



ELSEVIER

Contents lists available at ScienceDirect

Virology

journal homepage: [www.elsevier.com/locate/yviro](http://www.elsevier.com/locate/yviro)

# The E4 protein; structure, function and patterns of expression



John Doorbar\*

Division of Virology, National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, United Kingdom

## ARTICLE INFO

Available online 7 September 2013

### Keywords:

Papillomaviruses  
HPV  
E4  
Keratin  
Life-cycle  
Cervix  
Biomarkers

## ABSTRACT

The papillomavirus E4 open reading frame (ORF) is contained within the E2 ORF, with the primary E4 gene-product (E1<sup>E4</sup>) being translated from a spliced mRNA that includes the E1 initiation codon and adjacent sequences. E4 is located centrally within the E2 gene, in a region that encodes the E2 protein's flexible hinge domain. Although a number of minor E4 transcripts have been reported, it is the product of the abundant E1<sup>E4</sup> mRNA that has been most extensively analysed.

During the papillomavirus life cycle, the E1<sup>E4</sup> gene products generally become detectable at the onset of vegetative viral genome amplification as the late stages of infection begin. E4 contributes to genome amplification success and virus synthesis, with its high level of expression suggesting additional roles in virus release and/or transmission. In general, E4 is easily visualised in biopsy material by immunostaining, and can be detected in lesions caused by diverse papillomavirus types, including those of dogs, rabbits and cattle as well as humans. The E4 protein can serve as a biomarker of active virus infection, and in the case of high-risk human types also disease severity. In some cutaneous lesions, E4 can be expressed at higher levels than the virion coat proteins, and can account for as much as 30% of total lesional protein content. The E4 proteins of the Beta, Gamma and Mu HPV types assemble into distinctive cytoplasmic, and sometimes nuclear, inclusion granules.

In general, the E4 proteins are expressed before L2 and L1, with their structure and function being modified, first by kinases as the infected cell progresses through the S and G2 cell cycle phases, but also by proteases as the cell exits the cell cycle and undergoes true terminal differentiation. The kinases that regulate E4 also affect other viral proteins simultaneously, and include protein kinase A, Cyclin-dependent kinase, members of the MAP Kinase family and protein kinase C. For HPV16 E1<sup>E4</sup>, these kinases regulate one of the E1<sup>E4</sup> proteins main functions, the association with the cellular keratin network, and eventually also its cleavage by the protease calpain which allows assembly into amyloid-like fibres and reorganisation of the keratin network. Although the E4 proteins of different HPV types appear divergent at the level of their primary amino acid sequence, they share a recognisable modular organisation and pattern of expression, which may underlie conserved functions and regulation. Assembly into higher-order multimers and suppression of cell proliferation are common to all E4 proteins examined. Although not yet formally demonstrated, a role in virus release and transmission remains a likely function for E4.

© 2013 The Author. Published by Elsevier Inc. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

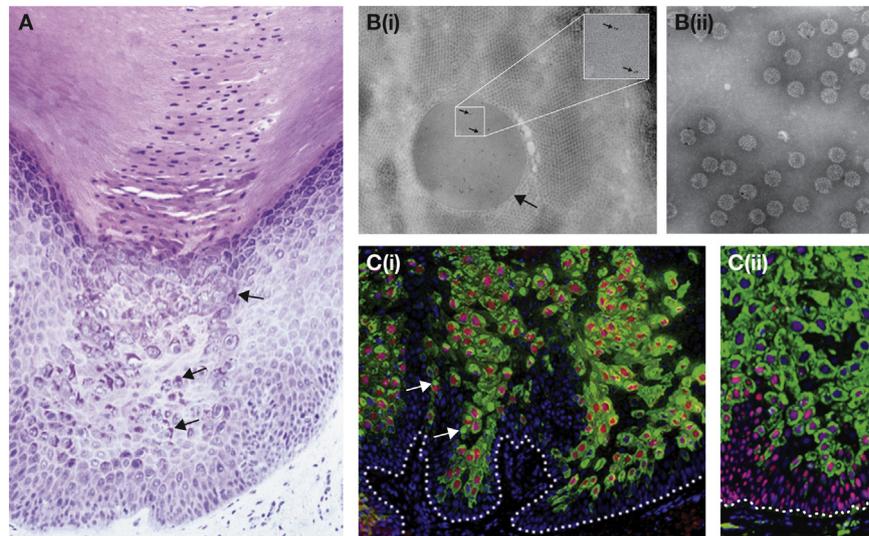
## The E4 protein

The E4 open reading frame (ORF) lies within the larger E2 ORF, and varies considerably in size between papillomavirus types (Doorbar and Myers, 1996) (see alignment in Figs. S1 and S2). In human papillomaviruses, the primary E4 gene product is expressed from a spliced mRNA (the E1<sup>E4</sup> message), in which

the initiation codon and first few amino acids are derived from the E1 ORF (Chow et al., 1987a,b; Doorbar et al., 1990; Milligan et al., 2007; Wang et al., 2011; Ozbun and Meyers, 1997). E1<sup>E4</sup> transcripts have been detected in biopsy material caused by diverse papillomavirus types, and typically represent the most abundant transcript in productive lesions. This abundance is also apparent at the protein level, with the E4 gene products accumulating to levels as high as 20–30% of the total protein content in some productive warts (Breitburd and Croissant Orth; Doorbar et al., 1986). In fact the E4 proteins were first observed in extracts of plantar warts induced by HPV1, as two small proteins of molecular weight 16 and 17 K that were visible by SDS gel electrophoresis at levels equivalent to those of the structural proteins of the cell (Croissant et al., 1985). Initially these small proteins were suspected to be

\* Fax: +44 20 8906 4477.

E-mail address: [jdoorba@nimr.mrc.ac.uk](mailto:jdoorba@nimr.mrc.ac.uk)



**Fig. 1.** E4 inclusion granules are typically found in cutaneous HPV types (Mu, Beta, Gamma Genera). Cutaneous HPV types from the Mu, Beta and Gamma Genera typically accumulate their E4 proteins to high level as inclusion granules visible by Hematoxylin & Eosin staining (granular structures arrowed in (A)). The precise composition of these inclusion granules is not known, but they are thought to be made up primarily of E4 (Breitburd and Croissant Orth; Rogel-Gaillard et al., 1992,1993; Doorbar et al., 1996). The appearance of such E4 inclusion granules varies depending on HPV type (Egawa, 1994; Gissmann et al., 1977; Gross et al., 1982). The lesion shown in A is caused by a Mu HPV type (HPV63), but E4 inclusion granules are also apparent in lesions caused by Beta and Gamma HPV types. The E4 inclusion granules are predominantly cytoplasmic, but can also be found in the nucleus. The individual electron-dense E4 granule shown in (B(i)) is imaged from the granular layer of a related Mu HPV type (HPV1) and is stained using a primary antibody to E4 and detected using a gold-conjugated secondary-antibody. E4-staining is visualised as dots on the E4 granule (arrowed). It appears that in cells with intact nuclei, there is little association between E4 and the virus particles, which form arrays around the intranuclear E4 inclusions. Following nuclear degeneration in the upper epithelial layers, virus particles enter the cytoplasm of the terminally differentiating epithelial cells and appear by immunostaining to be closely associated with the abundant cytoplasmic E4 proteins (Doorbar et al., 1997). In the case of the Mu HPV types, E4 can be detected in purified virus preparations (Doorbar and Gallimore, 1987), with virions often being associated with blebs (possibly E4) after density gradient centrifugation (B(ii)). In naturally-occurring papillomas caused by Mu HPV types, E4 expression begins in the lower epithelial layers as visualised by indirect immunofluorescence (C(i)) and (C(ii)). The green stain is E4, and the cytoplasmic inclusion granules are arrowed in (C(i)) (nuclei shown in blue after DAPI counterstain). The red stain in (C(i)) reveals the distribution of amplified viral DNA following Fluorescent In Situ Hybridisation (FISH) using a viral DNA probe. Genome amplification is generally confined to E4-positive cells, which in the lower layers are also positive for cellular replication proteins such as MCM (red stain in (C(ii))). It is generally thought that viral genome amplification occurs predominantly in E4-positive/replication-competent cells, which (in HPV1) are also positive for the G2 cell cycle marker CyclinB/Cdk1 (Davy et al., 2005).

breakdown products of the abundant skin keratins, but were subsequently shown to be the products of the viral E4 ORF (Breitburd and Croissant Orth; Doorbar et al., 1986,1988). Because of the great abundance of E4 in HPV1 warts (see Fig. 1), much of the early characterisation of E4 expression was carried out on this HPV type (Doorbar et al., 1986; Grand et al., 1989; Rogel-Gaillard et al., 1992, 1993; Roberts et al., 1993,1994), with studies on the high-risk HPV types following (Doorbar et al., 1991; Roberts et al., 1997; Pray and Laimins, 1995; Brown et al., 1994). Indeed, the abundance of E4 in lesions has suggested a role as a biomarker of active HPV disease (Middleton, 2003; Borgogna et al., 2012), and at the cervix, a marker of disease severity (Griffin et al., 2012; Doorbar and Cubie, 2005). The E4 proteins of all HPVs, and probably most animal HPV types, have some noticeable structural similarities (see detail provided in Tables 1–3, Katoh et al., 2002), but have diverged considerably in their primary amino acid sequences (see Figs. S1 and S2), and probably also the subtleties of their function.

#### E4 expression and its contribution to life-cycle success

Most molecular studies on E4 have focused on HPV1 from the Mu Genus (see Fig. 1), and HPV16 from the Alpha Genus (Species Group 9; Fig. 2, Table 1), with additional analysis coming from the study of other Alpha papillomaviruses including HPV11, 18 and 31 (Pray and Laimins, 1995; Brown et al., 1991; Frattini et al., 1997). The E4 proteins have however been detected in lesions caused by a much wider range of papillomavirus types, including HPV2 (Alpha), HPV3 (Alpha), HPV4 (Gamma), HPV5 (Beta), HPV6 (low risk (LR) Alpha), HPV8 (Beta), HPV11 (LR Alpha), HPV58 (HR Alpha), HPV59 (HR Alpha), HPV63 (Mu) and HPV65 (Gamma) (Fig. 3 (Borgogna et al., 2012; Middleton et al.,

2003; Doorbar and Gallimore, 1989; Brown et al., 2004)), as well as several animal papillomaviruses including Canine Oral Papillomavirus (COPV, (Nicholls et al., 2001)), Rabbit Oral Papillomavirus (ROPV, (Maglennon et al., 2011)), Cottontail Rabbit Papillomavirus (CRPV, (Peh et al., 2004)) and Bovine Papillomavirus (BPV, (Peh et al., 2002)). In each case, the detection of E4 by immunofluorescence is simple and straightforward, suggesting that these papillomaviruses resemble HPV1 to some extent in their ability to express their E4 proteins at high level. Such comparative analysis has also revealed broad similarities in the timing of expression, with all papillomaviruses that have been examined showing prominent E4 accumulation in the mid and upper epithelial layers during productive infection (Maglennon et al., 2011; Peh et al., 2002; Doorbar et al., 1997) (see protein distribution images in Figs. 1 and 2). A close association is apparent between the first detection of E4 by immunofluorescence and the onset of vegetative viral DNA replication or genome amplification, which when considered alongside its abundance, points to an important role for E4 in the late stages of the virus life cycle (Breitburd and Croissant Orth; Peh et al., 2002; Doorbar et al., 1996). Historically, the E4 ORF was classified as an early viral gene (Chen et al., 1982; Danos et al., 1982), because it lies in the early region of the viral genome and is embedded amongst viral genes that regulate cell-cycle entry and genome maintenance. No obvious function for E4 during the early stages of the virus life-cycle has yet been convincingly described however. Early studies on BPV showed that it was not involved in cell transformation by this virus (Neary et al., 1987), while E4 null mutants of CRPV, were not apparently compromised in their ability to produce papillomas in either domestic or cottontail rabbits (Peh et al., 2004), which is consistent with a primary role during the late rather than the early stages of infection. Amongst HPVs, only the high-risk HPV types have been extensively studied, partly because of the availability of convenient model systems for these viruses (Frattini et al., 1997;

**Table 1**  
The 16E1<sup>^</sup>E4 protein is structurally and functionally modified by cellular kinases and proteases as the infected cell undergoes terminal differentiation. These changes are mediated partly by the viral gene products that drive cell cycle entry and affect normal cell signalling (e.g. E6, E7 and E5), and partly by the differentiating environment within the epithelium. Changes in the cellular environment are thought to modulate the functions of viral and cellular proteins simultaneously in order to ensure virus genome amplification, virus assembly and virus release.

