane and cytoskeleton and at 80 kDa in the cytoplasmic fraction, all of which are stronger in the keratinocytes containing HPV genomes than in HeLa cells (Figure 4.9 B).

The N-terminal NAG antibody NAG is raised to a region of DLG1 encompassing the insertions I1A and I1B (McLaughlin-Drubin and Meyers, 2004; Lue et al., 1994). As with H60 and NT, the NAG antibody detected full length DLG1 and a number of smaller species (Figure 4.9 D). In particular, protein detected at approximately 65 kDa in the membrane and nuclear fractions are more prominent in genome containing cells than in HeLa (Figure 4.9 D).

Thus, protein analysis with a number of DLG1-specific antibodies suggested the existence of smaller (less than 120 kDa) species, including a form of approximately 60 kDa. However, the analysis of the I6 containing transcripts indicated a α isoform although it cannot be ruled out that β forms contain the I6 insertion too, and none of the existing antibodies recognised this isoform of DLG1. In order to assess the expression of DLG1α, an immunogenic peptide was designed from the N-terminus sequence unique to DLG1α by a commercial company (Generon, Figure 4.2). The peptide (KGSSDELQAEPEPSR) was used as an immunogen in rabbits and the subsequent
A

<table>
<thead>
<tr>
<th></th>
<th>HeLa</th>
<th>HPV18</th>
<th>HPV18</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>F2</td>
<td>F3</td>
<td>F4</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th></th>
<th>HeLa</th>
<th>Donor 1</th>
<th>Donor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>F2</td>
<td>F3</td>
<td>F4</td>
</tr>
</tbody>
</table>

F1, Cytoplasm  F2, Membrane  F3, Nucleus  F4, Cytoskeleton

Longer Exposure
Figure 4.9: Detection of DLG1 protein in cellular compartments of HeLa and HPV18 genome containing keratinocytes

The cytoplasmic (F1), membrane (F2), nuclear (F3) and cytoskeletal (F4) fractions were sequentially extracted from monolayer grown HeLa and HPV18 genome containing keratinocytes derived from two donor backgrounds, separated by SDS-PAGE and detected using various antibodies against regions of DLG1. Both 2D11 (A) and H-60 (B) are specific to the N-terminal region of DLG1β. The NT-antibody (C) detects…..The N-terminal antibody NAG (D) was developed in-house by the Roberts lab.

In each case, the bands corresponding to full-length DLG1 protein (100-120kDa) is indicated by the arrow head. Molecular weight marker indicates kDa.
antibody (DLG1α 1524) recognised the peptide in ELISA based assay (performed by Generon). The antibody was then tested upon protein extracted from keratinocytes containing HPV16 and HPV18 by western blotting. Rabbit serum extracted before exposure to imunogen (pre-immune) was used as a control for non-specific binding. No non-specific bands were detected using the pre-immune serum (Figure 4.10 A). In parallel experiments on the same lysates the 2D11 antibody confirmed the presence of full-length DLG1 migrating at approximately 120 kDa (Figure 4.10 B). The full-length protein was present in primary cells and in the genome containing keratinocytes when using the DLG1α antibody and a number of additional bands were detected in all lysates (Figure 4.10 C). Furthermore, in keratinocytes the bands detected by the DLG1α specific antibody migrated slightly faster than in HeLa, indicating differences in phosphorylation of the α forms between HeLa and other cells (Figure 4.10 C). Upon longer exposure, a profile of faster migrating bands were visible including several bands between 50 and 60 kDa, suggesting the presence of truncated forms of the proteins. However, there was no obvious change in the level of these smaller species between primary cells and those containing HPV genomes (Figure 4.10 C), which would be expected if the bands corresponded to the I6-containing form of DLG1 upregulated by HPV.

4.4.3 Investigation of DLG1 isoform expression

Chapter 3 details the expression of the HOOK domain containing transcripts of DLG1 in keratinocytes containing high-risk HPV genomes. Having identified that this corresponded to a region of DLG1, which was previously not known to be expressed, here expression analysis was extended to characterise the expression of DLG1 transcripts containing this novel insertion. In addition, the expression of the two main isoforms, DLG1α and DLG1β was also characterised because studies in Chapter 4 (Section 4.3.2) suggested that
<table>
<thead>
<tr>
<th></th>
<th>Donor 1</th>
<th>Donor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A

B

2D11

β-actin

Pre-immune

230

100

50

40

β-actin
Figure 4.10: Detection of DLG1α protein

Total protein was extracted from untransfected HFKs as well as HFKs containing HPV16 and HPV18 genomes from two donor backgrounds and from HeLa cells. All cells were grown in monolayer culture, harvested at 70% confluency and lysed in urea buffer. Protein was separated by SDS-PAGE and probed with pre-immune serum (A), commercial DLG1 antibody (2D11, B) and DLG1α 1524 (C). In each case, the bands corresponding to full-length DLG1 protein (100-120kDa) is indicated by the arrow head and loading control (β-actin) is positioned beneath. Molecular weight marker indicates kDa.
the novel insertion was expressed primarily as part of the α isoform. Moreover, our current knowledge of DLG1 is focused on the β isoform and little is understood of DLG1α expression. In qRTPCR analysis, primers were utilised that corresponded specifically to the DLG1α N-terminus, to the DLG1β L27 domain or to the HOOK domain I6 insertion (see Section 2.2.1.2, Table 2.2). cDNA was transcribed from DNAse treated RNA using oligo(dT) primers before being subjected to qRTPCR. Gene expression was calculated relative to the number of transcripts of the internal control β2microglobulin (B2M), before comparison to external control (e.g. untransfected keratinocytes), to generate ΔdCt fold change.

4.4.3.1 DLG1 isoform expression in keratinocytes containing high-risk HPV genomes

As previously described, a screen of DLG1 expression levels in keratinocytes containing high-risk HPV genomes was carried out by microarray, and was validated using qRTPCR (Chapter 3). The microarray included two probes corresponding to different regions of DLG1, the aforementioned HOOK region, and a probe that aligned to the N-terminus of the DLG1β isoform (Figure 4.11). Microarray analysis showed an increase in DLG1 transcripts containing the HOOK domain in monolayer-grown keratinocytes containing either HPV16 (11-fold) or HPV18 (7.5-fold) compared to untransfected keratinocytes from the same donor background (Table 4.6). Conversely, microarray analysis of DLG1β transcripts was not above the limit of 2.5 fold (Table 4.6). Gene expression analysis by qRTPCR verified the trend seen by microarray; in two donors a 18-20-fold increase in the DLG1 I6 insertion was observed in the presence of HPV16 genomes, and a 5-6-fold increase in HOOK transcripts was detected in HPV18 genome containing cells (Figure 4.11). Although there was no microarray probe corresponding to DLG1α transcripts, the use of specific primers to this isoform in qRTPCR demonstrated that DLG1α transcripts
Figure 4.11: *DLG1* isoform expression in primary keratinocytes containing α-HPV genomes

RNA was extracted from primary keratinocytes containing HPV16 or HPV18 genomes, alongside untransfected HFKs, from two donor backgrounds. Following DNAse treatment and translation using oligodT primers, cDNA was subjected to SYBR green qPCR, using primers specific to the DLG1α or DLG1β N-terminus, and to the I6-containing HOOK region.

Transcript levels were normalised to an internal B2M control and then to transcript levels in HFKs of the same donor background (indicated by dotted line). Error bars ±SEM of three experimental repeats and star (*) indicates a p value of <0.05.
were increased in both HPV16 and HPV18 genome containing keratinocytes, by more than ~10- and ~4-fold respectively (Figure 4.11).

Table 4.6: Microarray and qRTPCR analysis of DLG1 expression.

<table>
<thead>
<tr>
<th>Region of DLG1</th>
<th>HPV16</th>
<th>HPV18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Microarray</td>
<td>qRTPCR</td>
</tr>
<tr>
<td>DLG1α N-terminus</td>
<td>-</td>
<td>12.20</td>
</tr>
<tr>
<td>DLG1β N-terminus</td>
<td>1.70</td>
<td>1.97</td>
</tr>
<tr>
<td>HOOK Region</td>
<td>11.67</td>
<td>17.56</td>
</tr>
</tbody>
</table>

Microarray and qRTPCR analysis of DLG1 expression. Fold change relative to HFKs and average from 2 donor backgrounds.

Both DLG1α and I6-containing transcripts were elevated to greater levels in keratinocytes containing HPV16 genomes than cells containing HPV18, and this was consistent in two donor keratinocyte backgrounds (Figure 4.11). Data is summarised in Table 4.6.

4.4.3.2 DLG1 isoform expression during the productive life cycle

To establish whether the DLG1 isoforms are affected upon activation of the productive virus life cycle, HPV18 genome containing keratinocytes were grown in semi-solid media (1.5% methylcellulose), which stimulates differentiation and activates productive viral events including viral genome amplification (Wilson and Laimins, 2005). Total cellular RNA was extracted from cells harvested at 0, 24 and 48 h and DLG1 isoform transcripts analysed using the primers in RTPCR analysis. Transcript level quantification at 24 and 48 h was normalised to 0 h. DLG1α and DLG1β exhibited a ~3-fold increase in transcripts after 24 and transcript levels remained elevated at 48 h (Figure 4.12). DLG1 transcripts containing the novel HOOK region insertion demonstrated a 2-fold increase after 24 h growth in methylcellulose and remained at this levels at 48 h (Figure 4.12).
Figure 4.12: Expression of *DLG1* in differentiating keratinocytes containing HPV18

Primary keratinocytes containing HPV18 genomes cultured in semi-solid media (methylcellulose) and harvested at 24 and 48 h. Total RNA was extracted and transcribed to cDNA using oligodT primers, following treatment with DNase. Three primer sets utilized in SYBR green qPCR, to measure transcript levels of DLG1α, DLG1β and the I6-containing HOOK region. Fold change in expression was calculated relative to the internal control B2M, and then to corresponding 0 h timepoint. Error bars show ±SEM of three technical repeats.
4.4.3.3 Investigating HPV regulation of DLG1 isoform expression

Although changes in E6 PBM activity are unlikely to have an effect upon the transcription of DLG1 isoforms, the change in E6 PBM activity is associated with alterations in proliferation of the genome containing cells (Delury et al., 2013), which may be reflected by changes in DLG1 isoform expression. Therefore, qRTPCR analysis was carried out upon monolayer keratinocytes containing HPV18 WT genomes or cells harbouring the ΔPDZ and ΔPKA genomes. The data showed that transcript levels of DLG1 isoforms - DLG1α, DLG1β and HOOK (I6) - were unchanged in the presence of genomes that have alteration on E6 PBM activity compared to the WT cells (Figure 4.13). At this stage of the study, it is not understood how the virus regulates DLG1 isoform expression.

The studies to date have focussed on cells harbouring the viral episomes, which replicate as low copy extra-chromosomal plasmids. It may be that this effect on DLG1 is specific to this stage of virus replication and if so it would be expected to change if the episomes were lost and the viral genomes integrated into the host chromosomal DNA. To examine this, the expression of the different DLG1 isoforms was assessed in monolayer keratinocytes containing HPV18 genomes at early (P8, viral genomes are episomes) and late (P21, viral episomes are lost and viral DNA integrated) passages (E. Marsh, personal communication).

