JOURNAL OF VIROLOGY, Oct. 1994, p. 6432–6445
0022-538X/94/$04.00+0

Vol. 68, No. 10

Mutational Analysis of Human Papillomavirus E4 Proteins: Identification of Structural Features Important in the Formation of Cytoplasmic E4/Cytokeratin Networks in Epithelial Cells

SALLY ROBERTS,[1]* IAN ASHMOLE,[1] LORETTA J. GIBSON,[1] SUSAN M. ROOKES,[1] GEOFF J. BARTON,[2] AND PHILLIP H. GALLIMORE[1]

Cancer Research Campaign Laboratories, Department of Cancer Studies, The Medical School, University of Birmingham, Birmingham B15 2TJ,[1] and Laboratory of Molecular Biophysics, University of Oxford, Oxford OX1 3QU,[2] United Kingdom

Received 19 May 1994/Accepted 8 July 1994

We have previously demonstrated that human papillomavirus type 1 (HPV 1) and 16 (HPV 16) E4 proteins form cytoplasmic filamentous networks which specifically colocalize with cytokeratin intermediate-filament (IF) networks when expressed in simian virus 40-transformed keratinocytes. The HPV 16 (but not the HPV 1) E4 protein induced the collapse of the cytokeratin networks. (S. Roberts, I. Ashmole, G. D. Johnson, J. W. Kreider, and P. H. Gallimore, Virology 197:176–187, 1993). The mode of interaction of E4 with the cytokeratin IFs is unknown. To identify E4 sequences important in mediating this interaction, we have constructed a large panel of mutant HPV (primarily HPV 1) E4 proteins and expressed them by using the same simian virus 40-epithelial expression system. Mutation of HPV 1 E4 residues 10 to 14 (LLGLL) abrogated the formation of cytoplasmic filamentous networks. This sequence corresponds to a conserved motif, LLXLL, found at the N terminus of other E4 proteins, and similar results were obtained on deletion of the HPV 16 motif, LLKLL (residues 12 to 16). Our findings indicate that this conserved motif is likely to play a central role in the association between E4 and the cytokeratins. An HPV 1 E4 mutant protein containing a deletion of residues 110 to 115 induced the collapse of the cytokeratin IFs in a manner analogous to the HPV 16 E4 protein. The sequence deleted, DLDDFC, is highly conserved between cutaneous E4 proteins. HPV 1 E4 residues 42 to 80, which are rich in charged amino acids, appeared to be important in the cytoplasmic localization of E4. In addition, we have mapped the N-terminal residues of HPV 1 E4 16-kDa and 10/11-kDa polypeptides expressed by using the baculovirus system and shown that they begin at tyrosine 16 and alanine 59, respectively. Similar-sized E4 proteins are also found in vivo. N-terminal deletion proteins, which closely resemble the 16-kDa and 10/11-kDa species, expressed in keratinocytes were both cytoplasmic and nuclear but did not form cytoplasmic filamentous networks. These findings support the postulate that N-terminal proteolytic processing of the E1^E4 protein may modulate its function in vivo.

Human papillomaviruses (HPVs) are small DNA viruses which infect the epithelium to produce a variety of clinically different proliferative lesions at different anatomical sites (reviewed by Shah and Howley [38]). On the basis of their epithelial tropisms, HPVs can be broadly divided into two groups, those which infect cutaneous epithelia (e.g., HPV 1) and those which infect the mucosa (e.g., HPV 16). Propagation of these viruses is closely linked to epithelial differentiation, with vegetative viral DNA replication and virion assembly being restricted to terminally differentiating keratinocytes (reviewed by Howley [22]). A viral RNA transcript formed by a single splice between the beginning of the E1 open reading frame and the E4 open reading frame, E1^E4, is the major viral transcript found in HPV-induced lesions (6, 7, 28). The E1 exon encodes the extreme N-terminal amino acids (HPV 1, 6, and 11, five amino acids) of the E1^E4 protein. Although the E4 proteins are expressed in large amounts in papillomas (4, 5, 8, 12, 13, 43), the role(s) of the E4 protein in the virus life cycle has remained rather elusive, primarily because of the lack of an in vitro HPV replication system. Mutational analysis of the bovine papillomavirus 1 E4 gene showed that the E4 protein was not required for transformation or establishment of episomal viral DNA (29). In HPV-induced papillomas the E4 protein is localized primarily to cells of the differentiating layers of the epithelium (4, 8, 12, 13, 31). Taken together, these studies suggested that the E4 protein was more likely to have a role in productive infection, possibly by disrupting the normal process of differentiation to favor virion maturation (12) and/or viral DNA replication (4).