| Epithelial layer | 16E1 <sup>^</sup> E4 function   | 16E1 <sup>^</sup> E4 organisation |
|------------------|---|-----------------------------------|
| Upper            | 16E1 <sup>^</sup> E4 is cleaved by calpain following cell cycle exit. Additional phosphorylation around the loop by PKC is predicted and may stabilize the amyloid form in the upper layers. The high abundance of E4 is thought to result from its ability to assemble into amyloid fibres                       |                                   |
| Upper & middle   | 16E1 <sup>^</sup> E4 is phosphorylated by CDK (CyclinB/Cdk and CyclinA/Cdk) in the G2-phase-like/differentiating cells where genome amplification primarily takes place. This is predicted to open the loop and enhance calpain cleavage  |                                   |
| Middle           | 16E1 <sup>^</sup> E4 is phosphorylated by MAPK in the S-phase-like/differentiating cells where genome amplification begins. This compacts the loop and enhances keratin binding   |                                   |
| Middle & lower   | 16E1 <sup>^</sup> E4 levels rise in the mid epithelial layers following up-regulation of the viral late promoter. In this conformation, it is thought that the E1 <sup>^</sup> E4 protein does not accumulate to high-level in the cell and has a shorter half-life than the keratin-bound and/or multimeric form |                                   |
| Lower            | In the lower epithelial layers 16E1 <sup>^</sup> E4 is not easily detected. It is either not present or present at very low levels  |                                   |

Flores et al., 1999; Hummel et al., 1992) (Fig. 2). In experimental systems, E4 null mutants of HPV16, 18 and 31 are not significantly perturbed in their ability to be maintained in proliferating 'basal-like' keratinocytes in monolayer culture, and do not show consistent differences in their ability to drive cell proliferation (Nakahara et al., 2005; Wilson et al., 2005, 2007). While these studies do not rule out an early function for E4, they have not yet allowed such a function to be identified. By contrast, both animal studies using CRPV (Peh et al.,

2001), and organotypic raft studies on HR HPV types (Nakahara et al., 2005; Wilson et al., 2005, 2007) have suggested a role for E4 in modulating genome amplification and virus synthesis. A single study that applied the organotypic raft approach to HPV11 (a low-risk HPV type) failed to report a dramatic effect on these functions (Fang et al., 2006). This may reflect a less significant involvement of the low-risk E4 proteins in these late events and/or differences in genome maintenance and amplification requirements between the high and

**Table 2**  
E4 binding-partners and their potential significance during infection.

| E4 binding partners   |  |  |   |   |
|---|--|--|---|---|
| Binding partner   | Shown for HPV type   | Additional information   | Significance or potential significance  | Refs.   |
| <b>Cytokeratin</b><br>Cytokeratin intermediate filaments & cytokeratin monomers | <b>HPV1<sup>a</sup>&amp;HPV16<sup>a</sup></b><br>... but thought to be a key function of most, if not all, E4 proteins (McIntosh et al., 2010) | Direct binding to type 1 keratins (e.g. keratin 18) but not type 2 keratins or vimentin (Wang et al., 2004). Binding is dependent on the 16 amino acids at the N-terminus of the E1 <sup>^</sup> E4 protein, with the leucine cluster being absolutely required (Roberts et al., 1997)   | <ol style="list-style-type: none"> <li>1. Interaction has been shown <i>in vitro</i> and <i>in vivo</i> (Wang et al., 2004) and triggers a cellular stress-response with activation of stress-associated kinases (e.g p38 and pJNK (McIntosh et al., 2010))</li> <li>2. Interaction leads to the reorganisation and hyper phosphorylation of the cytokeratin IF network (McIntosh et al., 2010).</li> <li>3. Interaction is thought to contribute to virus escape and cell destruction in the upper epithelial layers</li> </ol>  | Roberts et al. (1994), Roberts et al. (1993), Roberts et al. (1997), McIntosh et al. (2010), Wang et al. (2004) |
| <b>HPV E4</b>   | <b>HPV1<sup>a</sup> &amp; HPV16<sup>a</sup></b><br>... self-association thought to be a conserved property of E4 proteins                      | For many E4 proteins, self-association is directed (at least in part) by C-terminal amino sequences with beta aggregation potential (McIntosh et al., 2008). Loss of C-terminal residues increases E4 solubility in the cell. Cutaneous HPV types may have additional upstream regions that direct multimerisation   | <ol style="list-style-type: none"> <li>1. Self-association and the formation of amyloid fibres underlies E4-accumulation <i>in vivo</i> (McIntosh et al., 2008)</li> <li>2. Multimerisation is thought to enhance the avidity of keratin association (Wang et al., 2004) and leads to the loss of normal keratin filament dynamics (McIntosh et al., 2010)</li> <li>3. In cutaneous HPV types such as HPV1 and HPV4, E4 assembles into distinctive cytoplasmic (and occasionally nuclear) inclusion granules</li> </ol>   | Roberts et al. (1997), McIntosh et al. (2010), Wang et al. (2004)   |
| <b>Kinases</b><br>CyclinB/Cdk1  | <b>HPV16 &amp; HPV18 HPV1</b>  | For HPV16 E1 <sup>^</sup> E4, association appears to involve residues T <sup>22</sup> and T <sup>23</sup> (putative cdk phosphorylation site) rather than the <sup>41</sup> RRL <sup>43</sup> motif (putative cyclin binding site) (Davy et al., 2005). In HPV18, the putative cyclin binding motif <sup>43</sup> RRL <sup>46</sup> (rather than T <sup>23</sup> ) has been reported to important for binding (Knight et al., 2011; Ding et al., 2013)<br>For HPV1 E4, G2 arrest function appears to reside with the 16K N-terminally-truncated form of the protein and requires residue T <sup>27</sup> (E1 <sup>^</sup> E4 position) which is also referred to as T <sup>13</sup> (its position in the 16K E4 species) (Knight et al., 2004). HPV1 E4-mediated G2 arrest is dependent on sustained Wee1 kinase activity which phosphorylates Cdk1 at T <sup>15</sup> (Knight et al., 2006) | <ol style="list-style-type: none"> <li>1. For HPV16, the sequestration of CyclinB/Cdk1 in the cytoplasm arrests proliferating cells (such as are found in neoplasia) in G2 (Davy et al., 2002, 2005), and contributes to genome amplification-efficiency (Wang, Jackson, Doorbar; Manuscript in preparation)</li> <li>2. In HPV18, CyclinB/Cdk1 sequestration is similarly important for G2 arrest. In productive HPV18 rafts, no role for 18 E1<sup>^</sup>E4 G2 arrest was reported (Knight et al., 2011). This may reflect a more limited suprabasal cell proliferation seen in this system when compared to HPV16</li> <li>3. In HPV1, CyclinB/Cdk1 activity is thought to be inhibited rather than being sequestered. It has been suggested that the full length E1<sup>^</sup>E4 protein contributes to a role for 16K-mediated G2-arrest in genome amplification by inhibiting re-replication of cellular DNA in the arrested cells (Knight et al., 2004). Cdk1 inhibition is mediated by Wee1(Knight et al., 2006). The significance for the HPV1 life cycle is thought to be in optimising genome amplification (Knight et al., 2006)</li> </ol> | Davy et al. (2005), Davy et al. (2002), Nakahara et al. (2002)  |
| CyclinA/Cdk2  | <b>HPV16&amp;HPV18</b>   | For HPV18, association requires the <sup>43</sup> RRL <sup>46</sup> motif (Knight et al., 2011) that also mediates CyclinB/ Cdk1 binding. In the case of HPV16 E1 <sup>^</sup> E4, residues T <sup>22</sup> and T <sup>23</sup> are essential for both CyclinA/Cdk2 binding and CyclinB/Cdk1 association (Davy et al., 2006). CyclinA/Cdk2 phosphorylates 16 E1 <sup>^</sup> E4 at Serine 32 (S <sup>32</sup> )  | <ol style="list-style-type: none"> <li>1. For HPV 16, CyclinA/Cdk2 is sequestered in the cytoplasm in G2 but not S-phase. It is thought that this may augment the G2 arrest activity mediated through CyclinB/Cdk1 association</li> <li>2. For HPV 18 CyclinA/Cdk2 sequestration does not occur in all cell types and does not appear essential for the G2 arrest phenotype (Knight et al., 2011)</li> <li>3. The HPV1 E4 16K protein appears to stimulate an elevation in CyclinA levels on G2 arrest (Knight et al., 2004)</li> </ol>   | Knight et al. (2011), Ding et al. (2013), Davy et al. (2006)  |
| CyclinE/Cdk2  | <b>HPV18</b>   | Requires <sup>43</sup> RRL <sup>46</sup> motif in 18 E1 <sup>^</sup> E4. Interaction seen so far only in <i>in vitro</i> binding assays (Ding et al., 2013)  | Interaction with CyclinE/Cdk2 not yet seen in mammalian cells. 18 E1 <sup>^</sup> E4 does not appear to regulate S-phase entry. Significance uncertain  | Ding et al. (2013)  |
| p42 MAPK  | <b>HPV16, HPV11</b>  | Transient association and phosphorylation of 16 E1 <sup>^</sup> E4 at Threonine (HPV16 T <sup>57</sup> ; HPV11 T <sup>53</sup> ) around the onset of genome amplification  | HPV16 E1 <sup>^</sup> E4 is phosphorylated and functionally modified to enhance keratin binding by p42 MAPK (Wang et al., 2009). It is thought that   | Wang et al. (2009), Knight et al. (2011), Bell et al. (2007)  |

Table 2 (continued)

| E4 binding partners                  |   |  |  |                                       |
|--------------------------------------|---|--|--|---------------------------------------|
| Binding partner                      | Shown for HPV type  | Additional information   | Significance or potential significance   | Refs.                                 |
| Protein kinase A                     | <b>HPV16, HPV18<br/>HPV11, HPV1</b>   | (Wang et al., 2009). HPV18 E1^E4 is reported not to be a p42 MAPK target (Knight et al., 2011; Bell et al., 2007)<br>Phosphorylation site not mapped. S <sup>43</sup> predicted in HPV16 (Wang et al., 2009) | this helps initiate E4-accumulation following activation of the differentiation-dependent promoter (Wang et al., 2009)<br>Possible regulator of E1^E4 structure. PKA sites reported in many E4 proteins (Grand et al., 1989; Wang et al., 2009; Bryan et al., 2000; Knight et al., 2011)<br>Significance uncertain | Knight et al. (2011)                  |
| Protein kinase C                     | <b>HPV16</b>  | Phosphorylation site not mapped. S <sup>32</sup> , S <sup>43</sup> and S <sup>44</sup> predicted in HPV16 (Wang et al., 2009)  | Thought to maintain phosphorylation at the CyclinB/Cdk1 site as the cell exits the cell cycle and differentiates   | Wang et al. (2009)                    |
| SR protein kinase1                   | <b>HPV1</b><br>HPV16 & 18 (E1^E4 proteins appear to associate less well and are not SRPK1 substrates) | Phosphorylation site not mapped. Kinase binding thought to depend on arginine-rich region with amino acids G <sup>44</sup> , R <sup>45</sup> and R <sup>47</sup> being important (Roberts et al., 2008)      | Possible regulator of HPV1 E1^E4 structure. Association with SRPK could affect RNA processing and/or nuclear structure<br>Significance uncertain   | Roberts et al. (2008)                 |
| <b>Other PV proteins</b>             |   |  |  |                                       |
| HPV E2                               | <b>HPV16&amp;HPV18</b>  | Co-expression leads to stabilisation of E2 and E4, and to changes in E2 distribution (Davy et al., 2009)   | May have effects on viral DNA replication, transcription efficiency or virus assembly. Timing and extent of E2-association during the virus life-cycle is not yet known<br>Possible role in virus assembly or transmission   | Davy et al. (2009)<br>Sorathia (2005) |
| HPV L1, L2                           | <b>HPV 16</b>   | Low affinity association suggested   |  |                                       |
| <b>Other cellular proteins</b>       |   |  |  |                                       |
| E4-DBP                               | <b>HPV16</b>  | E4-DBP is a putative DEAD-box RNA helicase   | Significance uncertain. Possible role in RNA processing?   | Doorbar et al. (2000)                 |
| Promyelocytic Leukemia protein (PML) | <b>HPV1</b>   | E4 expression causes the re-organisation of PML from ND10 domains to nuclear E4 inclusion granules. Interaction not thought to be direct (Roberts et al., 2003)  | Redistribution dependent on the presence of nuclear E4 inclusions, rather than the more prominent cytoplasmic E4 inclusions. Suggested involvement in viral genome amplification, perhaps by modulating anti-viral responses (Roberts et al., 2003)  | Roberts et al. (2003)                 |
| TopBP1                               | <b>HPV16</b>  | Interaction seen in two-hybrid screen in yeast (JD unpublished data)   | Significance uncertain (but TopBP1 is a target of the HPV16 E2 protein (Donaldson et al., 2007))   |                                       |
| MCM                                  | <b>HPV1</b>   | Interaction reported in <i>in vitro</i> binding experiments (Roberts et al., 2008)   | May play a role in inhibition of cellular DNA replication mediated by full-length HPV1 E1^E4 (Roberts et al., 2008)  | Roberts et al. (2008)                 |

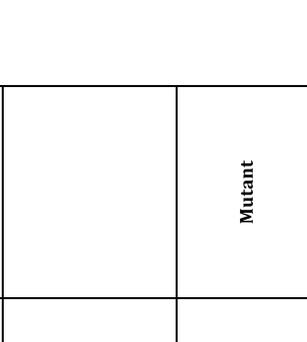
<sup>a</sup> Extensively validated in cell culture and in patient biopsies.