When compared to episomal genome containing keratinocytes, the levels of DLG1β transcripts remain unchanged (Figure 4.14). Although the trend indicates that DLG1α and HOOK domain containing transcripts were decreased in cells containing integrated HPV18; this is based upon three technical repeats in one donor and therefore requires further experimental repeats across multiple donors (Figure 4.14).
Figure 4.13: *DLG1* isoform expression in keratinocytes containing HPV18 genomes

RNA extracted from primary keratinocytes containing HPV18 WT, ΔPDZ or ΔPKA genomes was DNase treated and reverse transcribed using oligo-T primers. The transcript levels of *DLG1α*, *DLG1β* and the I6-containing HOOK region were measured by SYBR green qPCR, values were normalised to internal control (B2M) and then to levels of cells containing WT HPV18 genomes. Error bars show SEM of three technical repeats in two donors.
Figure 4.14: *DLG1* isoform expression in keratinocytes containing episomal and integrated forms of the HPV18 genome

RNA extracted from primary keratinocytes containing HPV18 genomes was DNAse treated and reverse transcribed using oligoT primers. The transcript levels of DLG1α, DLG1β and the I6-containing HOOK region were measured by SYBR green qPCR, values were normalised to internal control (B2M) and then to levels of cells containing episomal genomes. Error bars ±SEM of three technical repeats.
4.4.3.4 DLG1 isoform expression in normal tissue

DLG1 is expressed in a range of tissue types and isoforms are likely to be expressed in a tissue specific pattern (McLaughlin et al., 2002; Mori et al., 1998; Godreau et al., 2003). Thus, the expression of DLG1 isoforms DLG1α, DLG1β and HOOK region-containing transcripts were investigated in normal tissues of the brain, skin, kidney and live and in normal immune cells (T-cells). For comparison of transcript levels, fold change was calculated relative to DLGα transcripts. In all the tissues analysed, DLG1β transcripts were more abundant than DLG1α or the HOOK region (I6) containing transcripts. Expression of DLG1β in kidney and skin was more than 7-fold higher and the level of DLG1β transcript in the liver was 4-fold higher than DLG1α (Figure 4.15). In brain tissue, the difference in expression between the α and β isoforms was the largest of the tissues analysed; expression of DLG1β was 11.9-fold higher than DLG1α (Figure 4.15). The expression of HOOK domain containing transcripts varied in the tissue analysed. In brain and liver the transcripts were less abundant than DLG1α transcripts, whilst in kidney tissue HOOK domain containing transcript levels were more than 3-fold higher than the DLG1α−containing transcripts (Figure 4.15). HOOK domain containing transcripts were expressed equally with DLG1α in skin (Figure 4.15). Analysis of DLG1 expression in T-cells revealed equal levels of expression of the three transcripts assessed (Figure 4.15).

4.4.3.5 DLG1 isoform expression in cancer cell lines

Analysis of DLG2 expression has shown transcriptional upregulation of specific spliced variants in kidney cancer precursor growths (Zubakov et al., 2006). Furthermore, several studies have demonstrated differential expression of DLG1 in cancers (Watson et al., 2002; Frese et al., 2006; Cavatorta et al., 2004). To investigate the expression of DLG1 isoforms
Figure 4.15: DLG1 isoform expression in normal tissue

Total RNA was extracted from human brain, kidney, liver and skin tissue samples, as well as from T-cells. Following DNAse treatment and reverse transcription using oligodT primers, the levels of DLG1α, DLG1β and the I6-containing HOOK region transcripts were measured by SYBR green qPCR, values were normalised to internal control (B2M) and then to the levels of DLG1α transcripts. Error bars ±SEM of three technical repeats.
during growth of cancer cells, qRTPCR analysis was applied to epithelial cancer cell lines. These cervical cancer-derived cell lines comprise the HPV18 positive line HeLa, the HPV16 positive cell lines SiHa and Caski and the HPV negative cervical cancer cell line, C33A. The fold change in DLG1 expression in these cells was calculated relative to the primary cervical keratinocytes (Cx129), which were isolated from uterine cervical tissue removed for a non-cancerous condition and therefore are considered non malignant. None of the cervical cancer cell lines analysed exhibited a change in DLG1 isoform expression levels; with the exception of the HPV16 positive cell line Caski, which demonstrated a 4.9-fold increase in DLG1α transcripts (Figure 4.16). A trend towards a decrease in transcript levels was observed in the HPV negative cell line C33A, but this was not statistically significant (Figure 4.16). Expression of both DLG1β and HOOK domain containing transcripts remained unchanged in HeLa, SiHa and Caski (Figure 4.16).

Since high-risk HPVs have also been etiologically linked to the development of head and neck cancers (Gillison and Shah, 2001), DLG1 expression analysis was extended to HPV positive and HPV negative head and neck cancer derived cell lines. Four HNSCC cells lines were examined; SCC040 and Vu40T, which are both HPV negative, and the HPV16 positive cell lines SCC154 and Vu147 (Steenbergen et al., 1995; Rampias et al., 2009). Transcript levels in HPV positive cells were normalised to their HPV negative counterpart. All DLG1 transcripts detected were less abundant in SCC154 (HPV16+) compared to the SCC040 (HPV-), whereas DLG1α and HOOK domain containing transcripts are increased nearly 2-fold in the HPV positive Vu147 compared to Vu40T (Figure 4.17). The abundance of DLG1β was equal in Vu147 to Vu40T cells (Figure 4.17).

In summary, transcription of DLGβ is unaffected by HPV, whereas α is upregulated along with the I6-containing isoform in episomal genome containing cells. The DLG1α, DLG1β
Figure 4.16: *DLG1* isoform expression in cancer cell lines

Transcriptional analysis performed the HPV positive cervical cancer cell lines HeLa (HPV18), SiHa (HPV16) and Caski (HPV16), and the HPV negative cervical carcinoma-derived cell line, C33A. The levels of *DLG1α*, *DLG1β* and HOOK region containing transcripts were measured by SYBR green qRT-PCR, using DNAse treated RNA transcribed to cDNA with oligo(dT) primers. Expression levels were calculated relative to internal control (B2M), and then to non-cancerous cervical epithelial cells (Cx129 cells, indicated by dotted line). Error bars ±SEM of three experimental repeats, * denotes p<0.05.
**Figure 4.17: DLG1 isoform expression in head and neck cancer derived cell lines**

Transcriptional analysis performed on HPV negative (SCC040 and Vu40T) and HPV16 positive (SCC154 and Vu147) HNSCC derived cell lines. RNA was extracted, treated with DNase and transcribed to cDNA using oligodT primers. SYBR green qPCR was then carried out using three primer sets, to measure the levels of DLG1α, DLG1β and HOOK region transcripts. Relative expression levels were calculated relative to the internal control B2M, and then to the corresponding HPV negative cell line. Error bars show ±SEM of three technical repeats.
and I6-containing transcripts were detected in a range of cell types and, although the β form appears to be the predominant form in normal tissue, the observed change in the ratio between the two main isoforms in certain cancer derived cell lines indicates that perturbing the balance of DLG1 isoforms may play a role in cancer development.
4.5 Discussion

4.5.1 Identification of a novel insertion

DLG1 encodes five known insertions, occurring in two regions of the protein (McLaughlin et al., 2002). This study has identified a new exon in the DLG1 gene, which is inserted into an area of DLG1 where alternative splicing has already been identified (the HOOK region). The new insertion has been named I6. DLG1 transcripts containing the I6 insertion were detected in normal human liver, kidney, brain, skin and immune cells indicating that I6 is important in DLG1 function in normal tissue.

Interestingly, in primary human foreskin keratinocytes, transcripts containing I6 were upregulated by both high-risk viruses analysed, although the increase was greater in the presence of HPV16. The levels of these I6 containing transcripts were further increased upon activation of the productive virus life cycle. Together, these data indicate that I6-containing variants of DLG1 are important in the replication of the viral genome. Indeed when the expression of I6-containing transcripts was examined in cells containing integrated α-HPV genome containing cells, the levels were decreased compared to cells containing episomal genomes, perhaps suggesting that the inclusion of I6 is required for the maintenance of episomal genomes.

It may be that the increased inclusion of I6 in DLG1 is a function of the complete HPV genome. As integration disrupts the viral genome, and this data revealed a decrease in I6 upregulation upon integration, it follows that viral factors are perhaps involved in the upregulation of the I6 spliced variant. The viral E2 protein has been shown to affect the transcription and splicing of a number of human genes, particularly those implicated in cancer development and in cell motility (Gauson et al., 2014; Bouvard et al., 1994a). Investigation by (Gauson et al., 2014) identified 522 genes with differentially spliced exons.
in the presence of HPV16 E2. Although the mechanism of this is not known, E2 is known to bind to host transcription factors and may be able to bind and recruit splicing factors also (Johansson et al., 2012). Furthermore, E2 transcriptionally upregulates splicing factors, for example the factor SRSF1 (Mole et al., 2009). It is feasible that an increase in this factor increases splicing in the DLG1 HOOK region and promotes the inclusion of I6 in E2-expressing cells. In addition, work in the Roberts lab has recently identified relationships between HPV and splicing regulators (Prescott et al., 2014). As the virus is able to manipulate splicing machinery, it may be that such manipulations cause or lead to the alternative splicing in the HOOK domain and inclusion of I6. Furthermore, splice site analysis revealed that the I6 exon did not completely adhere to the established consensus sequence; the acceptor sequence conformed to consensus but the 5' donor was a weak match. This reflects a requirement for additional specific factors in splicing, a process which may be facilitated by the virus.

The question remains as to why this variant is upregulated in the presence the high-risk viruses. Perhaps the ability of the virus to target DLG1 protein via the PBM is a mechanism of control to counter virally induced overexpression of the protein, or perhaps upregulation of DLG1 containing I6 is a consequence of viral interactions with cellular splicing and transcription machinery. Investigation into the functional consequences of the variant should help to answer this.

As both the HOOK domain and α isoform were upregulated in HPV genome containing keratinocytes and were both downregulated upon integration of HPV18 genomes, it can be inferred that I6 is included in α N-terminus containing isoforms. Amplification using forward primers corresponding to DLG1α in combination with the reverse primer within the identified I6 insertion amplified a number of bands within the predicted size range of the mRNA (~1500 bp). Although sequencing of amplified fragments confirmed the
presence of the DLG1α 5’ sequence, further analysis is required to eliminate the possibility of an I6-containing DLG1β isoform. Cloning the entire cDNA would not only confirm this, but also allow endogenous expression of the variant and thus in depth analysis of the functional consequences of I6 inclusion.

4.5.2 Consequences of I6 inclusion on DLG1 function

Translation of I6-containing transcripts is predicted to produce a truncated form of DLG1 protein (Figure 4.18). There is a precedence for alternative splicing as a mechanism of generating PDZ protein diversity; notably alternative splicing of the related PDZ protein CASK results in the inclusion of a 40 residue loop within the GK domain and ERBIN is predicted to encode a variant with a truncated PDZ domain, which has yet to be characterised (Favre et al., 2001). Given the specificity of DLG1 isoform interactions in T-cells, where I1A-containing but not I1B-containing proteins are able to facilitate release of granules (Crocetti et al., 2014), and the tissue-specific nature of DLG1 expression, where I3 is predominantly in mammalian cardiac tissue whereas the I2-containing form appears to be restricted to neurons and liver (Mori et al., 1998; McLaughlin et al., 2002; Godreau et al., 2003), it is probable that I6 expression is tissue and cell dependent (Godreau et al., 2003; Lue et al., 1994).

The predicted truncated protein would lack the GK domain. Without the GK domain, DLG1 can no longer form associations with the cytoskeleton (Wu et al., 2000; Manneville et al., 2010). Furthermore, absence of the GK domain would remove the ability of DLG to bind to phosphorylated proteins LGN and LGL, thus affecting Scribble complex regulation. The GK domain is involved in localising DLG1 to the midbody of the cell during asymmetrical cell division. It has been observed that truncation of the Drosophila Dlg C-terminus negatively affects the ability of neuronal cells to divide asymmetrically but did
Figure 4.18: Predicted protein structure of I6-containing DLG1

(A) Translation of the new exon would produce a C terminal truncation, removing the DLG1 GK domain. We also suggest that this isoform lacks the N-terminal L27 domain.

(B) The absence of the GK domain may remove the ability of DLG1 to associate with cytoskeletal filaments, although homodimerisation is still possible with an intact SH3 domain. The presence of the PDZ and SH3 domains, as well as I2, may still influence DLG1 localisation and cellular interactions. Question marks imply suggested consequences of I6 inclusion.
not affect epithelial polarity (Bellaïche et al., 2001); mouse models expressing a truncated form of DLG1 lacking the SH3, GK and HOOK region display growth retardation, severe craniofacial abnormalities and perinatal death (Caruana and Bernstein, 2001; Klocker et al., 2002). Although no effect was observed by Caruana and Bernstein (2001) on synaptic clustering of the protein, the domain obviously facilitates important interactions during mammalian development. Interestingly studies in mice showed that, in the absence of the GK domain, DLG1 was restricted to adherens junctions (Naim et al., 2005). This may be due to the presence of the I3 insertion, which has been shown to direct membrane localisation of the protein (Naim et al., 2005; Roberts et al., 2007; McLaughlin et al., 2002). Altogether, in vivo studies indicate that the GK truncated form of DLG1 does not perform functions in cell division or development, eliminating a role for I6-containing variants in these functions.