More recently, several reports have shown that the E4 proteins are associated with the keratin cytoskeleton of cultured epithelial cells (15, 34, 37). Immunoelectron microscopy studies localized the E4 proteins to tonofilament-like structures in HPV 1-induced warts (37) and keratinized epithelium formed by HPV 16-containing keratinocytes (40). The observation that the HPV 16 E4 protein induced the collapse of the cytokeratin network in cultured cells (15, 34) indicated a possible role in aiding virus egress (15). However, a similar dramatic collapse of the keratin intermediate filaments (IFs) was not observed in HPV 1 E4-expressing cells (34, 37) or in HPV 16-containing stratified cervical keratinocytes (40).

Much of our knowledge of the E4 proteins has come from studies of the HPV 1 E4 protein (4, 12, 16, 20). In HPV 1-induced warts multiple E4 protein doublets of 10/11-kDa, 16/17-kDa, 21/23-kDa, and 32/34-kDa have been identified (4, 12). The 17-kDa species encodes the E1^E4 polypeptide,

* Corresponding author. Phone: 21-414-4471. Fax: 21-414-4486.

N-terminal processing of which is thought to give rise to the smaller proteins (16, 32), although the exact position of cleavage is unknown. The larger species represent E4 multimers (4, 12, 16). The different HPV 1 E4 species have a distinct pattern of expression in warts (4, 16). The 17-kDa protein is the first E4 species to be synthesized in cells of the parabasal layer, and its expression is coincident with the onset of vegetative viral DNA replication (4). The other species appear progressively and concomitantly with epithelial differentiation, the 10/11-kDa species being most abundant in the most superficial wart layers. The differential expression and post-translational modification of the E4 polypeptides may reflect functional modulation of the E4 protein (4, 16, 20).

Comparison of the E4 proteins showed only limited sequence homology, even between types which share similar epithelial specificity (13). Aside from a recent report which demonstrated distinct cellular localizations in cultured cells of HPV 1 E4 proteins differing in their N-terminal sequences (36), the relationship between the structure and function of the E4 proteins is largely unknown. In a previous study, we expressed the HPV 1 and 16 E1^E4 proteins, which have limited sequence homology and different epithelial specificities, in simian virus 40 (SV40)-transformed epithelial cells (34). In the cell line, SVJD, both E4 proteins were localized to extensive and organized cytoplasmic filamentous networks coincident with the cytokeratin filaments. The HPV 16 E4-cytokeratin networks, but not the HPV 1 E4-cytokeratin filaments, progressively collapsed to form a fibrous clump. To identify E4 sequences important in the interaction of the E4 protein with cultured keratinocytes and, specifically, with the cytokeratin networks, we have constructed mutations in the HPV 1 and 16 E1^E4 genes by using site-directed mutagenesis and expressed the mutant proteins by using the SV40-epithelial expression system. The effects of the mutations on localization of the mutant proteins were assayed by indirect-immunofluorescence microscopy.

MATERIALS AND METHODS

Cell lines. Cos-1 is an SV40-transformed cell line derived from CV-1, a simian kidney epithelial cell line (19), and SVJD, an SV40-transformed human adult skin keratinocyte cell line (35a). Both cell lines constitutively express SV40 large T antigen. Cell cultures were maintained in Dulbecco's modified Eagle's medium supplemented with 5% fetal calf serum.

Construction of HPV 1 and 16 E4 mutations. Full details of the construction of the HPV 1 and 16 E1^E4 cDNAs are given by Roberts et al. (34). Deletions and single and multiple substitutions were generated by oligonucleotide-directed in vitro mutagenesis by the method developed by Taylor et al. (42) and supplied in kit form by Amersham International. To prepare single-stranded template DNA we cloned the HPV 1 and 16 E1^E4 cDNAs into a unique *Bgl*II restriction site engineered into the phagemid vector, PTZ18 (Bio-Rad). The HPV 1 E4 deletion mutant, Δ2–60, was derived from the E1^E4 cDNA by the PCR technique. The forward PCR primer (5′-TCTAGCCACCATGGGCCTTACGGACG-3′) fused an initiation codon (underlined) in the correct context (25) to the E4 sequence beginning at codon 61. The PCR fragment and mutated cDNAs were sequenced in their entirety and subcloned into the SV40 early replacement vector, pAP16 (27), downstream of the SV40 early promoter as described previously (34, 35).

Preparation of SV40 recombinants. High-titer stocks of SV40 recombinants were prepared essentially as described previously (34). Briefly, the recombinant viral vectors were digested with *Eco*RI to remove the bacterial plasmid vector, recircularized with DNA ligase, and electroporated into the SV40 permissive cell line, Cos-1. Cells were harvested at maximum cytopathic effect and frozen and thawed three times, and the virus supernatant was stored at −70°C.