**Table 3**  
Key features of the HPV E4 proteins.

| HPV 1 E1 <sup>E4</sup>  |                   | HPV1 E1 <sup>E4</sup>  |     |                   |     |     |     |
|---|-------------------|--|-----|-------------------|-----|-----|-----|
| <p>10 20 30 40 50 60 70 80 90 100 110 120</p> <p>MADNKAPQGLLGLLQYTPPTQPYPRVTPPSNRRPSTTPNSQDRGRPRRSDDKSRKHLYADGLTDGEDPEVPEVEDEEKENQRPLGHPDLSLLRETLEVYTQRLKRDILQDLDDFCRKLGIHPWSV</p> <p><b>T</b> G2 cell cycle arrest is a characteristic of the 16K cleaved form. Threonine at position 27 (T<sup>13</sup> in the truncated 16K cleaved form) is essential for G2 arrest.<br/>Cell cycle-arrest in G2 involves a reduction in CyclinB/Cdk1 levels and activity</p> <p>Other reported phosphorylation sites -<b>STT</b>--<b>S</b>--<b>S</b>--<b>ET</b>--<b>T</b>--<b>S</b></p> <p>(phosphorylated by PKA. Specific roles not yet established)</p>                 |                   | Phosphorylation & Cell Cycle Arrest  |     |                   |     |     |     |
| <p><b>QGLLGLL</b> Leucine cluster motif</p> <p><b>MADNKAPQG</b> predicted Alpha helix</p> <p><b>LLGLLQY</b><br/><b>LGLLQYTPTT</b></p> <p><b>RVT</b></p> <p>Amino acids implicated in Zinc Binding <b>H</b></p> <p>Multimerisation region <b>SLLRETLEVYTQRLKRDILQDLDDFCRKLGIHPWSV</b></p> <p>Constrained loop-region (predicted) <b>SLLRETLEVYTQRLK</b> <b>QDLDDFCR</b> (Garnier-Robson)</p> <p>predicted Alpha helix <b>RKHLY</b> <b>EVPEVEDEEKENQR</b> <b>PEVEDEEKENQ</b> <b>SLLRE</b> <b>RDILQDL</b> (Chou-Fasman)</p> <p>predicted Beta strand <b>TLEVYTQRLK</b> <b>TL</b> (Garnier-Robson)</p> <p>predicted Beta strand <b>TLEVYTQRLK</b> (Chou-Fasman)</p> |                   | Motifs   |     |                   |     |     |     |
| <p><b>D</b> <b>K</b> <b>T</b> <b>TT</b> <b>T</b> <b>S</b> <b>STT</b> <b>S</b> <b>S</b> <b>S</b> <b>S</b> <b>T</b> <b>S</b> <b>T</b> <b>T</b> <b>S</b></p> <p><b>DR</b> <b>RR</b> <b>DR</b> <b>RR</b> <b>DKD</b> <b>RKH</b> <b>D</b> <b>D</b> <b>ED</b> <b>E</b> <b>E</b> <b>EDEEKE</b> <b>R</b> <b>H</b> <b>D</b> <b>RE</b> <b>E</b> <b>R</b> <b>KRD</b> <b>D</b> <b>DD</b> <b>RK</b> <b>H</b> <b>S</b></p> <p>-----Overall Positive Charge-----</p> <p>-----Overall Negative Charge-----</p>   |                   | Charge Distribution  |     |                   |     |     |     |
| <p><b>P</b> <b>P</b> <b>P</b> <b>P</b> <b>PP</b> <b>P</b> <b>P</b></p> <p>-----proline-rich domain 1-----</p> <p>-----proline-rich domain 2-----</p>   |                   | Proline-Rich Regions   |     |                   |     |     |     |
| <p><b>TOPYPRVTE</b> MoAb 7.76</p> <p><b>SNRRPSTTP</b> MoAb 9.95</p> <p><b>SNRRPSTTP</b> MoAb 9.60</p> <p><b>HLYADGLTD</b> MoAb 4.37</p> <p><b>VEDEEKENQ</b> MoAb 8.40</p>   |                   | Epitopes   |     |                   |     |     |     |
| <p><b>Keratin Binding/Leucine Cluster</b> <b>Proline-Rich</b> <b>Positive Charge</b> <b>loop</b> <b>Negative Charge</b> <b>Proline-Rich2</b> <b>Multimerisation Region</b></p>  |                   | Modular Structure  |     |                   |     |     |     |
| <p>Position 61* GLTDGEDPEVPEVEDEEKENQRPLGHPDLSLLRETLEVYTQRLKRDILQDLDDFCRKLGIHPWSV</p> <p>Position 59* ADGLTDGEDPEVPEVEDEEKENQRPLGHPDLSLLRETLEVYTQRLKRDILQDLDDFCRKLGIHPWSV</p> <p>Position 16* YTPPTQPYPRVTPPSNRRPSTTPNSQDRGRPRRSDDKSRKHLYADGLTDGEDPEVPEVEDEEKENQRPLGHPDLSLLRETLEVYTQRLKRDILQDLDDFCRKLGIHPWSV</p> <p>MADNKAPQGLLGLLQYTPPTQPYPRVTPPSNRRPSTTPNSQDRGRPRRSDDKSRKHLYADGLTDGEDPEVPEVEDEEKENQRPLGHPDLSLLRETLEVYTQRLKRDILQDLDDFCRKLGIHPWSV</p>   |                   | <table border="1"> <tr> <td>10K</td> <td rowspan="4">cleavage products</td> </tr> <tr> <td>11K</td> </tr> <tr> <td>16K</td> </tr> <tr> <td>17K</td> </tr> </table> | 10K | cleavage products | 11K | 16K | 17K |
| 10K   | cleavage products |  |     |                   |     |     |     |
| 11K   |                   |  |     |                   |     |     |     |
| 16K   |                   |  |     |                   |     |     |     |
| 17K   |                   |  |     |                   |     |     |     |

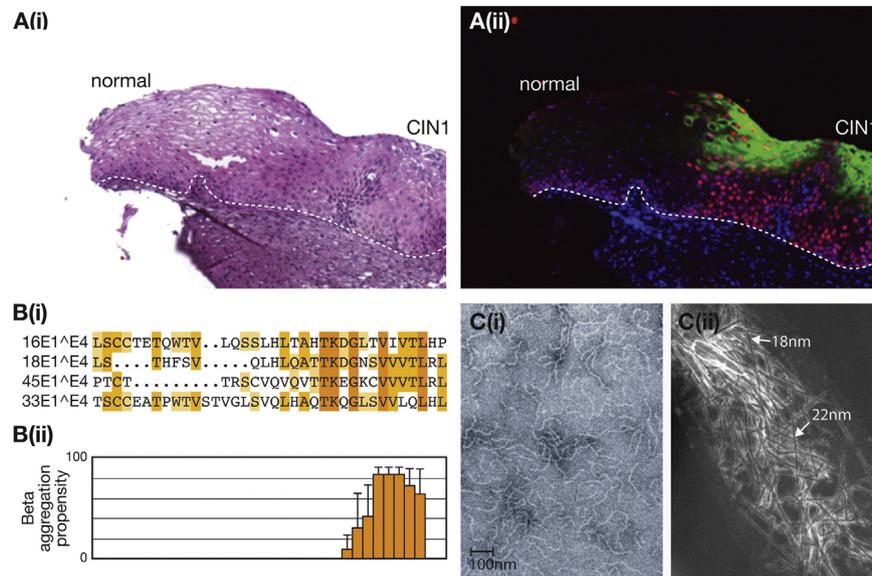




| Structural Organisation (NMR + FRET)   | E1 <sup>Δ</sup> E4 structure is constrained by sequences at the N and C-termini that also have additional functions in keratin-binding and multimerisation. These additional functions are thought to require relaxation of the constrained E1 <sup>Δ</sup> E4 structure by phosphorylation and N-terminal cleavage. The pattern of positively and negatively charged amino acids in E1 <sup>Δ</sup> E4 are thought to contribute to the constrained E4 structure |   |
|--|---|---|
| <p><b>HPV 16</b></p>    |   |   |
| <p><b>HPV 16 E1<sup>Δ</sup>E4 Mutants</b></p> <p>MADPAAATKYPLLKLLGSGTWPTPPRPPIPKPSWPAPKKHRRLLSSDQSQSTPPTPATPLSCCCTETQWTVLQSSHLTAHTKDKGLTIVITLHP</p>  | <p>10   20   30   40   50   60   70   80   90  </p>   |   |
| <p>ΔADPA</p> <p>ΔLLKLL</p> <p>Loss of keratin association. No keratin network disruption</p>   | <p>ΔTPRPPI</p> <p>ΔPLPKPS</p> <p>ΔPSFWAP</p> <p>ΔPKKRR</p> <p>ΔRLSSDQ</p> <p>ΔDQSQT</p>   | <p>ASLHLTA</p> <p>ΔTKD</p> <p>ΔLTVI</p> <p>ΔIVTLHP</p> <p>Reduced keratin association as a result of lower multimerisation-capability &amp; reduced binding-avidity</p> |
| <p>MADPAAATKYPLLKLLGSGTWPTPPRPPIPKPSWPAPKKHRRLLSSDQSQSTPPTPATPLSCCCTETQWTVLQSSHLTAHTKDKGLTIVITLHP</p> <p>N-terminal 16aa targets GFP to keratin network</p>  | <p>T<sup>22,23</sup></p> <p>Sequestration of CyclinB/Cdk1. Essential for 16 E1<sup>Δ</sup>E4's G2 arrest function.</p>  | <p>S<sup>86</sup></p> <p>S<sup>88</sup></p> <p>S<sup>90</sup></p> <p>P<sup>80</sup></p> <p>P<sup>83</sup></p> <p>Reduced ability to multimerise</p>                     |
| <p>MDFPAAATKYPLLKLLGSGTWPTPPRPPIPKPSWPAPKKHRRLLSSDQSQSTPPTPATPLSCCCTETQWTVLQSSHLTAHTKDKGLTIVITLHP</p> <p>Removal of E1 sequences Prevent keratin binding</p> <p>MLLKLL<sup>GFP</sup></p> <p>Mitochondrial targeting Keratin binding lost</p> | <p>ΔTETQWT</p>  | <p>Multimerisation/beta-aggregation</p>   |
| <p>MADPAAATKYPLLKLLGSGTWPTPPRPPIPKPSWPAPKKHRRLLSSDQSQSTPPTPATPLSCCCTETQWTVLQSSHLTAHTKDKGLTIVITLHP</p>  | <p>Proline domain 1</p> <p>Positive Charge</p> <p>Loop</p> <p>Negative Charge</p> <p>Proline domain 2</p>   | <p>Can not cross-link and collapse the keratin filament network</p>   |
| <p>MADPAAATKYPLLKLLGSGTWPTPPRPPIPKPSWPAPKKHRRLLSSDQSQSTPPTPATPLSCCCTETQWTVLQSSHLTAHTKDKGLTIVITLHP</p>  | <p>10   20   30   40   50   60   70   80   90  </p>   |   |







**Fig. 2.** High-risk Alpha HPV E4 expression as amyloid fibres in low-grade Cervical Neoplasia. The E4 proteins of HPV16 have been extensively studied because of the association of HPV16 with cervical cancer. A haematoxylin-eosin stained image showing the junction between normal uninfected epithelium (labelled normal) and a region of low-grade cervical neoplasia (labelled CIN1) is shown in (A(i)). The same piece of tissue is shown in (A(ii)) after staining for HPV16 E4 (green) and MCM (red). The HPV16 E4 protein is abundant in the upper epithelial layers in cells supporting viral genome amplification (Doorbar et al., 1997). Although the E4 protein of HPV16 does not assemble into prominent cytoplasmic inclusion granules as seen for the Mu, Gamma and Beta HPV types, the HPV16 protein is predominantly cytoplasmic with a low-level of nuclear staining, and associates with keratin filaments. This intracellular distribution is seen in lesions caused by other Alpha HPV types including HPV2, HPV11 and HPV18. The E4 proteins of many HPV types and the majority of Alpha types contain a predicted beta-aggregation motif at their C-terminus (B). Current thinking is that this multimerisation/aggregation motif is carefully regulated in the context of the full-length protein, and that the capacity to multimerise is controlled by phosphorylation and proteolytic cleavage (see Table 1). C-terminal homology is shown in (B(i)), with aggregation potential for key E4 proteins shown in (B(ii)). The assembly of 16 E1<sup>E4</sup> into amyloid-like fibres occurs slowly after expression and purification of WT E1<sup>E4</sup> in bacteria, and more quickly after expression of N-terminal E1<sup>E4</sup> deletions as shown in (C(i)) (McIntosh et al., 2008). Soluble WT 16 E1<sup>E4</sup> can be converted into amyloid-like fibres following cleavage with calpain, which removes the N-terminal 17 amino acids (C(ii)). It appears that individual E4 fibres can aggregate laterally to produce larger structures. The width of individual filaments are indicated with arrows in (C(ii)).