As the I2 insertion has been identified as predominantly nuclear, and the GK nuclear export signal is predicted to be absent in this truncated variant, there is an expectation that the I6-containing protein residues in the nucleus (Roberts et al., 2007; Kohu et al., 2002). However, the precise role of nuclear DLG1 has not yet been elucidated; it is known to be important in the growth of undifferentiated epithelial cells (Roberts et al., 2007). Notably, I2 has been implicated in chromatin-association (Kohu et al., 2002; McLaughlin et al., 2002); as the I6 isoform was identified downstream of I2, the insertion may contribute to DLG1 interactions with cellular chromatin. The nuclear protein NET1 has been shown to interact with DLG1 to promote translocation of the complex to the nucleus and, based on its function at the plasma membrane, the PDZ protein may act as a scaffold protein, organising protein complexes and nuclear filaments (Garcia-Mata et al., 2007; Sabio et al., 2010). The related MAGUK proteins CASK and ZO-1, which are involved in organisation of protein complexes at cell-cell junctions and synapses, are able to translocate to the
nucleus and contribute to the transcriptional regulation of specific genes; the truncated form of DLG1 may function similarly. Moreover, DLG1 has been shown to interact with the nuclear RNA-binding protein PSF (protein-associated splicing factor, Sabio et al., 2010). Via the PSF interaction, DLG1 binds to various RNAs (Sabio et al., 2010). Although the domain of DLG1 responsible has not been identified, this is another indication that nuclear DLG1 is involved in the regulation of gene transcription. This, with the observation that the GK domain forms interactions with a number of cytoskeletal filaments, could suggest that a nuclear form of GK-truncated DLG1 performs nuclear-specific functions.

Alternatively, the truncated form of DLG1 may act in a regulatory fashion and modify full-length protein function in a dominant negative manner. A number of PDZ proteins have been shown to act in this fashion. For example dynamin-2 sequesters GOPC at the cell periphery to block clathrin-mediated endocytosis (Cheng et al., 2004) and overexpression of the PDZ containing chromatin organiser SATB1 exerts a dominant negative effect upon itself, completely removing the transcriptional repressor functions (Notani et al., 2011). Indeed, overexpression of a dominant-negative isoform of CASK results in the mislocalisation of endogenous DLG1 (Lee et al., 2002). The I1A and I1B insertions, as well as the DLG1 SH3 domain have all been shown to facilitate DLG1 homodimerisation (McLaughlin et al., 2002; Hanada et al., 1997) and these are likely present in the I6-containing form. Interestingly, in Drosophila, the expression of a DlgA isoform lacking the GK domain bound endogenous Dlg1 to enable trafficking of the truncated protein, acting as a ‘hitchhiker’ (Mendoza et al., 2003; Mendoza-Topaz et al., 2008; Asaba et al., 2003). Such a mechanism may also be employed by truncated I6-containing forms of DLG1. Moreover, functional screening of the I6 protein sequence hypothesised an SH3 binding site, which has the potential to interact with the full-length DLG1 SH3 domain. Altogether, it is possible that the full-length and truncated forms of DLG1 are able to bind each other,
forcing other cellular proteins compete for binding to full-length protein and thus control DLG1 interactions.

Screening indicated that inclusion of I6 introduces a MAPK1 binding site to DLG1. Upon activation, MAPK1 phosphorylates nuclear proteins to activate or repress transcription of genes involved in proliferation, differentiation and apoptosis. The interaction with I6-containing DLG1 may facilitate interactions with, and thus phosphorylation of, specific transcription factors, and thus direct transcription of viral life cycle promoting genes. Interaction with MAPK1 may also facilitate phosphorylation of DLG; the PDZ protein is a known target of the MAPK cascade proteins ERK5, p38γ and cell cycle regulators CDK1 and CDK2 and the interaction with I6 may regulate the location and stability of the isoform in a similar fashion (Sabio et al., 2005; Narayan et al., 2008; Iñesta-Vaquera et al., 2010).

Likewise, screening has suggested an SH3 binding site within the insertion, which may allow binding of the kinases Src or PI3K to enable DLG1 phosphorylation or interaction with the membrane located proteins phospholipase C, cortactin or SORBS1 to facilitate the formation of protein complexes at the cell periphery.

It is interesting to speculate whether the inclusion of I6 affects DLG1 mRNA stability. Cavatorta et al. (2011) have identified two alternative 5'UTRs of DLG1 which are likely to affect the stability of transcribed mRNA and efficiency of translation. Elements within mRNA have been shown to effect the stability of β-catenin transcripts and isoforms of ZO-1 exhibit different mRNA stabilities (Reichert et al., 2000; Bachurski et al., 1994). Alternatively, I6 transcripts may act at an mRNA level to regulate other DLG1 transcripts. For example, genes encode long non-coding RNAs (lncRNA) or microRNAs (miRNA), which partially anneal with their target protein coding mRNAs, which can trigger degradation and regulate expression (Hannon et al., 2006; Ponting et al., 2009; Wilusz et al., 2006).
2009; Wu and Brewer, 2012). In this manner, the I6-containing transcripts may act as a
dominant-negative regulatory element. Although we hypothesize that this novel DLG1
spliced variant is expressed as protein, the precise function will be characterized in more
detail in future studies (see Section 4.4.4).

4.5.3 Differential expression of DLG1 isoforms

DLG1 isoforms are differentially expressed in human tissues and deregulated in cancers
(Godreau et al., 2003; Fuja et al., 2004). Thus, the expression of DLG1 α and β isoforms
was investigated in various stages of HPV-associated malignant disease. Previous studies
from the Roberts lab have shown that DLG1 protein is overexpressed and mislocalised at
early stages and then disappears in late stages of cervical cancer (Watson et al., 2002). In
this study, DLG1α, but not β, transcripts were increased by the presence of high-risk HPV
genomes in primary keratinocytes and upon integration of HPV genome, the abundance of
β transcripts decreased. However, in the cervical epithelial cells analysed, only the α form
was upregulated in Caski cells, other cell lines indicated no change in expression of either
of the DLG1 isoforms investigated. Cells containing HPV16 (Caski and SiHa) or HPV18
(HeLa) and HPV negative (C33A) cells were included in the analysis, which indicates that
the effect upon DLG1 expression is not HPV specific. Both Caski and SiHa are HPV16
positive, but differ in the number of integrated genomes; more than 600 have been
detected in Caski and one in SiHa (Raybould et al., 2014). The difference in DLG1
expression between the two HPV16-driven cancer cells may be attributed to the vast
difference in copy number, but it is more likely that there are other mutational changes that
have occurred in the different malignancies from which the cells are derived. It has been
noted that DLG1 expression in C33A cells is low and, as the PDZ protein is deregulated at
either the gene or protein level in a number of cancers, it is likely that the cancers from
which these cell lines were generated had developed mechanisms other than via HPV to disrupt DLG1 expression (Fuja et al., 2004; Watson et al., 2002; Cavatorta et al., 2004; Frese et al., 2006).

Although the presence of HPV in HNSCC cells lines correlated with effects upon the expression of DLG1 isoforms, it varied amongst the cell lines as to whether the levels decreased or increased. Both SCC154 and Vu147 are HPV16 positive; the former indicated a decrease in all DLG isoforms whereas there was a trend towards an increase in $\alpha$-DLG1 transcripts in Vu147 cells. The disparity between cell lines may be due to a difference in cell background; SCC154 and SCC040 are derived from tongue, whereas Vu40 and Vu147 are both from floor of mouth cancers (Steenbergen et al., 1995; Rampias et al., 2009). This implies that $DLG1$ expression is context specific, but needs to be confirmed experimentally by further repeats, and could be extended by analysis of further HNSCC derived lines.

This expression analysis included differentiated cells and healthy tissue, where DLG1$\beta$ was more abundant than the $\alpha$ form. As keratinocytes were stratified in culture, the abundance of both DLG1 isoforms increased, which correlates with the established knowledge that DLG1 is upregulated during differentiation of epithelium (Rivera et al., 2009). The combinations of insertions included in DLG1 are known to depend upon the tissue in which they are expressed (Godreau et al., 2003; McLaughlin et al., 2002; Roberts et al., 2007). The presence of an L27 domain is important in targeting the $\beta$ isoform to the membrane, via DLG2, and specifically to the baso-lateral membrane of epithelia via CASK (Lozovatsky et al., 2009; Hanada et al., 2003). The DLG1-mediated formation and maintenance of cell-cell contacts is a crucial function of the PDZ protein in cells and the abundance of the DLG1$\beta$ isoform may therefore be unsurprising.
The specific functions of the α isoform have not been extensively studied, although it is known that the α and β forms differentially traffic AMPA receptors at the synapse (Waites et al., 2009). The specific upregulation of the α-isoform during episomal HPV genome replication may reflect the viral requirement for increased nuclear functions of the protein. In the nucleus, DLG1 has been implicated in chromatin reorganization, in the regulation of transcription, splicing and as a scaffolding protein (Kohu et al., 2002; Lue et al., 1994; Huang et al., 2000). Upregulation of DLG1α in epithelia may therefore promote the transcription of growth-promoting genes and/or promote viral genome replication in order to create an environment beneficial to viral propagation. In addition, DLG1, specifically I2-containing isoforms, is known to be important in the growth of undifferentiated epithelial cells (Roberts et al., 2007) and thus HPV upregulation of DLG1α may be a viral mechanism to promote the growth of keratinocytes.

4.5.4 Future directions

Although protein analysis using a range of DLG1 antibodies revealed a number of lower molecular species, none were obviously suggestive of a truncated isoform. However, this is limited by the absence of an untransfected keratinocyte control, which would enable a comparison of expression levels and thus the identification of the expected truncated protein expressed to a lesser degree. Design and use of an α-specific antibody did not identify a smaller form that would be expected if the upregulated I6 form were detected. Thus, further work aims to dissect the downstream consequences of I6 inclusion in DLG1 mRNA and/or protein. In the short term, as the I2 sequenced has been identified upstream of I6, antibodies specific to I2 could be applied to protein extracted from primary cells and from cancer cell lines.
Further characterisation of I6 expression will be carried out; production of an I6-specific antibody has not yet been successful, but expression of the sequence can be achieved by cloning techniques. In this way, an investigation into the consequences of endogenous expression of this novel variant in cells could clarify the effect upon full-length DLG1 subcellular localisation, identify interacting proteins and infer function, as well as assess the phenotype of cells where this protein is overexpressed. Furthermore, it would be of interest to identify the presence of I6-containing isoforms in a larger range of tissue than was investigated here, including primary cancer tissue.

Having generated an antibody specific to the α isoform of DLG1, future work could investigate whether the ratio of α and β isoforms mRNA corresponds to protein levels and could more specifically identify the localisation of the isoforms in epithelia. Moreover, immunoprecipitation experiments with this antibody will enable the identification of cellular interactors and refine our knowledge of the functions of DLG1 isoforms.

There is also further characterisation of spliced variants of DLG1 to be carried out. An I3 and I6 containing isoform was identified but not studied in this thesis. Work by McLaughlin et al. suggests that transcripts containing both I3 and I2 produced a protein truncated after the I3 region in vitro, but the presence and consequence of combinations of insertions can be additional investigated. The spliced insertions present in DLG1 HOOK region are likely to exist in a range of combinations (there are more than 500 mathematical possibilities), the investigation of which across different tissues and cells would greatly expand current knowledge of DLG1 isoform expression.
CHAPTER 5: REGULATION OF THE HUMAN PAPILLOMAVIRUS E6 PDZ BINDING MOTIF DURING THE VIRUS LIFE CYCLE

5.1 Introduction

5.1.1 Cellular PDZ interactions

In vivo many cellular PDZ-PBM interactions occur, acting to orientate cells, receptors and organise signalling scaffolds (Fan and Zhang, 2002). The PDZ domain is composed of six β-sheets and two α-helices; a groove between the second β-strand and the second α-helix forms the hydrophobic pocket where the PBM peptide binds, serving as an extra β-strand (Doyle et al., 1996; Songyang et al., 1997; Niethammer et al., 1998). Further interactions between main chain residues stabilise and add specificity to the binding, for example a loop at the end of the binding groove is involved in positioning the PBM peptide and the terminal COOH of the PBM is coordinated by a network of hydrogen bonds to main-chain amide groups (Doyle et al., 1996). When bound within the PDZ groove, the side chain of the terminal and -3 residues of the PBM peptide point towards the base of the peptide-binding groove (Songyang et al., 1997; Doyle et al., 1996). In some cases, recognition of the PBM involves residues outside of the four amino acid consensus motif, enabling specificity of interactions unique to individual PDZ domains (Figure 1.11, Niethammer et al., 1998).