Immunofluorescence microscopy. Infection of SVJD cells with recombinant virus and processing for immunofluorescence microscopy were carried out as described previously (34). Cells were grown on glass slides and infected in situ with recombinant virus. At 72 h postinfection the cells were fixed in freshly prepared 4% paraformaldehyde for 8 min and then subjected to permeabilization in cold (−20°C) acetone for 10 min.

HPV 1 wild-type and mutant E4 proteins were detected by using the 4.37 mouse monoclonal antibody (MAb) (16) or an anti-E4 rat polyclonal serum, p1p7 (12). HPV 16 E4 proteins were detected with a mouse MAb, TVG402 (kindly provided by John Doorbar, University of Cambridge, Cambridge, United Kingdom [14]).

The anti-keratin antibodies used in this study were a rabbit antiserum to K5 and MAbs to K18 (LE65 [26] and CY-90 [Sigma Chemicals]). The rabbit K5 antiserum and LE65 were a kind gift of Birgitte Lane, University of Dundee, Dundee, United Kingdom. For LE65 activity, cells expressing mutant HPV 16 E4 proteins were processed in acetone (−20°C) for 5 min.

Specific antibody interactions were visualized with fluorescein- or rhodamine-conjugated antibodies purchased from Sigma Chemicals and Southern Biotechnology Associates. All reagents were used as specified by the manufacturers.

Protein sequencing. Aliquots (~10 to 30 μg) of HPV 1 E4 protein, purified from insect cells infected with a recombinant baculovirus as described by Roberts et al. (35), were fractionated on a sodium dodecyl sulfate-polyacrylamide gel and electroblotted onto Problott immobilization membrane (Applied Biosystems) in 10 mM CAPS [3-(cyclohexylamino)-1-propane sulfonic acid] (pH 11.0)–10% methanol at 200 mA for 12 to 16 h. The blot was rinsed in distilled water, and proteins were stained with 0.2% Ponceau S in 1% acetic acid. Following brief destaining in distilled water, the individual polypeptide bands were excised and sequenced on an Applied Biosystems sequencer model 437A.

Computer analysis. Alignment of the HPV E4 amino acid sequences was performed by using the sequence analysis software package, GCG version 7.3 (Genetics Computer Group Inc., University of Wisconsin). The E4 sequences were obtained from GenBank (release 81.0), EMBL (release 37.0), and NBRF-PIR (release 39.0) databases.

The HPV 1 E4 sequence was compared with the NBRF-PIR protein database (release 38.0) by using a rigorous local alignment algorithm (2). Amino acid substitutions were scored by using the Dayhoff 250 PAM mutation data matrix (10). The length-dependent gap penalty was set to 8.0. The significance of the highest-scoring alignments was evaluated by using a new statistic that considers the score and length of the alignment (2a) and was also checked using a conventional Needleman-Wunsch alignment (30). The standard deviation score of each alignment was calculated by comparing the score for native sequences with the distribution of scores for 100 shuffled sequence pairs of the same length and composition.

RESULTS

Construction of mutant HPV 1 and 16 E4 proteins prior to expression in epithelial cells. Our mutational analysis of the E4 proteins used the HPV 1 E1^E4 protein as the primary