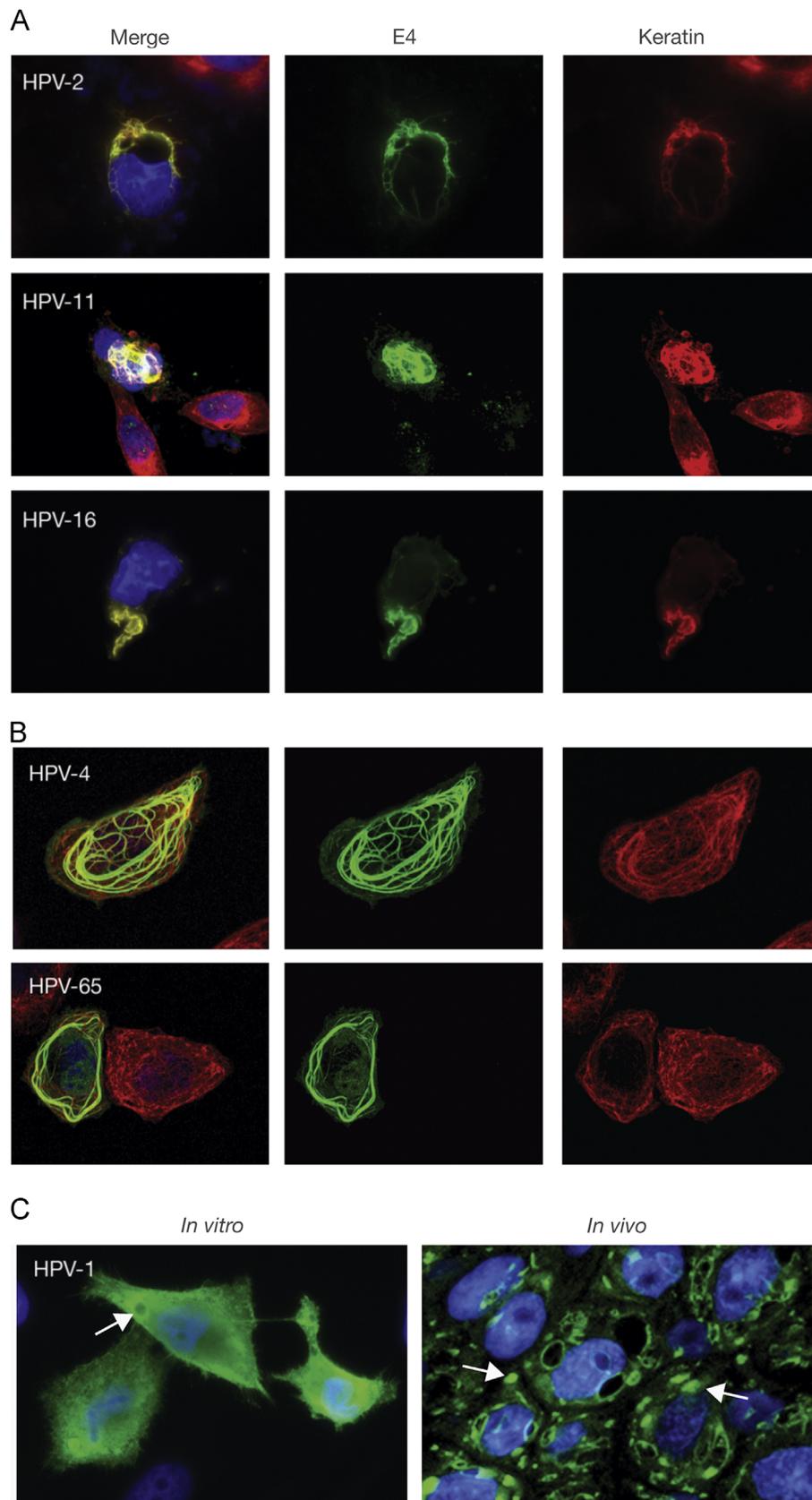
low-risk HPVs. LR HPV types such as HPV11 do not replicate well in standard keratinocyte monolayer culture however, which makes analysis of life-cycle events in organotypic culture difficult (Fang et al., 2006; Thomas et al., 2001; Oh et al., 2004). The analysis of HPV11 E4 in genital warts has however suggested that this protein can become cross-linked by endogenous transglutaminase to the cornified envelope and that the integrity of the cornified squame is compromised (Brown and Bryan, 2000; Brown et al., 2006; Bryan and Brown, 2000). It is thought that such 'fragile' squames may facilitate efficient person to person transmission (Bryan and Brown, 2001).

## E4 processing and functional regulation in infected epithelium

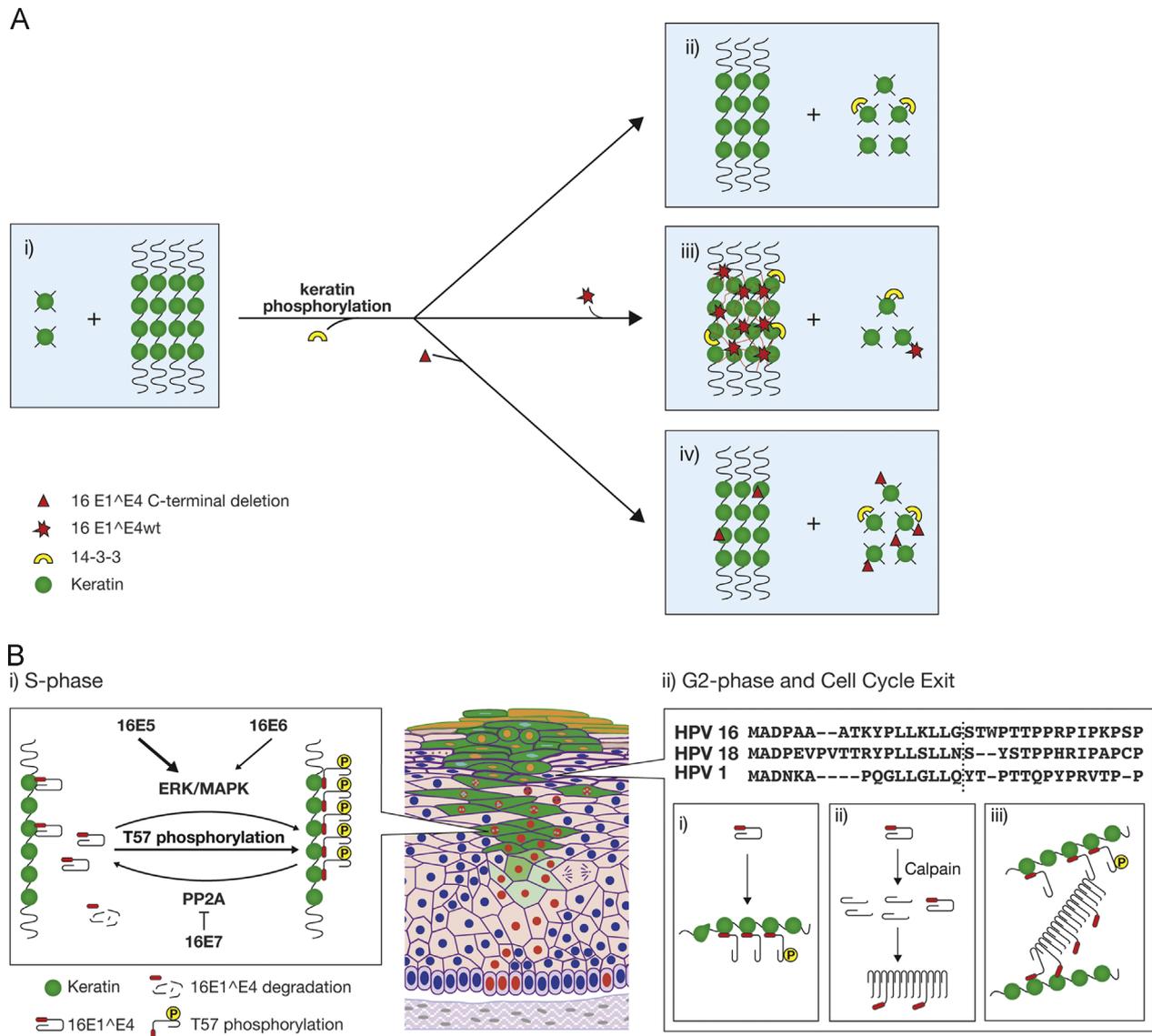
### Cutaneous infections (Mu, Gamma, Beta HPV types)

One of the key differences between the E4 proteins of different papillomaviruses lies in the structures that they form in the infected cell, and in the timing at which these structures appear (see Tables 1 and 3, Figs. 1 and 2). Many cutaneous HPV types such as HPV1 and 63 (Mu Genus), HPV 4 and 65 (Gamma Genus) and HPV 5 and 8 (Beta Genus) form productive papillomas that are characterised by the appearance of cytoplasmic inclusion granules (see Fig. 1 (Croissant et al., 1985; Egawa, 1993, 1994; Gissmann et al., 1977; Jablonska et al., 1985; Laurent et al., 1982)). These inclusion granules are typically first noticeable in the lower to mid epithelial layers (i.e. the spinous layers), and are in general, clearly visible in Haematoxylin and Eosin stains, although for some HPV types (e.g. HPV4) staining is poor and can give rise to a 'clear cell' appearance (Gross et al., 1982). These inclusion granules are the primary location of the HPV E4 proteins in such cutaneous lesions, with E4 antibodies staining these structures prominently (Fig. 1 (Breitburd and Croissant Orth; Croissant et al., 1985; Rogel-Gaillard et al., 1993; Peh et al., 2002; Doorbar et al., 1989)). The precise composition of these cytoplasmic structures is not

known, but analysis using chemical cross-linking agents that introduce reversible covalent links between proteins that are in close proximity to each other, suggests that they may be composed predominantly of E4, and that the great abundance of E4 in cutaneous lesions is likely to be attributed largely to the presence of these structures (Doorbar et al., 1996). E4 inclusion granules are a major contributor to the 'cytopathic effect' described by pathologists, with different HPV types showing different cytopathic morphologies (Rogel-Gaillard et al., 1993; Egawa et al., 1993; Gross et al., 1982). In HPV1, these structures have been extensively examined and are present as cytoplasmic (and nuclear) inclusion granules in cells supporting viral genome amplification immediately above the infected basal layer (see Fig. 1). The prominent HPV1 cytoplasmic inclusions probably associate transiently with the cytokeratin network during their formation (Rogel-Gaillard et al., 1992, 1993; Doorbar et al., 1996; Roberts et al., 2003). The less prominent nuclear HPV1 E4 inclusions (see Fig. 1) have been reported to be associated with promyelocytic leukemia protein (PML) at their periphery (Roberts et al., 2003). The E4 inclusion granules typically coalesce and increase in size as the infected cells are pushed towards the epithelial surface, eventually becoming less distinct as nuclear degeneration occurs in the upper cornified layers (Fig. 1, (Breitburd and Croissant Orth; Rogel-Gaillard et al., 1992, 1993)). Studies on the HPV1 E4 protein have shown that it becomes multiply phosphorylated during differentiation (Breitburd and Croissant Orth; Grand et al., 1989), and also becomes progressively cleaved from a 17 K full length species in the parabasal cell layers, to shorter 16 K (truncated before tyrosine 15), 11 K and 10 K species (truncated before alanine 59) that have lost sequences from their N-terminal region (Table 3, (Doorbar et al., 1996, 1988)). Interestingly, the 16 K HPV1 E4 protein has been reported to retard cell cycle progression at the G2 phase of the cell cycle by inhibiting the activity of Cdk1 (Knight et al., 2006), with the 17 and 16 K proteins working together to also inhibit cellular DNA replication (Knight et al., 2004, 2006) (Table 2). In cell-free culture systems, the HPV1 E4 proteins inhibit



**Fig. 3.** Expression of E4 proteins in epithelial cells; patterns and groupings. (A) The E4 proteins encoded by Alpha HPV types appear to have clear keratin-binding ability and an ability to reorganise the cellular keratin network (Roberts et al., 1993,1997; Doorbar et al., 1991; McIntosh et al., 2010; Wang et al., 2004). (B) The E4 proteins encoded by two Gamma HPV types (HPV4 and 65) show a partial association with keratins (Doorbar and Gallimore, 1989; McIntosh et al., 2010). The E4 proteins of these HPV types can form large cytoplasmic inclusions during natural infection (Egawa, 1994; Egawa et al., 1993). (C) Although the Mu HPV type, HPV1 can associate with keratins in some systems (Roberts et al., 1993), it appears to do this with a lower affinity, and assembles into cytoplasmic and nuclear inclusion granules (arrowed) both *in vitro* (Rogel-Gaillard et al., 1992,1993) and *in vivo* (Breitburd and Croissant Orth).