The specificity of PDZ:PBM must be regulated in order to coordinate interactions spatially and temporally. As such, many PDZ:PBM interactions are regulated by phosphorylation of the PDZ-binding peptide (Table 5.1 and Table 5.2). Due to the charge and size of a phospho group, its introduction can prevent the PBM sterically from binding within the PDZ pocket (Akiva et al., 2012, Figure 1.11). For example, phosphorylation of the PBM
serine residue of the glutamate receptor subunit GluR2 (SVKI) prevents its interaction with PICK (protein that interacts with C kinase 1) and reduces affinity for GRIP (glutamate receptor interacting protein, Shao et al., 2010; Matsuda et al., 1999). Interaction with PICK1 and GRIP modulates the location of the receptor and influences AMPAR clustering at the cell membrane, which in turn controls the transmission of synaptic impulses (Matsuda et al., 1999). Thus, phosphorylation of the GluR2 C-terminus may regulate postsynaptic clustering in neurons and downstream transmission of impulses (Matsuda et al., 1999; Shao et al., 2010).

Table 5.1: Examples of PBM phosphorylation

<table>
<thead>
<tr>
<th>PDZ containing protein</th>
<th>Binding partner</th>
<th>PBM</th>
<th>Kinase</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLG4 and TIP-1</td>
<td>Kir 2.3</td>
<td>RRESRI</td>
<td>PKA</td>
<td>Tanemoto et al., 2002</td>
</tr>
<tr>
<td>DLG4</td>
<td>Stargazin</td>
<td>RRTTPV</td>
<td>PKA/MA PK</td>
<td>Choi et al., 2002</td>
</tr>
<tr>
<td>DLG4 and DLG3</td>
<td>NR2B subunit of NMDAR</td>
<td>SSIESDV</td>
<td>CK2</td>
<td>Chung et al., 2004</td>
</tr>
<tr>
<td>GRIP1</td>
<td>GluR2 subunit of AMPAR</td>
<td>GIESVKI</td>
<td>PKC</td>
<td>Matsuda et al., 1999</td>
</tr>
<tr>
<td>DLG4 and DLG1</td>
<td>NR2B subunit of NMDAR</td>
<td>SLESEV</td>
<td>PKA/PK C</td>
<td>Chung et al., 2004</td>
</tr>
<tr>
<td>DLG4</td>
<td>LRP4</td>
<td>SSESQV</td>
<td>CaMKII</td>
<td>Tian et al., 2006</td>
</tr>
<tr>
<td>GRID2IP</td>
<td>GluR2 subunit of AMPAR</td>
<td>DPDRGTSI</td>
<td>PKA</td>
<td>Sonoda et al., 2006</td>
</tr>
<tr>
<td>DLG1</td>
<td>HTLV-1 Tax-1</td>
<td>SPPGLEPPSEKHFRETEV</td>
<td>CK2</td>
<td>Bidoia et al., 2010</td>
</tr>
</tbody>
</table>

Phosphorylated residue highlighted in bold, PBM underlined.

Conversely, phosphorylation of a residue adjacent to the PBM of the mutated in colorectal cancer (MCC) protein increases the affinity of MCC for the polarity protein SCRIB, indicating that phosphorylation can also promote interactions (Pango et al., 2012). The binding affinity is thought to increase due to the availability of additional interactions between MCC and the SCRIB PDZ domain upon PBM phosphorylation (Pango et al., 2012). It has been suggested that the interaction between SCRIB and MCC regulates the
formation of lamellipodia in migrating cells, and that this is regulated by MCC phosphorylation (Pangon et al., 2012).

In the case of the membrane protein apactin, phosphorylation of the PBM decreases affinity for the second PDZ domain of NHERF1 but increases binding affinity for the first PDZ domain and for fellow PDZ protein PDZK1 (Tandon et al., 2007). PBM phosphorylation also increases apactin association with the cytoskeleton, implying that the switch in interaction promotes the function of apactin in maintaining actin filament bundles at the apical surface (Tandon and De Lisle, 2004). Moreover, it has been suggested that, as apactin and NHERF1 function together in exocytosis, switching the PDZ binding preference allows NHERF1 to act as a scaffold and bring together an alternative combination of proteins via their PBMs (Tandon et al., 2007; Tandon and De Lisle, 2004).

Interestingly, regulatory phosphorylation outside of the PBM can affect PDZ binding, leading to a change in protein structure that affects the availability of the PBM for cognate binding. For example, regions of the pleckstrin homology (PH) domain of Kalirin-7 are heavily phosphorylated in neuronal cells and this is predicted to affect the interaction with a number of PDZ substrates, including DLG4, at the synapse (Kiraly et al., 2011). Furthermore, different patterns of phosphorylation depend upon cellular context; Kalirin-7 sites at the N terminus were phosphorylated in mouse brain by PKA, PKC and CaMKII but phosphorylation was not identified in non-neuronal cells (Kiraly et al., 2011). Thus, phosphorylation of PDZ-binding proteins both within and outside of the PBM can have varying affects upon PBM:PDZ interactions, and is likely to depend upon the cellular context and the proteins in question (Table 5.2).

5.1.2 Viral PBM regulation

Post-translational modifications of HTLV-1 Tax have been suggested to regulate the activity of the viral PBM (Bidoia et al., 2010; Lodewick et al., 2013). HTLV-1 Tax is
phosphorylated within the PBM at threonine 351 by CK2 and is acetylated proximal to the PBM at lysine 346 by p300 (Bidoia et al., 2010; Lodewick et al., 2009). Phosphorylation of the Tax PBM interrupts the proteins ability to interact with DLG1 and disrupts viral activation of the NFκB pathway, whilst acetylation promotes activation of the NFκB pathway (Suzuki et al., 1999; Bidoia et al., 2010). The NFκB family of transcription factors regulate the expression of a number of cellular genes associated with cell growth, apoptosis, embryonic development, immune responses, stress responses and oncogenesis (Pahl, 1999). HTLV-1 mediated activation of these genes is likely to promote viral replication, indeed the PBM is crucial to Tax-induced T-cell proliferation (Xie et al., 2006). Phosphorylation of the Tax PBM may be a mechanism to impair PBM binding when NFκB activation is not required (Bidoia et al., 2010). Furthermore, as Tax binding to MAGI-1 alters the localization of both proteins, interruption of this by phosphorylation may act to regulate Tax location and thus control the interactions that the viral protein is able to make spatially during stages of viral replication (Ohashi et al., 2004). However, the functional consequences of Tax PBM phosphorylation during HTLV-1 replication have yet to be fully characterized.

Crystal structures have revealed that, as well as the four residue class I PBM, residues upstream are also involved in recognition by DLG1 and MAGI-1 PDZ domain recognition (Zhang et al., 2007; Charbonnier et al., 2011; Mischo et al., 2013). Moreover, a peptide-binding loop of the PDZ domain specifically interacts with an arginine residue at the -5 position the HPV18 E6 peptide which, when mutated, was demonstrated to be crucial in HPV-mediated degradation of DLG1 and MAGI-1 (Zhang et al., 2007). The presence of the binding loop and the interaction with six of the E6 PBM residues may explain the specificity of high-risk HPV E6 proteins in their ability to target cellular PDZ domain containing proteins (Zhang et al., 2007; Charbonnier et al., 2011).
As phosphorylation is an established mechanism by which binding to PBMs is controlled, cellular kinases may also regulate the E6 PDZ binding function during the HPV life cycle. As with cellular proteins, phosphorylation within the PBM of E6 has been predicted to affect binding activity. Recognition sites of the serine/threonine kinases protein kinase A (PKA) and AKT have been identified proximal to the PBM (Figure 5.1). Structural analysis of the E6 PBM bound to PDZ proteins DLG1 and MAGI-1 indicates that a phosphate group would not fit in the interacting groove and would therefore disrupt E6-PBM binding to cellular PDZ proteins (Zhang et al., 2007; Charbonnier et al., 2011).

In vitro kinase assays have shown that the HPV18 E6 PBM is phosphorylated by PKA and studies have identified phosphorylated E6 in human epithelial cells (Kühne et al., 2000). Furthermore, stimulation of PKA activity inhibits the degradation of DLG1 induced by HPV18 E6, indicating that phosphorylation inhibits E6-PDZ interactions (Kühne et al., 2000). Moreover, mutating the PKA recognition site proximal to the PBM, by substituting an arginine residue for a leucine residue at position 153, removes the sensitivity of HPV-mediated DLG1 degradation to PKA activity, indicating that regulation of PBM activity is removed once the PKA recognition site is mutated (Kühne et al., 2000). It is suggested that this substitution enables E6 to degrade DLG1, and other PDZ targets constitutively, regardless of PKA activity (Kühne et al., 2000). Thus phosphorylation of the E6 PBM is able affect E6 PDZ interactions.

Different HPV E6 proteins exhibit different susceptibilities to PKA phosphorylation; HPV18 E6 shows a higher susceptibility to PKA phosphorylation than HPV16, and HPV31 was shown to be a weak PKA substrate (Kühne et al., 2000; Boon et al., 2014). The kinase AKT (also known as protein kinase B) is another candidate for recognizing and phosphorylating high-risk HPV E6 PBMs (Figure 5.1). Conservation of a kinase recognition site in high-risk HPV types suggests that phosphorylation of the PBM is
Group 1
HPV 16
CMSCC RSRRRETQTV
HPV 18
CHSCCNRARQE RLORRETQTV
HPV 31
CIVCW RRPRRETQTV
HPV 33
CAACW RS RRRETAL
HPV 35
CMSCW KPTRETREV
HPV 39
CRRCWTTKRED RRLRETQTV
HPV 45
CNTCCDQRARQE RRREEETQTV
HPV 51
CANCW Q RTRQNATQOV
HPV 52
CSECW RPRPVETQOV
HPV 56
CLGCW RQ TSPRETVSTV
HPV 58
CAVCW RPPRETQOV
HPV 59
CRGCRTRARHLRQQOARSETV

Group 2A
HPV 68
CRHCWTSKRED RRRTQETQOV

Group 2B
HPV 26
CTNCW RRREETQTV
HPV 30
CLQCW RHTSTETAV
HPV 34
CRQCW RPSATTVV
HPV 53
CLTCW RHTTATESAV
HPV 66
CLQCW RHTSRQATETSTV
HPV 67
CSVCW RPRPQTQTOV
HPV 69
CTNCW RPPREATETQOV
HPV 70
CRHCWTSNRED RRRTRETOQTV
HPV 73
CTRQCW RPSATVV
HPV 82
CANCRTCAA RQRSETQOV
HPV 85
CRRCMTRA QGQQSRRETQOV
HPV 97
CNSCNQSRQE RLKRRETQOV

Group 3
HPV 6
CLHCWTTCMED MLP
HPV 11
CLHCWTTCMED LLP

Key:

**PDZ binding motif (PBM)**

Bold and arrow indicates residue targeted

**PKA recognition sequence** (R-R-S/T-Y)

**Akt recognition motif** (R-X-R-X-S/T)

Consensus PBM: \( R^S_X^S_X^X_Y \)
Figure 5.1: Alignment of E6 C-terminal sequences
Alignment of the C terminus of the E6 protein from HPV types belonging to IARC Groups 1, 2A, 2B and 3. PBM shaded in grey, PKA recognition site boxed and AKT recognition site underlined, target residue for phosphorylation indicated in bold.

The consensus PBM sequence is shown. Number above PBM indicates residue nomenclature, numbered in reverse from the C-terminal amino acid.
important during the viral life cycle, and that it may be an important feature in viral oncogenesis.

