Virology 445 (2013) 80–98

Contents lists available at ScienceDirect

Virology

journal homepage: 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^E4) 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^E4 mRNA that has been most extensively analysed.

During the papillomavirus life cycle, the E1^E4 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^E4, these kinases regulate one of the E1^E4 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.

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^E4 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^E4 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

0042-6822
http://dx.doi.org/10.1016/j.virol.2013.07.008

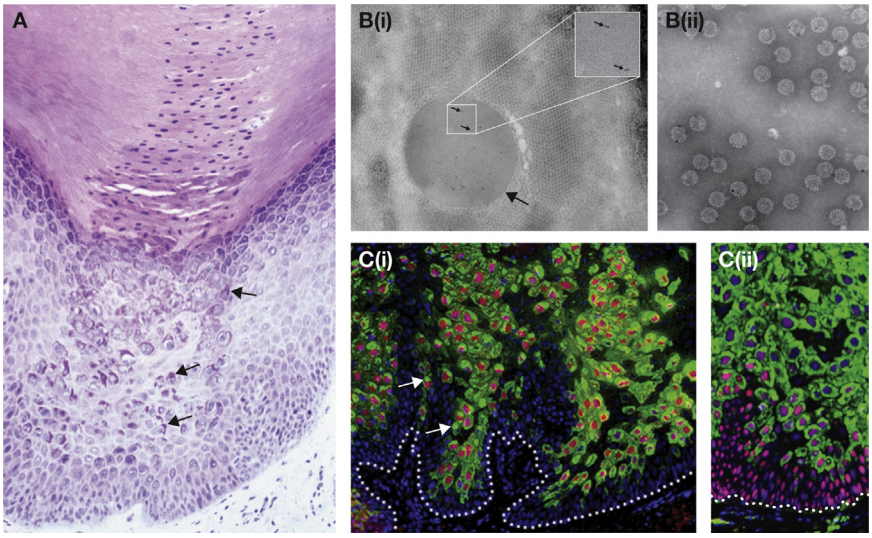

Fig. 1. E4 inclusion granules are typically found in cutaneous HPV types (Mu, Beta, Gamma Genera). Cutaneous HPV types from the Mu, Beta and Gamma Genera typically accumulate their E4 proteins to high level as inclusion granules visible by Hemotoxylin & Eosin staining (granular structures arrowed in (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;