**Fig. 4.** Model of Keratin Reorganisation mediated by E1^E4. (A) Reorganisation of the cellular keratin network is a well-characterised consequence of HPV E1^E4 expression in epithelial cells. In the cell, cytokeratins exist primarily as insoluble keratin filaments as well as soluble monomers (i). Keratin filaments are converted to keratin monomers as a result of phosphorylation and association with the cellular 'solubility factor' 14-3-3, with soluble and insoluble keratins existing in a dynamic equilibrium (ii). In the presence of E1^E4 multimers, the keratin network becomes cross-linked, and although the 14-3-3 solubility factor can still bind, phosphorylated keratin monomers cannot be properly solubilised (iii) and eventually the filament network becomes disorganised and can collapse to the nuclear periphery. In the absence of E1^E4's C-terminal multimerisation region, some degree of keratin association is still apparent, but the keratin network is not cross-linked and the dynamics of keratin movement within the cell is not significantly compromised (iv). (B) During the papillomavirus life cycle, keratin association and other functions of E4 are carefully regulated. This has been most thoroughly investigated for the HPV16 E1^E4 protein (McIntosh et al., 2008; McIntosh et al., 2010; Wang et al., 2004, 2009; Khan et al., 2011). Our current thinking is that the unphosphorylated form of E4 binds keratins poorly, but that in the S-phase environment that precedes genome amplification, the E1^E4 protein is phosphorylated by MAPK and that keratin-association is enhanced (i). In the upper epithelial layers the E4 protein is cleaved at its N-terminus by Calpain, which favours the assembly into amyloid-like fibres (ii). It is thought that these E4 structures participate in keratin cross-linking, virus release and possibly also transmission (Khan et al., 2011).

loading of the cellular MCM helicases onto chromatin, and it has been suggested that cellular DNA replication may be similarly inhibited *in vivo* (Roberts et al., 2008). It is not yet clear however whether inhibition of cellular DNA replication occurs as part of any HPV life-cycle, although other viruses are known to use this strategy to favour replication of their genomes over that of the cell (Roberts et al., 2008). The full length HPV1 E1^E4 protein is not readily detectable once cells have exited the cell cycle post genome-amplification, and in these cells the truncated forms of E4 predominate (Table 3). Mutagenesis and structural analysis has localised E4's multimerisation-capability to the C-terminus of the protein (Roberts et al., 1994, 1997; Doorbar et al., 1996; Ashmole et al., 1998), and in the case of Alpha HPV types, has suggested that interactions are mediated through beta-aggregation (McIntosh et al., 2008). A second key function of E1^E4 that is thought

to extend to all HPV types is its ability to associate with cytokeratin filaments and to reorganise the cytokeratin filament network in the cell (Table 2 and Fig. 3, (Doorbar et al., 1991; Roberts et al., 1993)). For HPV1 E1^E4, this interaction requires the N-terminal E1^E4 amino acids that are lost in the truncated E4 proteins (Table 3). This region is predicted to form an amphipathic alpha helix that usually contains a 'leucine-cluster' domain (LLXLL or a motif derived from this) that is essential for keratin association (Table 3). In productive papillomas caused by HPV1, normal cytokeratin organisation is abolished, with differentiation-dependent keratins being almost undetectable in cells expressing E4 (Breitburd and Croissant Orth; Rogel-Gaillard et al., 1993). This dramatic reduction in keratin staining is a characteristic of infection by many if not all HPV types, including cutaneous and mucosal types from the Alpha genus (Doorbar et al., 1997;

McIntosh et al., 2010; Wang et al., 2004). Indeed, our knowledge of E1<sup>E4</sup> processing during differentiation is in fact best understood for HPV16, where both structural information and the nature of regulatory phosphorylation is well understood.

### *Mucosal Infections (Alpha HPV types)*

Detailed analysis of the high-risk E4 proteins has focused primarily on HPV16, HPV18 and 31 (Tables 1–3 and Fig. 4), and it is assumed that the E4 proteins of related HPV types will have similar mechanisms of regulation. Amongst low-risk Alpha types, the HPV11 E1<sup>E4</sup> protein has been most extensively studied (Table 3) (Brown et al., 1994, 2006; Bryan and Brown, 2000; Brown et al., 1996). E1<sup>E4</sup>-mediated disruption of the cellular cytokeratin network (see Figs. 3 and 4) was first shown in epithelial cells grown in monolayer culture (Doorbar et al., 1991; Doorbar, 1991), with later work confirming the importance of this interaction in HPV-associated tissue biopsies from patients (Doorbar et al., 1997; Wang et al., 2004, 2009). In such infected tissue, E1<sup>E4</sup>'s effect on cytokeratins is a re-organisation, generally to the cell periphery, rather than collapse to a perinuclear bundle as seen in dividing epithelial cells in culture (Doorbar et al., 1991; Roberts et al., 1997). These differences underlie the limitations of using monolayer culture to model events during epithelial differentiation, and indeed, most recent studies have made use of organotypic raft culture and the analysis of clinical material to support basic cell culture experiments (Doorbar et al., 1997; Wilson et al., 2005, 2007; Wang et al., 2004, 2009; Davy et al., 2005; Khan et al., 2011). As with the E1<sup>E4</sup> protein of HPV1, the HPV16, 18 and 31 E1<sup>E4</sup> proteins all contain a 'leucine-cluster' motif close to their N-terminus, which is important (along with upstream amino acids) for keratin association (see Tables 2 and 3). In the constrained structure identified for the full length 16 E1<sup>E4</sup>, this motif is thought to be largely unavailable for cytokeratin binding, and associates instead with the E4 multimerisation-motif located at the C-terminus of the protein ((McIntosh et al., 2008), Table 1, Fig. 4). The C-terminus of the HPV11 E1<sup>E4</sup> protein is similarly important for multimerisation (Bryan et al., 1998) and a common structural organisation may be widely conserved amongst E4 proteins (Fig. 2, Table 3). In the constrained form, the 16 E1<sup>E4</sup> protein (and very likely other E1<sup>E4</sup> proteins as well) has only a limited ability to bind keratins and to self-associate, and does not accumulate to any significant extent within the cell. Indeed, during natural infection, the E1<sup>E4</sup> protein is not readily apparent prior to the onset of viral genome amplification, but may be present at very low level in this constrained form. The dramatic accumulation of E1<sup>E4</sup> at or around the onset of genome amplification, coincides with the activation of the differentiation-dependent promoter (i.e. p670 in HPV16) and an elevation in E1<sup>E4</sup> transcript levels, as well as with regulatory post-translational modification as cell-cycle activity declines and the infected cell progresses through its final S-phase and G2-phase-like stages before exiting the cell cycle completely and succumbing to true terminal differentiation (Table 1, Fig. 4). The basis of how these post-translational modifications affect E4 function during productive infection (see Fig. 4) can be understood to some extent by considering the positions of the post-translationally modified amino acids within the HPV16 E1<sup>E4</sup> amino acid sequence (Table 1). In low-grade neoplasia (such as CIN1), late events such as genome amplification begin as E6/E7-mediated proliferative activity wanes (Isaacson Wechsler et al., 2012). In such cells, it appears that the S-phase kinase p42 ERK (Wang et al., 2009) (and possibly other members of the MAP kinase family (McIntosh et al., 2010)) phosphorylate 16 E1<sup>E4</sup> at amino acid Threonine 57, which adds a negative charge within a region of the protein that is already rich in negatively charged amino acids (McIntosh et al., 2008; Wang et al., 2009). A similar MAP kinase phosphorylation site has been identified in the HPV11 E1<sup>E4</sup> protein at position Threonine 53, and its loss compromises the normal

distribution of the protein within the cell ((Bryan et al., 2000), Table 3). For HPV16 E1<sup>E4</sup>, the consequence of this, as revealed by FRET analysis, is a tightening of the central loop of the protein, which facilitates keratin binding and the initial multimerisation of the E4 protein ((McIntosh et al., 2008), Table 1, Fig. 4), and it is likely that the HPV11 MAP kinase site regulates this protein similarly. The similar amino acid charge distribution seen across divergent E1<sup>E4</sup> sequences suggests a similar kinase-mediated regulation of structure, although for HPV18, the kinase responsible appears not to be p42 MAPK, which was unable to phosphorylate the 18 E1<sup>E4</sup> protein in *in vitro* assays ((Knight et al., 2011; Ding et al., 2013) see Tables 2 and 3). Experiments using antibodies that specifically recognise the HPV16 phospho T57 epitope suggest that this modification is however sustained only transiently during the productive virus life cycle, and is lost as the infected cell slips into the G2-like phase of the cell cycle where viral genome amplification is thought to occur (Table 1 and Fig. 4 (Wang et al., 2009)). For 16 E1<sup>E4</sup>, a second well-characterised phosphorylation event accompanies this S to G2-phase-like transition, and is mediated by the cytoplasmic kinase Cdk1 ((Davy et al., 2005, 2006), Table 1, Fig. 4). Although, several Cdk sites are predicted in the 16 E1<sup>E4</sup> protein, the most significant appears to be at position 32, which is located amongst a generally positively-charged stretch of amino acids in the N-terminal half of the protein (McIntosh et al., 2008, 2010; Wang et al., 2009). Although not located at precisely the same position, a Cdk1 site (at threonine 23) has been revealed experimentally in the HPV18 E1<sup>E4</sup> protein (Table 3 (Knight et al., 2011; Ding et al., 2013)). For HPV16 (and probably also HPV18 as well), the electrostatic interactions in this region are important for maintaining the constrained loop-structure of the E4 protein, which is thought to become more relaxed following CyclinB/Cdk1 (and possibly also CyclinA/Cdk2) phosphorylation (see Table 1 and Fig. 4). It is during this G2-like phase that E4 begins to accumulate to very high levels, facilitated by cleavage of sequences from its N-terminus in a manner akin to that described above for the HPV1 E4 proteins (see Table 1). For HPV16 E1<sup>E4</sup>, this is mediated in organotypic rafts (and probably also *in vivo*) by the cysteine protease calpain, which cleaves the 16 E1<sup>E4</sup> protein between amino acids 17 and 18 (Khan et al., 2011). The HPV 18 E1<sup>E4</sup> protein shows a similar susceptibility to calpain cleavage, which removes the proteins N-terminus and abolishes its constrained structure (McIntosh et al., 2008). As a result of this, the C-terminal multimerisation motifs identified in Alpha HPV types (Figs. 2 and 4, Tables 1 and 3) become available for self-association, allowing the cleaved E1<sup>E4</sup> protein to assemble into amyloid-like fibres which affect cell structure and cytokeratin organisation, and indeed beta-aggregation potential is predicted in all Alpha HPV E1<sup>E4</sup> proteins that we have examined (see Figs. 2 and 4). Although further analysis of post-translational modifications may reveal additional layers of complexity, it appears that as with other papillomavirus proteins (such as E1 for instance (Ma et al., 1999; Deng et al., 2004; Yu et al., 2007)), that E1<sup>E4</sup> function is regulated by the specific kinases and proteases that become activated as the infected cell migrates from the basal layer to the epithelial surface. Furthermore, the recognisable modular structure of E1<sup>E4</sup> that is apparent across different papillomavirus types suggests that the broad principles of E4 regulation may in fact be conserved, with differences (e.g. in kinases and phosphorylation sites) being driven by individual epithelial tropisms and routes of transmission.