Table 5.2: Consequences of cellular PBM Phosphorylation

<table>
<thead>
<tr>
<th>PBM-containing protein</th>
<th>Protein function</th>
<th>Consequences of phosphorylation</th>
<th>Kinase</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCC protein</td>
<td>Cell cycle regulator</td>
<td>Increase binding affinity with SCRIB</td>
<td>PAK1</td>
<td>Pangon et al., 2012</td>
</tr>
<tr>
<td>GluR2</td>
<td>Glutamate receptor subunit</td>
<td>Disrupted interaction with PICK1, Disrupted binding to glutamate receptor interacting protein</td>
<td>CK2</td>
<td>Shao et al., 2010; Matsuda et al., 1999</td>
</tr>
<tr>
<td>Kir2.3</td>
<td>Potassium ion channel</td>
<td>Interruption of colocalisation with DLG4 at postsynaptic membrane, TIP-1 interaction through its PDZ domain is abolished</td>
<td>PKA</td>
<td>Yan et al., 2009a; Cohen et al., 1996</td>
</tr>
<tr>
<td>MMAC1/PTEN</td>
<td>Membrane tethered phosphatase</td>
<td>Promotes binding with DLG1 and hMAST205</td>
<td>Unknown</td>
<td>Adey et al., 2000</td>
</tr>
<tr>
<td>Ephrin B</td>
<td>Receptor tyrosine kinase</td>
<td>Syntenin binding hindered, PTPN13 binding unchanged, RGS3 binding unchanged, Par6 interaction disrupted</td>
<td>PKA</td>
<td>Lu et al., 2001; Lee et al., 2008</td>
</tr>
<tr>
<td>Jagged-1</td>
<td>Membrane bound ligand</td>
<td>Binding to AF-6 disrupted</td>
<td>Unknown</td>
<td>Popovic et al., 2012; Chetkovich et al., 2002; Stein and Chetkovich, 2010</td>
</tr>
<tr>
<td>Stargazin</td>
<td>Calcium ion channel</td>
<td>Binding to DLG4 inhibited</td>
<td>PKA</td>
<td>Chetkovich:2002 wg, Stein:2010gb</td>
</tr>
<tr>
<td>NR2B</td>
<td>Subunit of NMDA receptor</td>
<td>Affects binding to DLG4 and DLG2</td>
<td>CaMKII</td>
<td>Chung et al., 2004</td>
</tr>
<tr>
<td>Apactin</td>
<td>Membrane glycoprotein</td>
<td>Decreased affinity for NHERF1 PDZ2, increased affinity for PDZK1 and NHERF1 PDZ1</td>
<td>Unknown</td>
<td>Tandon et al., 2007</td>
</tr>
<tr>
<td>Kalirin-7</td>
<td>RhoGEF kinase</td>
<td>Decreased affinity for DLG4 suggested</td>
<td>Fyn, PKA, PKC, CK2, CaMKII</td>
<td>Kiraly et al., 2011</td>
</tr>
<tr>
<td>β2-adrenergic receptor</td>
<td>GPCR</td>
<td>NHERF1 interaction disrupted, leading to mis-sorting of receptor</td>
<td>GRK5</td>
<td>Cao et al., 1999</td>
</tr>
<tr>
<td>LRP4</td>
<td>Wnt signalling regulator</td>
<td>Suppressed its interaction with DLG1 and DLG4</td>
<td>CaMKII</td>
<td>Tian et al., 2006</td>
</tr>
<tr>
<td>Syndecan-1</td>
<td>GPCR coreceptor</td>
<td>Interaction with syntenin-1 promoted by dephosphorylation</td>
<td>Unknown</td>
<td>Sulka et al., 2009</td>
</tr>
<tr>
<td>HTLV-1 Tax</td>
<td>Viral oncoprotein</td>
<td>Impair binding to DLG1, affects NFkB pathway interaction</td>
<td>CK2</td>
<td>Bidoia et al., 2010</td>
</tr>
</tbody>
</table>
As well as interrupting PDZ interactions, phosphorylation of the E6 PBM enables alternative protein interactions (Figure 5.2). Phosphorylation of exogenously expressed HPV18 E6 by PKA promotes an interaction with the adapter protein 14-3-3ζ (Boon and Banks, 2012a; Boon et al., 2014). This suggests that phosphorylation of the PBM can act as a switch, changing the binding profile of the PBM. The same study demonstrated that binding of 14-3-3ζ increased the stability of the viral oncoprotein (Boon and Banks, 2012b).

Thus, the PBMs of high-risk α-HPV E6 proteins are highly likely to be phosphorylated during the viral life cycle. However, the affect of PBM phosphorylation upon E6 interactions during the viral life cycle is largely unknown.
Figure 5.2: The HPV18 E6 PDZ binding motif

(A) The amino acid sequence of the HPV18 E6
(B) The HPV18 E6 PBM sequence and current binding models: A number of E6 oncoproteins have been shown to bind cellular PDZ proteins via a C-terminal PBM, leading to their proteasome mediated degradation. The PBM overlaps with a PKA recognition site, phosphorylation of which sterically hinders PDZ protein binding. Boon et al. have shown that the phosphorylated PBM can bind to 14-3-3ζ.
5.3 Hypothesis and objectives

Studies to date have shown that the PBM of HPV18 E6 is phosphorylated by PKA, and that this alters the interactions that E6 is able to form with cellular factors. The experiments described in this chapter aimed to study phosphorylation of the PBM within a model of the HPV18 life cycle: Specific objectives were to:

i. Identify kinase recognition sites overlapping the PBMs of high-risk α-HPV E6 oncoproteins

ii. Determine whether the HPV18 E6-PBM is phosphorylated by PKA in the HPV18 life cycle model

iii. Determine whether phosphorylation affects HPV18 E6 protein levels in viral genome containing cells

iv. Determine whether E6 PBM phosphorylation affects PBM-PDZ interactions during the HPV life cycle.
5.4 Results

5.4.1 Computational screening predicts phosphorylation of α-HPV E6 PBMs

To date, twelve HPV types have been defined by the International Agency for Research on Cancer (IARC), a department of the WHO, as ‘high-risk’ with respect to causing cancer (group 1), and a further ten types are recognised as ‘probably and ‘possibly’ cancer-causing (group 2A and 2B, Table 1.3, Bouvard et al., 2009). To initially assess the susceptibility of high-risk HPV E6 proteins to phosphorylation, a Scansite motif screen was performed on the amino acid sequences of HPV E6 proteins from the IARC groups 1, 2A and 2B (Figure 5.3). The screen assessed the likelihood of peptide phosphorylation to produce a probability score within the range of 0 to 1. All of the group 1 HPV types are predicted in this screen to be targeted by a kinase within, or proximal to the PBM, with a probability score of 0.49 or more (Figure 5.3). Furthermore, the cellular kinases predicted to perform this phosphorylation included the AGC kinase family members PKA, AKT, protein kinase C (PKC) and AMPK (AMP-activated protein kinase) (Figure 5.3). Both PKA and AKT were predicted to phosphorylate the HPV16 PBM with scores of 0.49 and 0.52, whereas the PBM of HPV18 E6 was a predicted target of PKA alone, with a score of 0.46 (Figure 5.3). Phosphorylation of HPV16 by AKT produced a higher probability score than for PKA (0.53 cf. 0.49). During this study, it was confirmed \textit{in vitro} that AKT is able to phosphorylate purified HPV16 E6 at the PBM and that PKA is able to target the PBM of both HPV16 and HPV18 E6 proteins (Boon \textit{et al.}, 2014; Boon and Banks, 2012b). Boon \textit{et al.} confirmed that α-HPV E6 proteins are susceptible to phosphorylation at the PBM and further investigations in this study focussed on this in the context of whole genome replication in primary cells. In contrast to the group 1 viruses, not all members of group 2
<table>
<thead>
<tr>
<th>Virus</th>
<th>Kinase</th>
<th>Scansite Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV 16</td>
<td>Protein Kinase A</td>
<td>0.4969</td>
</tr>
<tr>
<td></td>
<td>Akt Kinase</td>
<td>0.5272</td>
</tr>
<tr>
<td>HPV 18</td>
<td>Protein Kinase A</td>
<td>0.4626</td>
</tr>
<tr>
<td>HPV 31</td>
<td>Akt Kinase</td>
<td>0.5565</td>
</tr>
<tr>
<td>HPV 33</td>
<td>Protein Kinase A</td>
<td>0.4939</td>
</tr>
<tr>
<td>HPV 35</td>
<td>Protein Kinase A</td>
<td>0.5472</td>
</tr>
<tr>
<td>HPV 39</td>
<td>Protein Kinase A</td>
<td>0.5054</td>
</tr>
<tr>
<td></td>
<td>Protein Kinase C</td>
<td>0.5663</td>
</tr>
<tr>
<td>HPV 45</td>
<td>Protein Kinase A</td>
<td>0.4207</td>
</tr>
<tr>
<td></td>
<td>Akt Kinase</td>
<td>0.5451</td>
</tr>
<tr>
<td>HPV 51</td>
<td>Akt Kinase</td>
<td>0.6631</td>
</tr>
<tr>
<td>HPV 52</td>
<td>Akt Kinase</td>
<td>0.6997</td>
</tr>
<tr>
<td>HPV 56</td>
<td>Akt Kinase</td>
<td>0.5733</td>
</tr>
<tr>
<td>HPV 58</td>
<td>Protein Kinase</td>
<td>0.4892</td>
</tr>
<tr>
<td>HPV 59</td>
<td>Akt Kinase</td>
<td>0.616</td>
</tr>
<tr>
<td>HPV 68</td>
<td>Akt Kinase</td>
<td>0.6313</td>
</tr>
<tr>
<td>HPV 26</td>
<td>Protein Kinase A</td>
<td>0.6086</td>
</tr>
<tr>
<td></td>
<td>Akt Kinase</td>
<td>0.6564</td>
</tr>
<tr>
<td></td>
<td>AMP Kinase</td>
<td>0.661</td>
</tr>
<tr>
<td>HPV 34</td>
<td>AMP kinase</td>
<td>0.7454</td>
</tr>
<tr>
<td>HPV 66</td>
<td>PKC</td>
<td>0.4996</td>
</tr>
<tr>
<td>HPV 67</td>
<td>AMP kinase</td>
<td>0.6841</td>
</tr>
<tr>
<td>HPV 69</td>
<td>AMP kinase</td>
<td>0.7299</td>
</tr>
<tr>
<td>HPV 70</td>
<td>PKA</td>
<td>0.4931</td>
</tr>
<tr>
<td></td>
<td>Akt</td>
<td>0.62</td>
</tr>
<tr>
<td>HPV 82</td>
<td>Akt</td>
<td>0.5272</td>
</tr>
<tr>
<td></td>
<td>AMP kinase</td>
<td>0.6779</td>
</tr>
<tr>
<td>HPV 85</td>
<td>Protein Kinase A</td>
<td>0.5563</td>
</tr>
</tbody>
</table>

**Figure 5.3: Computational screen of E6 C-terminal sequences**

A screen of the peptide sequence of E6 proteins from HPV types classified by the IARC as cancer causing (Group 1) and probably (2A) and possibly (2B) cancer causing was carried out using Scansite software [http://scansite.mit.edu/](http://scansite.mit.edu/). The probability score for each kinase targeting the E6 C-terminus is indicated by green bar and number.
had predicted kinase sites within the C-terminus of E6. These included HPV types 53, 73 and 97 (Figure 5.3).

5.4.2 Investigation into the effect of PBM phosphorylation upon E6

5.4.2.1 E6 is phosphorylated in primary keratinocytes harbouring HPV genomes

In order to assess whether phosphorylation of the PBM occurs during the HPV life cycle, primary keratinocytes containing HPV18 genomes were grown in monolayer and treated with the PKA activator forskolin (FK), or with vehicle (DMSO) for 6 hours. The HPV18 DNA positive HeLa cells were included in the assay since the E6:PDZ protein interactions in theses cells can be modulated by PKA activation (Kühne et al., 2000). Total protein was extracted from the cells in urea buffer and then subjected to western blot analysis using an antibody that recognises the phosphorylated form of the HPV18 E6 PBM (E6P, a kind gift from L. Banks). Under conditions of PKA activation, phospho E6 was detected in both HeLa and genome containing cells; the levels of phosphorylated E6 (E6P) detected were higher in cells exposed to PKA activator, whereas it was barely detectable in vehicle-treated cells (Figure 5.4 A). Furthermore, when the same treatment was applied to wild-type genome containing cells alongside cells containing ΔPKA genomes (the mutation abrogates the PKA recognition site), the levels of phosphorylated E6 (E6P) were again higher in cells harbouring the wild type genomes exposed to PKA activator (Figure 5.4 B). In this analysis, there was no significant detection of the mutated protein (ΔPKA) (Figure 5.4 B). Untransfected HFKs were included to show that neither the E6 nor the E6P antibodies detected proteins in these cells (Figure 5.4 B). An increase in E6 levels in HeLa and WT genome containing cells was observed in many, but not all studies, where cells were treated with PKA activator (Figure 5.4 A). This suggested an effect of the kinase upon E6 protein stability.
Figure 5.4: Phosphorylation of E6 in HPV18 positive cells

(A) HeLa and HPV18 genome containing keratinocytes, grown in monolayer, were treated with the PKA activator forskolin (FK,+) or the vehicle DMSO (-) in serum-free media. After 6 hours, protein was harvested and separated by SDS-PAGE. Membranes were blotted with mouse anti-HPV18 E6 (Santa-Cruz) or rabbit anti-HPV18 E6P, alongside β-actin as a loading control. Molecular weight markers indicated in kDa.

(B) From Delury et al. 2013. HPV18 genome containing keratinocytes were grown in monolayer alongside ΔPKA genome containing cells and were treated with the PKA activator forskolin (FK) or the vehicle DMSO (-) in serum-free media. After 6 hours, protein was harvested and separated by SDS-PAGE. Membranes were blotted with mouse anti-HPV18 E6 (Santa-Cruz) or rabbit anti-HPV18 E6P, alongside β-actin as a loading control. HFK was included as a negative control.

Numbers indicate fold change in E6P relative to E6, following normalisation of each to the respective loading control.
In order to assess whether specific cellular pools of E6 were phosphorylated by PKA, HPV18 genome-containing cells were grown in monolayer culture and subjected to subcellular fractionation, alongside the HPV18 positive cancer cell line HeLa. Housekeeping gene analysis confirmed enrichment of four subcellular compartments; cytoplasmic (F1), membrane (F2), nuclear (F3) and cytoskeletal (F4). As E-cadherin was not detectable in HeLa cells, ZO-1 was included to demonstrate enrichment of the membrane fraction in these cells (Figure 5.5 A).

The E6 oncoprotein was predominantly nuclear with a small amount of protein detected in the cytoplasmic, membrane and cytoskeletal fractions (Figure 5.5 B). Immunoblotting with the phospho-specific E6 PBM antibody detected protein primarily in the nuclear fraction of the HPV18 containing cells. Similarly, in HeLa cells the nuclear fraction of E6 was detected with the phospho E6 antibody, and this was consistent across three experiments (Figure 5.5 B and C).

Whilst the profile of E6 distribution is similar between genome containing cells and HeLa cells, where the viral DNA is integrated into the host genomes, a very small increase in E6 present in the cytoplasm was observed in the cancer cells (Figure 5.5). This modest change did not appear to relate to phosphorylation of E6, but this may be unclear due to the sensitivity of the E6P antibody. To establish whether localisation of the other viral oncoprotein E7 was altered upon disease progression, the fractionation lysates were immunoblotted with an anti-E7 antibody. The HPV18 protein E7 was detected in all subcellular fractions of primary keratinocytes harbouring the HPV18 genomes (Figure 5.5). E7 was detected primarily in the membrane fraction, with less protein observed in the cytoplasm and nucleus and a small amount in cytoskeletal fraction (Figure 5.5 B). In HeLa cells, E7 was predominantly detected in the cytoplasm, with small amounts of the protein...
A

<table>
<thead>
<tr>
<th></th>
<th>Donor 1</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>F4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Donor 2</td>
<td>F1</td>
<td>F2</td>
<td>F3</td>
<td>F4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HeLa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GRB2
E-cadherin
ZO-1
p53
K18

Proportion of protein
F1, Cytoplasm
F2, Membrane
F3, Nucleus
F4, Cytoskeleton

Donor 1
Donor 2
HeLa

GRB2
Ecadherin
p53
K18
Figure 5.5: Distribution of phosphorylated E6 in monolayer-grown cells

HeLa and HPV18 genome containing keratinocytes, form two donor backgrounds, were grown in monolayer to approximately 70% confluency before subjection to subcellular fractionation. Western blots were quantified by densitometry and are representative of three repeats.

(A) Housekeeping gene analysis indicated sufficient enrichment of the cytoplasmic, membrane, nuclear and cytoskeletal fractions (F1-F4), as described in Section 3.

(B) HPV protein distribution was analysed by subcellular fractionation followed by western blot using mouse anti-HPV18 E6 (Santa Cruz), rabbit anti-HPV18 E6P or mouse anti-HPV18 E7.

(C) Three subcellular fractionation repeats with detection of E6P.
5.4.2.2 E6 protein levels are affected by PKA activity

In some experiments, treatment of the HPV18 genome containing cells with a PKA activator led to an increase in E6 protein suggesting that PKA mediated phosphorylation of the E6 PBM regulates E6 stability (Figure 5.4 A). Indeed, it has been suggested by knockdown of 14-3-3ζ expression in HeLa cells that the association between the viral oncoprotein and the cellular adapter protein regulates E6 protein stability (Boon and Banks, 2012b). Thus, to investigate the relationship between PKA activity and E6 protein stability in the viral genome containing cells, monolayer-grown cells were treated with the PKA inhibitor dihydrochloride (H89). The cells were harvested after 4 and 6 hours incubation with H89 and E6 protein levels assessed by western blot. Following both 4 and 6 hours of treatment with H-89 led to a marked decrease in E6 protein in two different donor backgrounds (Figure 5.6).

Because the inhibitor H89 is not a selective inhibitor of PKA but has been shown to inhibit at least 8 other kinases, including AKT, ribosomal S6 kinase (RSK) and Rho-associated protein kinase II (ROCKII, Lochner and Moolman, 2006), the effect of PKA on E6 stability was also examined by transfecting the cells with a plasmid expressing the PKA inhibitor peptide (mPKI). This plasmid (a kind gift from M. Diskar) expresses a GFP fusion of protein kinase inhibitor peptide (PKI), which has been identified as an endogenous specific inhibitor of PKA (Dalton and Dewey, 2006). Transfection efficiency was estimated via observation of the GFP reporter, typically 50-60% in HeLa and 30-40% in HFKs following incubation for 48 hours. In genome containing cells and in Hela cells transfected with the specific inhibitor, the levels of the E6 protein were decreased in the membrane and in the nucleus (Figure 5.5 B). These data indicate that the subcellular localization of both oncoproteins changes upon disease progression.
Figure 5.6: The effect of PKA inhibition upon HPV18 E6 protein

Keratinocytes containing HPV18 genomes were treated with the PKA inhibitors H89 or PKI for 4 and 6 hours. Control cells were treated with the vehicle DMSO alongside H89 or 1% BSA alongside PKI. Protein extracted in urea buffer from cells derived from 2 donor backgrounds. Western blot analysis was applied using antibodies specific to HPV18 E6 protein and β–actin loading control. Representative of two repeats.
compared to the untransfected cells (Figure 5.6). Together these data suggest that PKA phosphorylation regulates E6 steady state stability.

The next step aimed to examine whether the observed change in E6 levels reflected a change in a specific cellular pool of E6. Firstly, HeLa cells were transfected with the mPKI plasmid, incubated for 48 hours and then equal numbers of treated and untreated cells were subjected to subcellular fractionation. Firstly, housekeeping gene analysis confirmed that the enrichment of the cytoplasmic, membrane, nuclear and cytoskeletal compartments was successful (Figure 5.8). As noted previously (Figure 5.4) HPV18 E6 was present in all four fractions analysed and was detected predominantly in the nucleus of HeLa cells (Figure 5.7). Where PKA was inhibited (+mPKI), less E6 was detected in the cytoplasmic fraction along with a subtle shift from the membrane to the nucleus compared to untransfected cells (-mPKI) (Figure 5.7). The distribution of the phosphorylated form remained nuclear and cytoskeletal in the presence and absence of mPKI, although less E6P was detected where PKI was expressed, indicating that phosphorylation of E6 was inhibited (Figure 5.8). As an equal number of cells were subjected to fractionation and less total E6 was detected where PKA was inhibited, this supports the earlier data where PKA phosphorylation stabilised E6 protein levels (Figure 5.6).

To determine whether these changes in distribution and abundance were specific to E6, the distribution of the viral protein E7 was included in the analysis and was shown to reside predominantly in the cytoplasmic and membrane fractions, with lower levels detected in the nuclear and cytoskeletal fractions, irrespective of mPKI expression (Figure 5.7).

Next, the sensitivity of E6 localization to changes in PKA signalling was examined in HPV18 genome-containing cells. Keratinocytes were transfected with the GFP-mPKI expressing plasmid and fractionated as described. Detection of GRB2, E-cadherin, p53 and K18 confirmed successful enrichment of the subcellular compartments (Figure 5.8 A). As
Figure 5.7: HPV oncoprotein expression in HeLa expressing mPKI
HeLa were transfected with plasmid expressing peptide inhibitor of PKA (+PKI) or vehicle (-PKI), quenched with serum-containing media after 8 hours and protein harvested subcellular fractionation after a further 40 hours. Left-hand key describes each fraction number corresponding to a subcellular compartment.

(A) Lysates probed for GRB2, ZO-1, p53 and K18 to show enrichment of the cytoplasmic, membrane, nuclear and cytoskeletal fractions, respectively.

(B) Fractions probed for analysis of HPV protein distribution.

Densitometry was performed on scanned film using ImageJ. Representative of three repeats.
Figure 5.8: HPV oncoprotein expression in primary cells expressing mPKI
HPV18 genome containing keratinocytes derived from two donor backgrounds were treated with transfection vehicle (-mPKI) or plasmid expressing peptide inhibitor of PKA (+mPKI). Following 48 hours, cytoplasmic (F1), membrane (F2), nuclear (F3) and cytoskeletal (F4) fractions were sequentially extracted and subject to western blot analysis. The top panel (A) shows an example of housekeeping gene analysis to confirm enrichment of each fraction. The lower panels (B and C, on next page) shows detection of HPV18 E6, E6P and E7 in lysates. Densitometry analysis perform using ImageJ software. Representative of 2 repeats per donor background.
in HeLa cells, E6 was predominantly localized to the nucleus of keratinocytes, with the remaining protein distributed across the other three cellular compartments (Figure 5.8 B and C). This fractionation also agrees with previous detection of E6 in fractions of genome containing cells (Figure 5.5). In cells derived from two donor backgrounds, no major change in E6 distribution was observed where PKA was inhibited (Figure 5.8 B and C). Phospho-E6 was primarily detected in the nucleus of HPV18 positive keratinocytes, with some protein detected in the cytoskeletal fraction of keratinocytes, and this was not significantly affected by the expression of mPKI (Figure 5.8 B and C). The E7 protein remained within the cytoplasmic and membrane fractions in the presence and absence of mPKI (Figure 5.8). Together these data suggest that phosphorylation by PKA does not alter the stability of a specific subcellular pool of viral oncoproteins.

5.4.2.3 E6 levels are affected by PDZ binding activity

If phosphorylation of the E6 PBM mediates interactions with cellular factors that stabilize E6, then mutations of the E6 domain that prevent phosphorylation would be expected to affect E6 protein levels. To test this hypothesis, E6 protein levels were examined in keratinocytes harbouring the HPV18 mutant genomes that either lack the E6 PBM (ΔPDZ), or genomes that express a mutant form of E6 that is not sensitive to PKA phosphorylation (ΔPKA). Interestingly, comparison of levels of the E6 oncoprotein to wild type genome containing cells showed decreased levels of E6 where the PBM was removed and this was consistent between two donors. However, there was no consistent change in E6 levels in the presence of the ΔPKA genomes; in one donor E6 protein levels were comparable to WT genome containing cells and in the second donor E6 levels were decreased (Figure 5.9). This indicates that the PDZ binding activity of E6 is influential in the protein’s stability. Moreover, the alteration in viral protein levels was not restricted to
Figure 5.9: E6 protein in HPV18 genome containing HFKs

Protein was extracted from HFKs containing WT HPV18, ΔPDZ or ΔPKA genomes, from two donor backgrounds, and subject to western blot analysis to assess HPV18 E6 and E7 oncoprotein levels. Untransfected keratinocytes (HFK) included as a control. Densitometry performing in ImageJ, values first normalised to β-actin and then to levels of protein in WT genome containing cells to calculate a relative fold change. Representative of two repeats.
E6, analysis of E7 protein levels also revealed a decrease in protein levels in cells containing ΔPDZ genomes (Figure 5.9). This decrease was sustained in the presence of ΔPKA genomes, albeit to a larger extent in donor 1 (0.5-fold) than donor 2 (0.7-fold, Figure 5.9).