## **The role of E4 in the virus life cycle**

### *Genome amplification*

Initial studies on the HPV1 E4 protein suggested involvement in virus assembly and/or release, and led to the proposition that the

protein might eventually be reclassified as the third late protein (L3 (Doorbar et al., 1986)). Although this classification was not adopted, it has since become clear that the major role for E4 lies in the late stages of the virus life cycle, and that these initial predictions may have been largely accurate. Although the gross effects on cyokeratin architecture clearly suggest a role in virus release, only a few studies have so far attempted to examine how cell disruption could facilitate the escape of infectious virions (Brown and Bryan, 2000; Brown et al., 2006; Bryan and Brown, 2001), with none yet addressing the possible effects on transmission success. So far, our understanding of E4 function has come from standard molecular analysis carried out in cells in monolayer culture, and from the analysis of E4's role during the virus life cycle using model systems and the analysis of biopsy material. A role for E4 in viral genome amplification, but not in the early stages of infection was first shown using mutant Cottontail Rabbit papillomavirus (CRPV) genomes that contain translation-termination linkers in the E4 ORF (Peh et al., 2004). Such genomes were not markedly compromised in their ability to produce papillomas, but when propagated in cottontail rabbits (which support the full CRPV productive cycle), the loss of E4 reduced the efficiency of genome amplification and the synthesis of capsid proteins (Peh et al., 2004). E4's role in genome amplification and the expression of capsid proteins, was subsequently shown for HPV16 (Nakahara et al., 2005), 31 (Wilson et al., 2005) and 18 (Wilson et al., 2007) using organotypic raft culture systems, and although our current work indicates that the protein is not in fact essential for these events, its presence certainly appears to optimise life-cycle completion. Such a role has not yet been demonstrated for low-risk HPV types however, although relevant work on these viruses has so far been limited (Fang et al., 2006; Fang et al., 2006; Thomas et al., 2001; Oh et al., 2004).

The precise mechanisms by which E4 contributes to genome amplification-success and capsid protein synthesis have not yet been clearly reported in the literature. However, from the analysis of known E4 functions, some important insights can be gleaned. Expression of both HPV16 and 18 E1<sup>E4</sup> in proliferating epithelial cells in monolayer culture causes a dramatic arrest in G2 (Davy et al., 2005; Nakahara et al., 2002; Davy et al., 2002), with similar findings also being reported for the HPV 1 E1<sup>E4</sup> proteins (Knight et al., 2004, 2006). In the case of HPV 16, G2-arrest appears to involve the sequestration of CyclinB/Cdk1 in the cytoplasm (through amino acids Threonine 22 and 23 in 16 E1<sup>E4</sup>; amino acids <sup>43</sup>RRL<sup>46</sup> in HPV18 E1<sup>E4</sup> (see Tables 2 and 3)), which is thought in its self to be sufficient to prevent nuclear translocation and the phosphorylation of proteins involved in mitotic progression. Sequestration of CyclinB/Cdk1 is not always apparent in neoplasia caused by high-risk Alpha types where life-cycle events are de-regulated, but is typically seen in benign papillomas caused by diverse low-risk types, in good agreement with the idea that E4 accumulation and genome amplification are G2-specific (Deng et al., 2004; Yu et al., 2007; Davy and Doorbar, 2007). For high-risk types, the G2 arrest function of E1<sup>E4</sup> may be expected to potentially inhibit E6/E7-mediated cell proliferation in the mid epithelial layers, and in this way contribute to the timely onset of vegetative viral genome amplification. Indeed, mutant HPV16 genomes expressing an 'arrest-deficient' form of E4 shows lower levels of genome amplification in organotypic raft culture, supporting this hypothesis (our unpublished data). For HPV18, the contribution of E4's G2-arrest function is apparently less significant (Knight et al., 2011), which may reflect differences in the extent of suprabasal E6/E7-driven cell proliferation in organotypic rafts induced by these two HPV types. Indeed, the ability of HPV11 to drive cell cycle entry, but not suprabasal cell proliferation, may partially explain why for this HPV type, the loss of E4 (and any potential G2-arrest capability) has only minimal effects on vegetative genome amplification (Fang et al., 2006). In addition to its dramatic effects on the cell cycle, it is clear that for HPV16 E1<sup>E4</sup> at least, that the protein might have other functions that are important for genome amplification but which are less well understood. These

include its association with cytoplasmic CyclinA/Cdk2 (via amino acids T22 and T23 in HPV16 and <sup>43</sup>RRL<sup>46</sup> in HPV 18 (Table 2)), which is thought to enhance 16 E1<sup>E4</sup>'s ability to inhibit mitotic progression (see Table 2 (Davy et al., 2006)), and the association with E2 (Davy et al., 2009), which is directly required for genome amplification (Table 2). The significance of the E2/E4 association during the HPV 16 life cycle is as yet unknown, but is likely to be carefully regulated during the productive cycle of the virus. In cell monolayer experiments however, co-expression of E2 and E4 leads to an elevation in the levels of both proteins, and to a partial sequestration of E2 in the cytoplasm (Davy et al., 2009). A final component of E4 function that may clearly contribute to the optimisation of genome amplification success is its effect on cellular kinases (Table 2, Fig. 4). The key kinases that modulate 16 E1<sup>E4</sup> function are the same as those that regulate E1 accumulation in the nucleus, and it is likely that E4's ability to sequester certain members of the MAP (McIntosh et al., 2010; Wang et al., 2009) and Cyclin-dependent kinase family (Davy et al., 2002, 2005; Nakahara et al., 2002) will result in the functional modification of diverse viral and cellular gene products including E1. Differences in kinase site position and susceptibility to phosphorylation seen in the E1<sup>E4</sup> protein of HPV18 (Knight et al., 2011; Ding et al., 2013) may reflect subtle differences in the epithelial tropism of HPV18 and related types. When considered together, it is perhaps not surprising that E4-loss in the context of the viral genome acts to compromise genome amplification success. In our hands however, E4 is certainly not essential for either genome amplification or virus assembly, prompting us to suspect that the E4 protein may have evolved primarily to fill other important functions.

#### *Virus assembly, virus release and transmission*

Although the E4 proteins of papillomaviruses are always primarily cytoplasmic, they can clearly be found in the nuclei too (Nicholls et al., 2001), and can sometimes associate with recognisable structures (Fig. 2 (Roberts et al., 2003)). It is not yet clear whether the presence of nuclear E4 is an unavoidable consequence of its high abundance in the cell, or whether the nuclear E4 fraction represents a particular functional form of the protein that is necessary for a particular activity such as association with E2 (Davy et al., 2009), or with viral or cellular proteins necessary for virus assembly and/or maturation. Of significant importance however when considering why E4 accumulates late in infection, is the observation that the expression of capsid proteins appears to only occur in cells that are already expressing E4 (Brown et al., 1994,1995), and that during the productive life cycle, E4 expression precedes L2 expression, which in itself appears to precede-slightly the expression of L1 (Florin et al., 2002). Thus the order of expression of viral gene products is arranged in such a way that infectious virus particles are only ever produced in E4-positive cells. Although the role of E4 in virus assembly may initially suggest a need for nuclear E4, we should also remember that L1 and L2 are expressed in the cytoplasm before locating to the nucleus, and that the initial assembly of L1 molecules into capsomeres is a cytoplasmic event (Bird et al., 2008). Although associations between E4, L1 and L2 can be demonstrated *in vitro*, it appears that E4 loss does not necessarily compromise virus assembly significantly in the organotypic raft system (our unpublished observations). Importantly however, the nuclear assembly of virus particles is followed eventually by nuclear degeneration and the release of assembled virions into the terminally-differentiating cell which contains abundant E4 and a reduced level of cellular cyokeratins. Work carried out by Brown and co-workers has revealed important defects in the cornified envelope in cells expressing the HPV11 E1<sup>E4</sup> protein, (Brown et al., 2006; Bryan and Brown, 2000; Lehr et al., 2002, 2004) and has clearly highlighted the point that the major L1 capsid protein is expressed from a bicistronic mRNA that contains E1<sup>E4</sup> as the first cistron (Brown et al., 1996). The natural follow-on from this is that the E1<sup>E4</sup>

proteins, which in the upper epithelial layers are present in their N-terminally truncated form as amyloid fibres, in some way serve to optimise the transmission success of virions shed from the epithelial surface in virus-laden squames. Driven largely by the need to fully understand how prophylactic vaccination offers protection against infection, a considerable effort has already been put into the study of virus attachment and the mechanisms of virus entry, but only few studies have yet made use of virions produced in differentiating epithelial cells, and none have yet addressed in any significant way the contribution of E1<sup>E4</sup> to this process. The absence of convenient epithelial systems for the study of virus production and maturation in the presence or absence of E4 have so far limited such important studies.

### The modular structure of the E4 proteins suggests conserved functions

While most of the above analysis focuses on the E4 proteins of the prominent high-risk HPV Alpha types, the comparison with HPV1 E1<sup>E4</sup> at a functional level, and the analysis of sequence similarities between these E1<sup>E4</sup> proteins and those of other papillomaviruses, suggests that while E4 proteins are highly divergent at the primary amino acid sequence level, they are much more obviously conserved when broad 'structural' modules are compared (Tables 2 and 3). When large-scale alignment is carried out, it is clear that the most significant regions of similarity lie at the N-terminus, where a predicted amphipathic alpha helix and the leucine-cluster motifs reside (see alignment in Figs. S1 and S2), and at the C-terminus, where sequences involved in E4/E4 self-association are primarily located (Table 3). These C-terminal sequences of the Alpha HPV types typically have high amyloidogenic/Beta aggregation potential and can be identified as such using amyloid peptide prediction programs such as Tango (McIntosh et al., 2008). For HPV16 and HPV1 E1<sup>E4</sup>, the loss of the C-terminal region and the binding-avidity that this provides (McIntosh et al., 2008; Wang et al., 2004), prevents prominent E1<sup>E4</sup> accumulation in the cytoplasm (Roberts et al., 1994,1997), and can a redistribution of E1<sup>E4</sup> to the nucleus, despite the retention of N-terminal motifs that mediate direct keratin binding (Fig. 4 (Wang et al., 2009)). Although the affinity of E1<sup>E4</sup>/keratin binding has not yet been measured, the leucine-cluster alone is insufficient to mediate specific cytokeratin-binding, but instead directs reporter molecules such as GFP to the mitochondria (Raj et al., 2004). For HPV16 E1<sup>E4</sup>, cytokeratin-association requires the leucine-cluster as well as upstream amino acids, and it is likely that specificity is controlled in a similar way amongst other HPV types (Table 3). Thus it appears that despite variation at the level of primary amino acid sequence, the E1<sup>E4</sup> proteins of all HPV types contain clearly recognisable 'functional motifs' at their N and C-termini. Variation between E4 sequences lie primarily in their length, which is mediated for the most part by an extension of the central portion of the E1<sup>E4</sup> molecule. In the case of the HPV16 E1<sup>E4</sup> protein, this central region (which is surrounded by proline-rich segments) comprises a generally positively-charged region, a loop-region, and a generally negatively-charged region which form a folded structure ((McIntosh et al., 2008), Fig. 3). These defined structural elements are recognisable not only in the putative E1<sup>E4</sup> proteins of related HPV types, but also in the E1<sup>E4</sup> proteins from other genera, suggesting that E1<sup>E4</sup> proteins from diverse HPVs may have similar mechanisms of regulation. Interestingly, the loop region (where the amino acid sequence undergoes a turn) contains the major immune dominant epitopes within the protein, and is the site to which many of the most useful E1<sup>E4</sup> monoclonal antibodies bind (Fig. 3). It is possible that this region in particular is available for interaction with cellular targets in the context of the full-length protein.