5.4.3 14-3-3ζ

Phosphorylation of the HPV18 PBM allows an alternative interaction with 14-3-3ζ (Boon and Banks, 2012b). To investigate whether interaction with the viral protein alters 14-3-3ζ levels and/or localization in the life cycle, the overall abundance of the scaffold protein was assessed in primary keratinocytes containing high-risk α-HPV types 16 and 18 and in the cervical cancer cell line HeLa. Although 14-3-3ζ levels decreased in HeLa compared to untransfected HFKS, there was no change in abundance of the scaffold protein where intact α-HPV genomes were present (Figure 5.10 A). Furthermore, perturbing the activity of the HPV18 E6 PBM activity had no effect upon the abundance of 14-3-3ζ (Figure 5.10 B).

To determine if PKA activity was inhibited via the expression of mPKI in HeLa or HPV18 genome containing cells. There was no effect upon the distribution of the scaffold protein in monolayer cells, where the majority of the protein was in the cytoplasmic and membrane fractions, with some protein detected in the nucleus and a little at the cytoskeleton (Figure 5.10 C).

The distribution of 14-3-3ζ was further investigated in differentiated keratinocytes by immunofluorescence of cells stratified in raft culture. Together, these data indicate that the expression levels or the localization of cellular factor is not affected by changes in E6 PBM activity.
### Figure A

<table>
<thead>
<tr>
<th></th>
<th>Donor 1</th>
<th>Donor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HeLa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Western Blot:**
- 14-3-3
- β-Actin

**Relative fold change**

### Figure B

<table>
<thead>
<tr>
<th></th>
<th>Donor 2</th>
<th>Donor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔPDZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔPKA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HFK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔPDZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔPKA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Western Blot:**
- 14-3-3
- β-Actin

**Relative fold change**
Figure 5.10: 14-3-3ζ protein expression in HPV genome containing keratinocytes

(A) Total protein was extracted from primary keratinocytes and from HPV16 or HPV18 genome containing cells from two donor backgrounds and subject to western blot analysis. Representative of two repeats. Fold change calculated relative to levels in HFKs of the same donor background, following normalisation to β-actin loading control.

(B) Total protein was extracted from HFKs and from HPV18 WT, ΔPDZ and ΔPKA genome containing cells from two donor backgrounds and subject to western blot analysis. Representative of two repeats. Fold change calculated relative to levels in WT genome containing cells of the same donor background, following normalisation to β-actin loading control.

(C) Analysis of 14-3-3ζ distribution in monolayer-grown HeLa cells and keratinocytes containing HPV18 genomes, from two donor backgrounds. Cells were transfected with vehicle (-mPKI) or mPKI plasmid and subjected to subcellular fractionation. Representative two repeats. Densitometry analysis performed using ImageJ software.
5.5 Discussion

This study showed for the first time that the E6 PBM is phosphorylated during the life cycle of the high-risk virus HPV18. The data suggests that phosphorylation occurs principally to the nuclear pool of E6. The nuclear localization of E6 observed in monolayer grown keratinocytes agrees with established knowledge; three nuclear localization signals have been identified in HPV16 E6 and exogenous expression of the oncoprotein in cervical cells leads to a predominantly nuclear subcellular location (Masson et al., 2003; Tao et al., 2003; Liang et al., 1993). However, as the E6P antibody detected the phospho form weakly, it is possible that the phosphorylated protein was only detected in nuclear fractions because this is where the majority of the E6 protein resides. If E6 phosphorylation occurs mainly to protein in the nucleus, then activation of PKA could be expected to lead to an increase in nuclear DLG1, and inhibition of PKA to a decrease of the cellular protein. Due to time constraints, this study was unable to investigate the distribution of DLG1, and other PDZ targets, during the expression of mPKI, and so this requires further characterisation, perhaps with additional investigation including small molecule inhibitors and activators of PKA.

Activation of PKA using forskolin led to an increase in E6 abundance and correspondingly inhibition of PKA by both specific and less-specific inhibitors led to a decrease in oncoprotein levels. Although no change in a specific pool of phosphorylated E6 was observed upon PKA inhibition of keratinocytes, there may be an alteration in E6 distribution in mPKI treated cancer cells. The differences observed between cancer cells and primary cells may be attributable to differences in mPKI transfection efficiency, which could be overcome using small molecule inhibitors in combination with the subcellular distribution analysis.
It has been demonstrated that E6 binding to the PDZ protein SCRIB stabilizes the oncoprotein (Nicolaides *et al.*, 2011). Phosphorylation of the PBM is likely to disrupt this interaction and thus destabilize the protein. However, since phosphorylation presents novel interactions for E6, it is feasible that this compensates for any interruption in the stability provided by PDZ interactions (Figure 1.12). In this study, lower levels of E6 were observed where the PBM was removed, indicating that an interaction 14-3-3ζ may stabilise E6 in primary cells. However, if this were the case E6 levels where the PKA recognition site is altered would exhibit the same decrease in protein levels, and this was not consistently observed. As the ΔPKA mutation allows constitutive binding to PDZ proteins and interaction between E6 and SCRIB has been shown to promote the stability of both proteins (Nicolaides *et al.*, 2011), perhaps the inability of the ΔPDZ mutant to interact with either PDZ proteins or 14-3-3ζ has negative consequences on protein stability (Figure 5.11).

It should also be noted that other sites within the E6 protein may be recognised by PKA, for example HPV18 threonine 86 was identified in the Scansite screen as a potential target for basophilic kinases (data not shown), phosphorylation of which may have an effect upon E6 stability. In addition deletion of the E6 PBM was associated with a moderate decrease E7 oncoprotein levels (Figure 5.9) and so an effect of the mutation upon E6E7 mRNA transcription cannot be ruled out completely.

It may also be that binding to either PDZs or 14-3-3ζ functions to control the stability of E6 at different times during the viral life cycle. The aforementioned analysis of the effect of phosphorylation on PDZ protein expression levels and distribution could aid this investigation. Further work could also include the analysis of PDZ protein levels and/or 14-3-3 isoforms in differentiated keratinocytes, perhaps in the presence of PKA inhibitors.
Figure 5.11: The consequences of E6 phosphorylation

A number of E6 oncoproteins have been shown to bind cellular PDZ proteins via a C-terminal PBM.

(A) Work by Nicolaides et al. has shown that the interaction with SCRIB stabilises the viral protein.

(B) Phosphorylation of the PBM sterically hinders PDZ protein binding and Boon et al. have shown that the ability of 14-3-3ζ to bind the phosphorylated PBM promotes the stability of HPV18 E6.

(C) Removal of the PBM removes the ability of E6 to interact with both cellular PDZ proteins and 14-3-3ζ, and may therefore be the cause of the observed E6 destabilisation.
or activators. In this study, investigation into the degree of E6 phosphorylation in differentiated cells was investigated by applying western blot analysis to protein extracted from keratinocytes grown in organotypic raft culture. However, the results from multiple experiments were inconsistent and thus inconclusive. Due to this, and the challenging nature of immuno staining for E6, calcium-induced and/or methylcellulose differentiation may present a more practical option to observe E6 interactions during productive viral life cycle.

It seems that the steady state stability of E6 is affected by PKA phosphorylation. This may be due to the increased ability to interact with 14-3-3ζ, which has been previously shown to stabilise E6 (Boon and Banks, 2012b). To examine this further and identify whether an interaction between E6 and 14-3-3ζ occurs in primary keratinocytes, attempts were made to co-immunoprecipitate E6 and 14-3-3ζ in cells were treated with vehicle (DMSO) or with PKA activators. These were unsuccessful, but with further optimisation pose an interesting avenue of investigation.

As E6 exists in more than one subcellular compartment, it is likely that the different pools perform different functions during viral replication. Phosphorylation of E6 may affect pools of protein differently, depending on the proteins available for interactions. For example, the interaction with 14-3-3ζ has been shown to stabilise E6 protein and this may be specific to E6 located at the membrane (Boon and Banks, 2012b). Such phosphorylation of the PBM may promote 14-3-3ζ binding to the PBM and thus organise E6 with other junctional proteins at the nuclear membrane. Equally, in the nucleus, PBM phosphorylation may prevent binding to cellular PDZ proteins where (or when) the interaction is not required, thereby allowing spatial and temporal regulation of E6 PBM
interactions. This speculation requires further investigation, perhaps in stratified cells for better observation of E6 distribution during differentiation.

Work by Delury et al. showed that loss of the PKA recognition site proximal to the E6 PBM was associated with an increased cell growth rate, whilst WT genome containing cells treated with PKA activators exhibited a slowed growth rate. Here, no difference was observed in E6 levels between ΔPKA and HPV18 WT genome containing cells, which may indicate that the difference in growth rate phenotypes are due to the regulation of E6 interactions. The absence of PBM phosphorylation confers constitutive PDZ binding and an increased capacity of E6 to promote cell growth, whereas, in WT cells, PKA phosphorylation regulates E6:PDZ interactions to prevent uncontrolled growth. This could be further investigated using specific kinase inhibitors such as mPKI. In conditions of PKA inhibition, it is expected that growth levels of WT genome containing cells increase as PDZ protein interactions would no longer be sterically hindered by a phosphate group, resembling the constitutive binding conditions of the ΔPKA mutant. The PBM function is linked to the maintenance of episomal genomes and thus observation of the state of HPV genomes under conditions of PKA activation and inhibition may also be of interest (Delury et al., 2013).

A kinase screen of high-risk HPV E6 proteins revealed that the group 1 proteins are targeted at a common site within the PBM by either PKA or AKT. Less of the group 2 E6 proteins were predicted to be phosphorylated by PKA or other kinases, and in some cases (e.g. HPV 34 and 53) this phosphorylation is predicted to occur outside of PBM. The conservation of this site in these high-risk viruses may reflect a role for the modification in HPV-mediated cancer development. Indeed, both kinases have been implicated in growth factor-induced proliferative pathways (Jang et al., 2012). However, the recognition sites of these kinases differ by an extra amino acid between the double arginine and target
serine/threonine residues (PKA, RRX\textsuperscript{T}/\textit{s} and AKT, RRXX\textsuperscript{T}/\textit{s}) and both HPV16 and HPV58 are phosphorylated by both \textit{in vitro} (Boon \textit{et al.}, 2014). Furthermore, HPV31 has been shown to be phosphorylated by PKA within the core of the E6 protein with no effect on PDZ interactions (Boon \textit{et al.}, 2014). Thus, the Scansite screen is limited and unable to recapitulate \textit{in vivo} interactions, it is important to confirm PBM phosphorylation \textit{via in vitro} assays and to test the role of phosphorylation in the primary keratinocyte model of HPV propagation.

It has been demonstrated that HPV16 is phosphorylated by the kinase AKT and that expression of HPV16 E6 in HFKs causes accumulation of the active phosphorylated form of AKT (Spangle and Münger, 2010; Boon \textit{et al.}, 2014). Furthermore, deletion of the HPV16 E6 PBM was unable to activate AKT and degradation of NHERF-1 mediated by HPV16 E6 correlates with the activation of the AKT signalling pathway (Accardi \textit{et al.}, 2011; Spangle and Münger, 2010). This indicates a strong relationship between the HPV16 PBM function and cellular kinase activity, and has been suggested to be important during the differentiation dependent stage of the virus life cycle. While analysis of the PKA pathway in the life cycle of HPV18 is currently under investigation in the Roberts lab, it may also be of interest to the field to investigate the phosphorylation of the HPV16 PBM, perhaps using mutations in the PBM analogous to those used in the context of HPV18 during this thesis.
6.1 Overview of findings

High-risk HPV types 16 and 18 encode a C-terminal PDZ-binding motif in the E6 oncoprotein. Although this motif has been shown to direct the degradation of a number of cellular PDZ domain containing proteins, is known to influence the maintenance of viral genomes and is essential to the productive viral life cycle, the precise function of the interactions in the context of the virus life cycle has not been fully characterized. More recently, HPV16 studies demonstrated that expression of the E6 and E7 oncoproteins had no effect upon degradation nor altered cellular localization of endogenous PDZ proteins in primary keratinocytes (Choi et al., 2013). The study described in this thesis investigated the expression of known and suggested PDZ targets of E6 in the presence of whole HPV genomes. In the main, results in this thesis agree with Choi et al., 2013; the major known targets of HPV genomes were unaffected by the replicating virus at a protein level. Unexpectedly, the tight junction proteins ZO-1 and ZO-2 were affected by HPV18 and HPV16 respectively; ZO-2 protein was decreased whilst ZO-1 protein increased in keratinocytes. Both ZO-1 and ZO-2 proteins have been suggested to interact with E6, although this has yet to be confirmed in primary cells (Hernandez-Monge et al., 2013; Belotti et al., 2013). Thus, the next step would be to identify an interaction with the E6 PBM and utilize transcriptional analysis to identify whether HPV is exerting an effect at a gene level.