### Other E4 species generated by splicing and proteolytic cleavage

In general, sequence alignment carried out using the predicted products of the full length E4 ORF show recognisable conservation only downstream of the E1<sup>E4</sup> splice acceptor site (Doorbar and Myers, 1996). Such observations suggest that sequences upstream of the splice acceptor are not in fact E4 coding sequences, and indeed, there is no generally conserved ATG codon at the start of all E4 reading frames to suggest this. It seems instead that in most cases, the abundant E4 proteins are expressed from a spliced E1<sup>E4</sup> mRNA transcript, and that the initiation codon for E4 gene expression comes from the E1 ORF. It also appears, for both HPV16 and HPV1 E4, that smaller E4 species with modified functions can be generated by proteolytic cleavage as the infected cell migrates towards the epithelial surface, and that these can be further modified by phosphorylation (Breitburd and Croissant Orth; Doorbar et al., 1988; Roberts et al., 1993,1997; Doorbar et al., 1996; McIntosh et al., 2008; Wang et al., 2004; Khan et al., 2011). Although an abundant E1<sup>E4</sup> transcript has been identified in most papillomaviruses that have so far been examined (Chow et al., 1987a,b; Doorbar et al., 1990; Milligan et al., 2007; Wang et al., 2011; Ozbun and Meyers, 1997; Nasseri et al., 1987; Baker and Howley, 1987; Barksdale and Baker, 1995; Palermo-Dilts et al., 1990), recent work has suggested that there may in fact be additional E4 gene-products expressed from distinct transcripts. These including a putative E6<sup>E4</sup> gene product in HPV16 (Milligan et al., 2007), as well as two distinct E2<sup>E4</sup> polypeptides (E2<sup>E4</sup>S and E2<sup>E4</sup>L) (Tan et al., 2012) and an alternative form of E1<sup>E4</sup> (E1<sup>E4</sup>S) in HPV18 (Kho et al., 2013). Such transcripts are generally low-abundance when compared to the primary E1<sup>E4</sup> species, with their pattern of expression and significance during productive infection not yet known. Interestingly, the E4 ORF encodes not only the hinge domain of E2, but in the case of HPV16, it also contains a series of ASF/SF2 sites that regulate usage of the splice acceptor at position 3358 (SA3358 (Somberg and Schwartz, 2010)). SA3358 is in itself a poor splice acceptor, despite being efficiently used during the HPV life-cycle to direct the production of the abundant E1<sup>E4</sup> mRNA. Current thinking suggests that increasing ASF/SF2 acts to down regulate the activity of SA2709 (which is required for E2 expression) in order to favour the use of SA3358 which allows expression of the abundant E1<sup>E4</sup> transcripts, and it has been suggested that the eventual decline in ASF/SF2 binding in the upper layers of the epidermis may eventually facilitate the switch to SA5639 usage and allow the expression of capsid proteins (Somberg and Schwartz, 2010; Rush et al., 2005; Johansson et al., 2012).

### Summary

The functions of the papillomavirus E4 proteins in the virus life cycle have been difficult to resolve. Although located in the early part of the viral genome, the E4 proteins are primarily expressed during the late stages of infection, at or around the time that genome amplification is initiated. In general, the E4 proteins of different papillomaviruses share a recognisable modular organisation despite showing divergence at the primary amino acid sequence level (Table 3), and can accumulate to very high levels in the upper epithelial layers where virus particles assemble. Any consideration of E4 function must take this into account, and it has been suggested that E4's ability to disrupt the cellular keratin network and the formation of the cornified envelope may facilitate virus release and/or transmission (Figs. 1–4 (Bryan and Brown 2000,2001)). Formal validation of the extent to which E4 facilitates virus transmission and infection remains to be carried out. In addition to this potentially significant role, the E4 proteins of several HPV types have been shown to inhibit cell proliferation in G2 (Davy and Doorbar, 2007), and to participate in efficient genome amplification, and perhaps as a consequence of this,

virus synthesis as well (Peh et al., 2004; Nakahara et al., 2005; Wilson et al., 2005). The contribution of E4 to genome amplification has not yet been fully worked out, but may involve E4's interaction with kinases, its association with E2, and its cell cycle-arrest capabilities. The abundance of E4 in lesional tissue has recently suggested a useful role as a HPV-specific biomarker, particularly when used in combination with cellular markers of deregulated viral gene expression (Griffin et al., 2012). The preservation of the E4 ORF across diverse papillomavirus types, coupled with its abundance suggests an evolutionary important role in improving viral fitness. It appears that the E4 proteins may have multiple roles in the virus life-cycle, but that it plays an important role in ensuring efficient virus release and transmission.

## Acknowledgments

Thanks to Heather Griffin, Woei Ling Peh and Pauline McIntosh (NIMR) for providing some of the immunofluorescence and electron micrograph images used in this publication (Figs. 1–3), and to Koenraad van Doorslayer (NAAID, Bethesda, USA) for producing the very large E1<sup>^</sup>E4 amino acid sequence alignment shown in Fig. S2. John Doorbar is funded by the United Kingdom Medical Research Council.

## Appendix A. Supplementary information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.virol.2013.07.008>.

## References

- Ashmole, I., Gallimore, P.H., Roberts, S., 1998. Identification of conserved hydrophobic C-terminal residues of the human papillomavirus type 1 E1E4 protein necessary for E4 oligomerisation in vivo. *Virology* 240 (2), 221–231.
- Baker, C.C., Howley, P.M., 1987. Differential promoter utilization by the bovine papillomavirus in transformed cells and in productively infected wart tissues. *EMBO J.* 6, 1027–1035.
- Barksdale, S.K., Baker, C.C., 1995. Differentiation specific alternative splicing of bovine papillomavirus late mRNAs. *J. Virol.* 69, 6553–6556.
- Bell, I., Martin, A., Roberts, S., 2007. The E1<sup>^</sup>E4 protein of human papillomavirus interacts with the serine-arginine-specific protein kinase SRPK1. *J. Virol.* 81 (11), 5437–5448.
- Bird, G., et al., 2008. Possible role for cellular karyopherins in regulating polyomavirus and papillomavirus capsid assembly. *J. Virol.* 82 (20), 9848–9857.
- Borgogna, C., et al., 2012. Characterization of beta papillomavirus E4 expression in tumours from Epidermodysplasia Verruciformis patients and in experimental models. *Virology* 423 (2), 195–204.
- Breitburd, F., O. Croissant, G. Orth. Expression of human papillomavirus type-1 E4 gene products in warts. In: *Cancer Cells* 51987, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York. pp. 115–122.
- Brown, D.R., Bryan, J.T., 2000. Abnormalities of cornified cell envelopes isolated from human papillomavirus type 11-infected genital epithelium. *Virology* 271 (1), 65–70.
- Brown, D.R., et al., 1991. Detection of human papillomavirus types 6 and 11 E4 gene products in condylomata acuminatum. *J. Med. Virol.* 34, 20–28.
- Brown, D.R., et al., 1994. Colocalization of human papillomavirus type 11 E1<sup>^</sup>E4 and L1 proteins in human foreskin implants grown in athymic mice. *Virology* 201, 46–54.
- Brown, D.R., et al., 1995. Human papillomavirus type 11 E1<sup>^</sup>E4 and L1 proteins colocalize in the mouse xenograft system at multiple time points. *Virology* 214, 259–263.
- Brown, D.R., et al., 1996. Virus-like particles and E1<sup>^</sup>E4 protein expressed from the human papillomavirus type 11 bicistronic E1<sup>^</sup>E4<sup>^</sup>L1 transcript. *Virology* 222, 43–50.
- Brown, D.R., Brown, C.R., Lehr, E.E., 2004. Intracellular expression patterns of the human papillomavirus type 59 E1/E4 protein in COS cells, keratinocytes, and genital epithelium. *Intervirology* 47 (6), 321–327.
- Brown, D.R., et al., 2006. The human papillomavirus type 11 E1<sup>^</sup>E4 protein is a transglutaminase 3 substrate and induces abnormalities of the cornified cell envelope. *Virology* 345 (1), 290–298.
- Bryan, J.T., Brown, D.R., 2000. Association of the human papillomavirus type 11 E1<sup>^</sup>E4 protein with cornified cell envelopes derived from infected genital epithelium. *Virology* 277 (2), 262–269.
- Bryan, J.T., Brown, D.R., 2001. Transmission of human papillomavirus type 11 infection by desquamated cornified cells. *Virology* 281 (1), 35–42.
- Bryan, J.T., Fife, K.H., Brown, D.R., 1998. The intracellular expression pattern of the human papillomavirus type 11 E1–E4 protein correlates with its ability to self associate. *Virology* 241 (1), 49–60.
- Bryan, J.T., et al., 2000. The human papillomavirus type 11 E1E4 protein is phosphorylated in genital epithelium. *Virology* 268 (2), 430–439.
- Chen, E.Y., et al., 1982. The primary structure and genetic organization of the bovine papillomavirus type 1 genome. *Nature* 299 (5883), 529–534.
- Chow, L.T., et al., 1987a. Identification and mapping of human papillomavirus type 1 RNA transcripts recovered from plantar warts and infected epithelial cell cultures. *J. Virol.* 61, 1913–1918.
- Chow, L.T., et al., 1987b. Human papillomavirus types 6 and 11 mRNAs from genital condylomata acuminata. *J. Virol.* 61, 2581–2588.
- Croissant, O., Breitburd, F., Orth, G., 1985. Specificity of the cytopathic effect of cutaneous human papillomaviruses. *Clin. Dermatol.* 3 (4), 43–55.
- Danos, O., Katinka, M., Yaniv, M., 1982. human papillomavirus 1a complete DNA sequence: a novel type of genome organisation amongst papovaviridae. *EMBO J.* 1, 231–236.
- Davy, C., Doorbar, J., 2007. G2/M cell cycle arrest in the life cycle of viruses. *Virology* 368 (2), 219–226.
- Davy, C., et al., 2009. A novel interaction between the human papillomavirus type 16 E2 and E1<sup>^</sup>E4 proteins leads to stabilization of E2. *Virology* 394 (2), 266–275.
- Davy, C.E., et al., 2002. Identification of a G(2) arrest domain in the E1<sup>^</sup>E4 protein of human papillomavirus type 16. *J. Virol.* 76 (19), 9806–9818.
- Davy, C.E., et al., 2002. Identification of a G(2) arrest domain in the E1<sup>^</sup>E4 protein of human papillomavirus type 16. *J. Virol.* 76 (19), 9806–9818.
- Davy, C.E., et al., 2005. Human papillomavirus type 16 E1<sup>^</sup>E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/Cyclin B1 complexes. *J. Virol.* 79 (7), 3998–4011.
- Davy, C.E., et al., 2006. HPV16 E1<sup>^</sup>E4 protein is phosphorylated by Cdk2/cyclin A and relocalizes this complex to the cytoplasm. *Virology*.
- Deng, W., et al., 2004. Cyclin/CDK regulates the nucleocytoplasmic localization of the human papillomavirus E1 DNA helicase. *J. Virol.* 78 (24), 13954–13965.
- Ding, Q., Li, L., Whyte, P., 2013. Human papillomavirus 18 E1<sup>^</sup>E4 protein interacts with cyclin A/CDK 2 through an RXL motif. *Mol. Cell. Biochem.* 373 (1–2), 29–40.
- Donaldson, M.M., Boner, W., Morgan, I.M., 2007. TopBP1 regulates human papillomavirus type 16 E2 interaction with chromatin. *J. Virol.* 81 (8), 4338–4342.
- Doorbar, J., 1991. An Emerging Function for E4. *Papillomavirus Report*, 2(6): 145–148.
- Doorbar, J., Cubie, H., 2005. Molecular basis for advances in cervical screening. *Mol. Diagn.* 9 (3), 129–142.
- Doorbar, J., Gallimore, P.H., 1987. Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. *J. Virol.* 61 (9), 2793–2799.
- Doorbar, J., et al., 1986. Identification of the human papilloma virus-1a E4 gene products. *EMBO J.* 5 (2), 355–362.
- Doorbar, J., et al., 1988. Analysis of HPV-1 E4 gene expression using epitope-defined antibodies. *EMBO J.* 7 (3), 825–833.
- Doorbar, J., Coneron, I., Gallimore, P.H., 1989. Sequence divergence yet conserved physical characteristics among the E4 proteins of cutaneous human papillomaviruses. *Virology* 172 (1), 51–62.
- Doorbar, J., et al., 1990. Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. *Virology* 178 (1), 254–262.
- Doorbar, J., et al., 1991. Specific interaction between HPV-16 E1<sup>^</sup>E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature* 352 (6338), 824–827.
- Doorbar, J., Medcalf, E., Naphthine, S., 1996. Analysis of HPV1 E4 complexes and their association with keratins in vivo. *Virology* 218 (1), 114–126.
- Doorbar, J., et al. Correlation of E4 expression with the onset of other late functions in vivo using high affinity monoclonals prepared by phage display. In: 15th International Papillomavirus Workshop. 1996. Brisbane, Australia.
- Doorbar, J., Myers, G., 1996. The E4 protein. In: Myers, G., et al. (Eds.), *Human Papillomaviruses* 1996. Los Alamos National Laboratory, Los Alamos, New Mexico, USA, pp. 58–80.
- Doorbar, J., et al., 1997. Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4. *Virology* 238 (1), 40–52.
- Doorbar, J., et al., 2000. The E1<sup>^</sup>E4 protein of human papillomavirus type 16 associates with a putative RNA helicase through sequences in its C terminus. *J. Virol.* 74 (21), 10081–10095.
- Doorbar, J.C.I., Gallimore, P.H., 1989. Sequence divergence yet conserved physical characteristics among the E4 proteins of cutaneous human papillomaviruses. *Virology* 172 (1), 51–62.
- Egawa, K., 1994. New types of human papillomaviruses and intracytoplasmic inclusion bodies: a classification of inclusion warts according to clinical features, histology and associated HPV types. *Br. J. Dermatol.* 130, 158–166.
- Egawa, K., et al., 1993. Two novel types of human papillomavirus, HPV63 and HPV65: comparison of their clinical and histological features and DNA sequences to other HPV types. *Virology* 194, 51–62.
- Fang, L., et al., 2006. Induction of productive human papillomavirus type 11 life cycle in epithelial cells grown in organotypic raft cultures. *Virology* 347 (1), 28–35.
- Fang, L., et al., 2006. The human papillomavirus type 11 E1/E4 protein is not essential for viral genome amplification. *Virology* 351 (2), 271–279.
- Flores, E.R., et al., 1999. Establishment of the human papillomavirus type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line. *Virology* 262 (2), 344–354.
- Florin, L., et al., 2002. Assembly and translocation of papillomavirus capsid proteins. *J. Virol.* 76 (19), 10009–10014.

- Frattoni, M.G., et al., 1997. Induction of human papillomavirus type 18 late gene expression and genomic amplification in organotypic cultures from transfected DNA templates. *J. Virol.* 71 (9), 7068–7072.
- Gissmann, L., Pfister, H., Zur Hausen, H., 1977. Human papilloma viruses (HPV): characterization of four different isolates. *Virology* 76 (2), 569–580.
- Grand, R.J., et al., 1989. Phosphorylation of the human papillomavirus type 1 E4 proteins in vivo and in vitro. *Virology* 170 (1), 201–213.
- Griffin, H., et al., 2012. E4 antibodies facilitate detection and type-assignment of active HPV infection in cervical disease. *PLoS One* 7 (12), e49974.
- Gross, G., et al., 1982. Correlation between human papillomavirus (HPV) type and histology of warts. *J. Invest. Dermatol.* 78, 160–164.
- Hummel, M., Hudson, J.B., Laimins, L.A., 1992. Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes. *J. Virol.* 66, 6070–6080.
- Isaacson Wechsler, E., et al., 2012. Reconstruction of human papillomavirus type 16-mediated early-stage neoplasia implicates E6/E7 deregulation and the loss of contact inhibition in neoplastic progression. *J. Virol.* 86 (11), 6358–6364.
- Jablonska, S., et al., 1985. Cutaneous warts. Clinical, histological, and virological correlations. In: Jablonska, S., Orth, G. (Eds.), *Clin. Dermatol.*, 3. J. B. Lippincott, Philadelphia, pp. 71–82.
- Johansson, C., et al., 2012. HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. *EMBO J.*
- Katoh, K., et al., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30 (14), 3059–3066.
- Khan, J., et al., 2011. Role of calpain in the formation of human papillomavirus type 16 E1<sup>E4</sup> amyloid fibers and reorganization of the keratin network. *J. Virol.* 85 (19), 9984–9997.
- Kho, E.Y., et al., HPV-18 E6 mutants reveal p53 modulation of viral DNA amplification in organotypic cultures. In: *Proceedings of the National Academy of Sciences of the United States of America*, 2013. 110(19): pp. 7542–9.
- Knight, G.L., et al., 2004. Cooperation between different forms of the human papillomavirus type 1 E4 protein to block cell cycle progression and cellular DNA synthesis. *J. Virol.* 78 (24), 13920–13933.
- Knight, G.L., Turnell, A.S., Roberts, S., 2006. Role for Wee1 in inhibition of G2-to-M transition through the cooperation of distinct human papillomavirus type 1 E4 proteins. *J. Virol.* 80 (15), 7416–7426.
- Knight, G.L., Turnell, A.S., Roberts, S., 2006. Role for Wee1 in inhibition of G2-to-M transition through the cooperation of distinct human papillomavirus type 1 E4 proteins. *J. Virol.* 80 (15), 7416–7426.
- Knight, G.L., et al., 2011. A cyclin-binding motif in human papillomavirus type 18 (HPV18) E1<sup>E4</sup> is necessary for association with CDK-cyclin complexes and G2/M cell cycle arrest of keratinocytes, but is not required for differentiation-dependent viral genome amplification or L1 capsid protein expression. *Virology* 412 (1), 196–210.
- Laurent, R., et al., 1982. Two anatomical types of warts with plantar location: specific cytopathic effects of papillomavirus type 1 (HPV1) and type 2 (HPV2). *Arch. Dermatol. Res.* 274, 101–111.
- Lehr, E., Jarnik, M., Brown, D.R., 2002. Human papillomavirus type 11 alters the transcription and expression of loricrin, the major cell envelope protein. *Virology* 298 (2), 240–247.
- Lehr, E., et al., 2004. Infection with Human Papillomavirus alters expression of the small proline rich proteins 2 and 3. *J. Med. Virol.* 72 (3), 478–483.
- Ma, T., et al., 1999. Interaction between cyclin-dependent kinases and human papillomavirus replication-initiation protein E1 is required for efficient viral replication. *Proc. Nat. Acad. Sci. U.S.A.* 96 (2), 382–387.
- Maglennon, G.A., McIntosh, P., Doorbar, J., 2011. Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology* 414 (2), 153–163.
- McIntosh, P.B., et al., 2008. Structural analysis reveals an amyloid form of the human papillomavirus type 16 E1<sup>E4</sup> protein and provides a molecular basis for its accumulation. *J. Virol.* 82 (16), 8196–8203.
- McIntosh, P.B., et al., 2010. E1<sup>E4</sup>-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium. *J. Cell Sci.* 123 (16), 2810–2822.
- Middleton, K., 2003. Analysis of papillomavirus E1<sup>E4</sup> expression with respect to epithelial proliferation and differentiation in productive and neoplastic papillomavirus lesions., in *Virology*. University of London, London p. 270.
- Middleton, K., et al., 2003. Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers. *J. Virol.* 77 (19), 10186–10201.
- Milligan, S.G., et al., 2007. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* 360 (1), 172–181.
- Nakahara, T., et al., 2002. Modulation of the cell division cycle by human papillomavirus type 18 E4. *J. Virol.* 76 (21), 10914–10920.
- Nakahara, T., et al., 2002. Modulation of the cell division cycle by human papillomavirus type 18 E4. *J. Virol.* 76 (21), 10914–10920.
- Nakahara, T., et al., 2005. Human papillomavirus type 16 E1<sup>E4</sup> contributes to multiple facets of the papillomavirus life cycle. *J. Virol.* 79 (20), 13150–13165.
- Nasseri, M., et al., 1987. A human papillomavirus type 11 transcript encoding an E1<sup>E4</sup> protein. *Virology* 159, 433–439.
- Neary, K., Horwitz, B.H., DiMaio, D., 1987. Mutational analysis of open reading frame E4 of bovine papillomavirus type 1. *J. Virol.* 61 (4), 1248–1252.
- Nicholls, P., et al., 2001. Detection of viral DNA and E4 protein in basal keratinocytes of experimental canine oral papillomavirus lesions. *Virology*, 82–98283, 82–98.
- Nicholls, P.K., et al., 2001. Detection of viral DNA and E4 protein in basal keratinocytes of experimental canine oral papillomavirus lesions. *Virology* 284 (1), 82–98.
- Oh, S.T., Longworth, M.S., Laimins, L.A., 2004. Roles of the E6 and E7 proteins in the life cycle of low-risk human papillomavirus type 11. *J. Virol.* 78 (5), 2620–2626.
- Ozbun, M.A., Meyers, C., 1997. Characterisation of late gene transcripts expressed during vegetative replication of human papillomavirus type 31b. *J. Virol.* 71 (7), 5161–5172.
- Palermo-Dilts, D.A., Broker, T.R., Chow, L.T., 1990. Human Papillomavirus type 1 produces redundant as well as polycistronic mRNAs in plantar warts. *J. Virol.* 64, 3144–3149.
- Peh, W., et al. Production of papillomas in cottontail and domestic rabbits using CRPV knockout mutants reveals an essential role for E4 in viral replication and capsid synthesis. In: 19th International Papillomavirus Conference. 2001. Florianopolis, Brazil.
- Peh, W.L., et al., 2002. Life cycle heterogeneity in animal models of human papillomavirus-associated disease. *J. Virol.* 76 (20), 10401–10416.
- Peh, W.L., et al., 2004. The viral E4 protein is required for the completion of the cottontail rabbit papillomavirus productive cycle in vivo. *J. Virol.* 78 (4), 2142–2151.
- Pray, T.R., Laimins, L.A., 1995. Differentiation-dependent expression of E1–E4 proteins in cell lines maintaining episomes of human papillomavirus type 31b. *Virology* 206 (1), 679–685.
- Raj, K., et al., 2004. E1<sup>E4</sup> protein of human papillomavirus type 16 associates with mitochondria. *J. Virol.* 78 (13), 7199–7207.
- Roberts, S., et al., 1993. Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells. *Virology* 197 (1).
- Roberts, S., et al., 1993. Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells. *Virology*, 176–187197, 176–187.
- Roberts, S., et al., 1994. Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. *J. Virol.* 68 (10), 6432–6445.
- Roberts, S., et al., 1997. Mutational analysis of the human papillomavirus type 16 E1<sup>E4</sup> protein shows that the C terminus is dispensable for keratin cytoskeleton association but is involved in inducing disruption of the keratin filaments. *J. Virol.* 71 (5), 3554–3562.
- Roberts, S., et al., 2003. The ND10 component promyelocytic leukemia protein relocates to human papillomavirus type 1 E4 intranuclear inclusion bodies in cultured keratinocytes and in warts. *J. Virol.* 77 (1), 673–684.
- Roberts, S., et al., 2008. Identification of an arginine-rich motif in human papillomavirus type 1 E1/E4 protein necessary for E4-mediated inhibition of cellular DNA synthesis in vitro and in cells. *J. Virol.* 82 (18), 9056–9064.
- Rogel-Gaillard, C., Breitburd, F., Orth, G., 1992. Human papillomavirus type 1 E4 proteins differing by their N-terminal ends have distinct cellular localizations when transiently expressed in vitro. *J. Virol.* 66 (2), 816–823.
- Rogel-Gaillard, C., et al., 1993. Cytopathic effect in human papillomavirus type 1-induced inclusion warts: in vitro analysis of the contribution of two forms of the viral E4 protein. *J. Invest. Dermatol.* 101, 843–851.
- Rush, M., Zhao, X., Schwartz, S., 2005. A splicing enhancer in the e4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression. *J. Virol.* 79 (18), 12002–12015.
- Somberg, M., Schwartz, S., 2010. Multiple ASF/SF2 sites in the human papillomavirus type 16 (HPV-16) E4-coding region promote splicing to the most commonly used 3'-splice site on the HPV-16 genome. *J. Virol.* 84 (16), 8219–8230.
- Sorathia, R., 2005. Identification of human papillomavirus type16 (HPV16) E1<sup>E4</sup> binding partners and the characterisation of the E1<sup>E4</sup>/E2 interaction, National Institute for Medical Research. University of London, London.
- Tan, C.L., et al., 2012. HPV-18 E2<sup>E4</sup> chimera: 2 new spliced transcripts and proteins induced by keratinocyte differentiation. *Virology* 429 (1), 47–56.
- Thomas, J.T., et al., 2001. Cellular changes induced by low-risk human papillomavirus type 11 in keratinocytes that stably maintain viral episomes. *J. Virol.* 75 (16), 7564–7571.
- Wang, Q., et al., 2004. Functional analysis of the human papillomavirus type 16 E1<sup>E4</sup> protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *J. Virol.* 78 (2), 821–833.
- Wang, Q., et al., 2009. Phosphorylation of the human papillomavirus type 16 E1<sup>E4</sup> protein at T57 by ERK triggers a structural change that enhances keratin binding and protein stability. *J. Virol.* 83 (8), 3668–3683.
- Wang, X., et al., 2011. Construction of a full transcription map of human papillomavirus type 18 during productive viral infection. *J. Virol.* 85 (16), 8080–8092.
- Wilson, R., Fehrmann, F., Laimins, L.A., 2005. Role of the E1<sup>E4</sup> protein in the differentiation-dependent life cycle of human papillomavirus type 31. *J. Virol.* 79 (11), 6732–6740.
- Wilson, R., et al., 2007. The full-length E1<sup>E4</sup> protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression. *Virology* 362 (2), 453–460.
- Yu, J.H., et al., 2007. Mitogen-activated protein kinases activate the nuclear localization sequence of human papillomavirus type 11 E1 DNA helicase to promote efficient nuclear import. *J. Virol.* 81 (10), 5066–5078.