A number of microarray studies have identified changes in protein expression in the presence of HPV genomes, and when expressing HPV oncoproteins alone (Duffy et al.,
The deregulated genes identified in these studies are mostly downstream of the p53, pRb/E2F or interferon-responsive pathways, and thus changes are likely due to the effect of the E6 and E7 oncoproteins upon pathway components (Chang and Laimins, 2001; 2000; Nees et al., 2001; Karstensen et al., 2006). Work in this thesis demonstrates that HPV genomes exert an influence over the transcription of specific PDZ-domain containing protein genes, including DLG1, PTPN13, NHERF1 and PATJ. HPV-mediated upregulation of these genes may occur indirectly, via p53, pRb etc. or directly via a virally encoded factor.

Moreover, HPV differentially affects the transcription of DLG1 isoforms, with the DLG1β form largely unaffected by viral genomes and upregulation of an isoform containing a previously unidentified exon. In this study, an exon was identified in the DLG1 HOOK region, which exists in DLG1α transcripts. This insertion (I6) is predicted to include a 20-amino acid sequence and produce a truncated form of the DLG1 protein. However, it is also possible that I6 acts at an mRNA level.

Finally, studies described in this thesis investigated the role of PBM phosphorylation in the context of genome replication. The phosphorylation of the HPV18 E6 PBM in primary cells was confirmed. Although this did not affect E6:PDZ interactions in monolayer grown cells, much of this preliminary work forms some foundation for ongoing investigations in the Roberts lab which focuses upon the influence and importance of PKA in the HPV life cycle. As the motif is conserved amongst high-risk α-HPV types, these data suggest that phosphorylation of the PBM performs a crucial role in regulating E6 interactions during the virus life cycle.

As tight junctions regulate epithelial integrity and polarity, viral targeting of these proteins may be advantageous to HPV by interrupting differentiation of the infected basal
keratinocytes that would otherwise cause exit from the cell cycle. The PDZ proteins are multifunctional and thus there are likely to be functions, including the control of proliferation and pro-proliferative transcription factors, which are required at different stages of viral propagation. Especially since HPV targets subsets of PDZ proteins depending upon post-translational modifications, cellular localization or differential splicing of cellular proteins. As there is such specific PDZ targeting by E6 during the HPV life cycle, it can be assumed that the deregulation of the viral protein observed in transformed cells, and the ensuing disruption of polarity complexes, enhances HPV-mediated malignant progression.

6.2 Future directions

Despite the increase in transcript levels, no change was observed in the abundance or distribution of major known E6:PDZ targets, DLG1 and SCRIB during the early stages of the HPV life cycle. Although no changes in the protein abundance of PDZ targets (DLG1, PTPN13, NHERF1, PATJ) was observed in monolayer grown cells, some changes in protein distribution were observed in differentiated cells. However, more thorough investigation of degradation pathways may identify an increase in transcript correlates with translation, but also with an increase in protein turnover. Alternatively, the transcripts may be unstable or again turned over by infected cells more quickly to counter the detrimental effects of increased PDZ mRNA. Studies by Watson et al., 2002 revealed an increase in cytoplasmic DLG1 in CIN1. Further investigation of DLG1, SCRIB and of other PDZ protein gene expression during neoplastic progression, for example following viral genome integration and in increasing CIN stages, would provide an insight into these interactions during the development of HPV-mediated cancers.

As well as the consequences of PDZ gene upregulation, it is also of interest to identify the viral mechanisms promoting upregulation of PDZ genes in such a selective fashion. This is
especially interesting in the case of DLG1 isoform expression; the ability of HPV to promote the inclusion of exons and the transcription of certain isoforms merits further exploration.

Replication of high-risk α-HPV genomes led to an increase in transcription of selected PDZ domain containing targets, and affected proteins ZO-1 and ZO-2. The effect of HPV appears to concentrate upon the tight junctions and their protein partners, which may aid viral propagation by promoting growth factor signalling or the release of proliferatory transcription factors (Balda and Matter, 2000; Betanzos et al., 2004). These proteins warrant further study in the HPV life cycle model and it may be of interest to investigate the consequences of kinase activation or inhibition on the tight-junction proteins.

Analysis of PDZ gene expression during the HPV life cycle identified the upregulation of a novel variant of DLG1. Expression of this form identified differential expression in HPV genome containing cell but not in cancers, suggesting that the form is not required during carcinogenesis but may be important in viral propagation. This work provides an exciting foundation for further investigation into the expression of this variant, and indeed of DLG1 isoforms, during HPV replication, in carcinogenesis and across tissue and cell contexts.

Although challenging to study, E6 PBM phosphorylation is likely a key influence in PDZ interactions. Thus, future studies investigating the effect of phosphorylation in differentiating cells, following genome integration and during the progression of HPV-driven cancers will further our understanding of E6:PDZ interactions.

6.3 Final statement

Overall, this study has contributed to current knowledge of HPV interactions with cellular PDZ proteins by performing a systematic analysis of PDZ proteins in the context of viral
replication. Although a number of major targets were unaffected by α-HPV types at a protein level, it was shown the virus acts on DLG1, PATJ, NHERF1 and PTPN13 at a gene transcription level and on ZO-1 and ZO-2 at a protein level. Furthermore, it was observed that HPV genomes promoted inclusion of a novel insertion of DLG1.

Together, data from this study offers a new perspective on the interactions between oncogenic HPV types and PDZ domain-containing targets, indicating that deregulation of PDZ protein function by the virus occurs via multiple mechanisms.

Disruption of cell-cell junctions contributes to invasive growth and, ultimately, metastases. A thorough knowledge of E6:PDZ interactions and their regulation may aid their manipulation during HPV-mediated cancers and thus the development of therapeutic strategies targeting the PDZ:PBM interaction.
REFERENCES


Ang, K.K. and Sturgis, E.M. (2012) Human Papillomavirus as a Marker of the Natural History and Response to Therapy of Head and Neck Squamous Cell Carcinoma. Seminars
in Radiation Oncology, 22 (2): 128–142


Boon, S.S. and Banks, L. (2012a) High Risk HPV E6 oncoproteins interact with 14-3-3z in a PDZ binding Motif (PBM)-dependant manner. pp. 1–25
Boon, S.S. and Banks, L. (2012b) High Risk HPV E6 Oncoproteins Interact with 14-3-3ζ in a PDZ Binding Motif (PBM)-Dependent Manner. *Journal of Virology*


human homologue of drosophila discs large oncosuppressor in histologic samples from human papillomavirus-associated lesions as a marker for progression to malignancy. *International Journal of Cancer*, 111 (3): 373–380


Chetkovich, D.M., Chen, L., Stocker, T.J., et al. (2002) Phosphorylation of the postsynaptic density-95 (PSD-95)/discs large/zona occludens-1 binding site of stargazin regulates
binding to PSD-95 and synaptic targeting of AMPA receptors. The Journal of neuroscience : the official journal of the Society for Neuroscience, 22 (14): 5791–5796


D’Souza, G., Kreimer, A.R., Viscidi, R., et al. (2007) Case-control study of human...


Das, B.C., Sharma, J.K., Gopalakrishna, V., et al. (1992) Analysis by Polymerase Chain-Reaction of the Physical State of Human Papillomavirus Type-16 Dna in Cervical Preneoplastic and Neoplastic Lesions. 73: 2327–2336


Complex. Genes & development, 5 (9): 1657–1671


transforming growth factor beta signaling and triggers its degradation by the proteasome. *Journal of Virology*, 79 (7): 4229–4237


binds zinc. **Oncogene**, 4 (9): 1089–1093


β-HPV5 E6 involves targeting of activated BAK for proteolysis by recruitment of the HERC1 ubiquitin ligase. International Journal of Cancer, pp. n/a–n/a


Hui, S., Xing, X. and Bader, G.D. (2013) Predicting PDZ domain mediated protein interactions from structure. Bmc Bioinformatics, 14 (1)


physiology, 227 (12): 3756–3767


Johansson, C., Somberg, M., Li, X., et al. (2012) HPV-16 E2 contributes to induction of
HPV-16 late gene expression by inhibiting early polyadenylation. *Embo Journal*


modulation of viral DNA amplification in organotypic cultures. *Proceedings of the National Academy of Sciences U.S.A.*


Leonard, A.S., Davare, M.A., Horne, M.C., et al. (1998) SAP97 is associated with the alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor GluR1 subunit. *Journal of
Biological Chemistry, 273 (31): 19518–19524


Embo Journal, 8 (13): 4099–4105


Newman, R.A. and Prehoda, K.E. (2009) Intramolecular Interactions Between the Src Homology 3 and Guanylate Kinase Domains of Discs Large Regulate Its Function in


phosphorylated at position -1 and controls lamellipodia formation in colon epithelial cells. 


Ramoz, N., Rueda, I.A., Bouadjar, B., et al. (1999) A susceptibility locus for epidermodysplasia verruciformis, an abnormal predisposition to infection with the oncogenic human papillomavirus type 5, maps to chromosome 17qter in a region


and E7 function in transgenic mouse models of cervical carcinogenesis. Cancer research, 63 (16): 4862–4871


Round, J.L., Humphries, L.A., Tommasian, T., et al. (2007b) MAGUK family member Dlgh1 coordinates TCR-dependent p38 activation, specifically linking TCR engagement to
activation of NFAT but not NF-kappa B. **Cytokine**, 39 (1): 36–36


Sonoda, T., Mochizuki, C., Yamashita, T., et al. (2006) Binding of glutamate receptor delta 2 to its scaffold protein, Delphilin, is regulated by PKA. Biochemical and Biophysical Research Communications, 350 (3): 748–752


Soubies, S.M., Hoffmann, T.W., Croville, G., et al. (2013) Deletion of the C-terminal ESEV domain of NS1 does not affect the replication of a low-pathogenic avian influenza virus H7N1 in ducks and chickens. The Journal of general virology, 94 (Pt 1): 50–58


Spoden, G., Kühlung, L., Cordes, N., *et al.* (2013) The Human Papillomaviruses Type 16, 18, and 31 share similar endocytic requirements for entry. *Journal of Virology*


Stöppler, M.C., Straight, S.W., Tsao, G., *et al.* (1996) The E5 gene of HPV-16 enhances
keratinocyte immortalization by full-length DNA. *Virology*, 223 (1): 251–254


Tanemoto, M., Fujita, A., Higashi, K., *et al.* (2002) PSD-95 mediates formation of a


Thomas, M., Kranjec, C., Nagasaka, K., et al. (2011) Analysis of the PDZ binding specificities of Influenza A Virus NS1 proteins. Virology Journal, 8 (1)


Tsubata, C., Higuchi, M., Takahashi, M., et al. (2005) PDZ domain-binding motif of human T-cell leukemia virus type I Tax oncoprotein is essential for the interleukin 2 independent growth induction of a T-cell line. Retrovirology, 2 (1)


Wang, W.-L., Yeh, S.-F., Chang, Y.-I., et al. (2003) PICK1, an anchoring protein that specifically targets protein kinase Calpha to mitochondria selectively upon serum stimulation in NIH 3T3 cells. The Journal of Biological Chemistry, 278 (39): 37705–37712


**Websites:**
http://www.cancerscreening.nhs.uk/cervical/about-cervical-screening.html#cost
Copyright Public Health England 2013 [Accessed 02/06/13]

http://pave.niaid.nih.gov/ [Accessed 22/12/14]