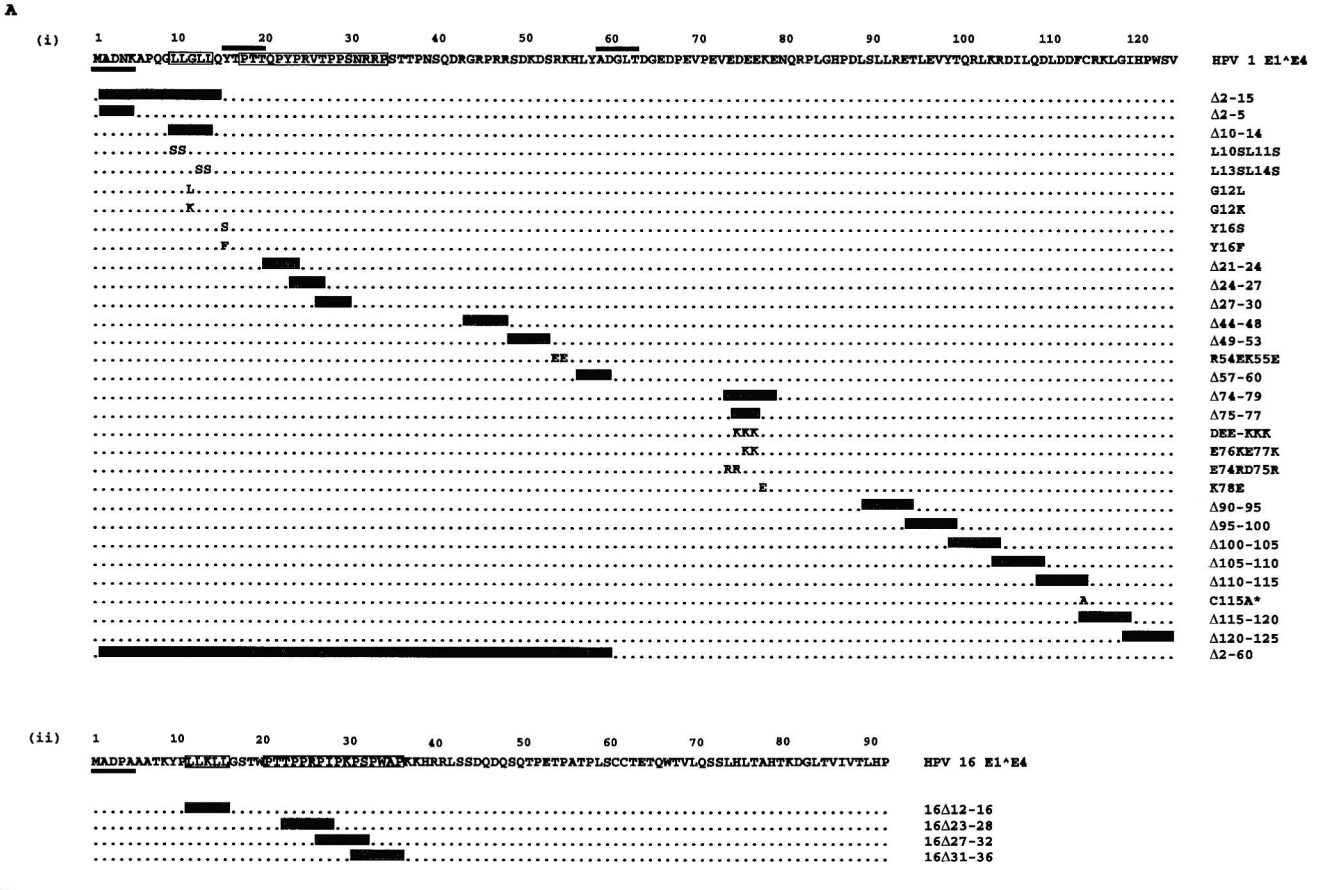


FIG. 1. (A) Structure of the HPV 1 (i) and 16 (ii) E1^E4 mutant proteins. The amino acid sequence of the E1^E4 proteins is shown. The N-terminal amino acids derived from the E1 exon are underlined. The HPV 1 and 16 E4 leucine clusters and proline-rich regions are boxed. The N-terminal amino acid sequences obtained from the baculovirus-expressed 16- and 10/11-kDa HPV 1 E4 polypeptides are indicated by solid lines above the sequence. The positions of deletions are indicated by solid bars. Each mutant is named in the right-hand column. C115A*, construction of this mutant protein has been described elsewhere (35). The single-letter amino acid code is used. (B) Sequence alignment of the C-terminal sequences of cutaneous HPV E4 proteins. Conserved residues are shaded. Note the high incidence of conserved hydrophobic amino acids, ***H*** (L, I, V, and F). The single-letter amino acid code is used.

template and, generally, targeted regions containing sequences conserved between E4 proteins. The leucine-rich sequence, LLGLL (residues 10 to 14), near the N terminus of the HPV 1 E1^E4 protein and the relatively proline-rich sequences found C-terminal to the leucine motif both represent conserved characteristics of the E4 proteins, including the HPV 16 E1^E4 protein (11, 16) (these regions are boxed in Fig. 1A). Therefore, these sequences in both the HPV 1 and 16 E1^E4 proteins were mutated. The HPV 1 E4 sequence between residues 42 and 80 is dominated by charged amino acids (~56%) and can be further divided into a basic (approximately residues 42 to 56) and an acidic (approximately residues 64 to 80) domain. Charged domains are present in other E4 proteins (13), the most unusual being the HPV 4 E4, which encodes a string of 10 glutamic acid residues. To determine the importance of these residues in the formation of HPV 1 E4-cytokeratin networks, we constructed a small panel of deletions and substitutions across this region. Sequence alignment of the E4 proteins present in the database showed significant conservation of the C-terminal regions of cutaneous HPV E4 proteins, a notable feature of which is the preservation of hydrophobic residues (Fig. 1B). The importance of this region in the HPV 1 E4 protein was examined by constructing a series of small overlapping deletions from residue 90 to the C terminus.

The structures of the mutant HPV 1 and 16 E1^E4 proteins are illustrated in Fig. 1A. Substitutions are named by using the single-letter symbol for the altered amino acid followed by the position and the amino acid introduced at that position. Deletions are named by a Δ symbol followed by the number of the first and last amino acids deleted. HPV 16 mutations are distinguished by "16" preceding the mutant name.

In addition to the mutational analysis of the E4 proteins outlined above, we wanted to study the expression, in keratinocytes, of the smaller HPV 1 E4 polypeptides found in vivo

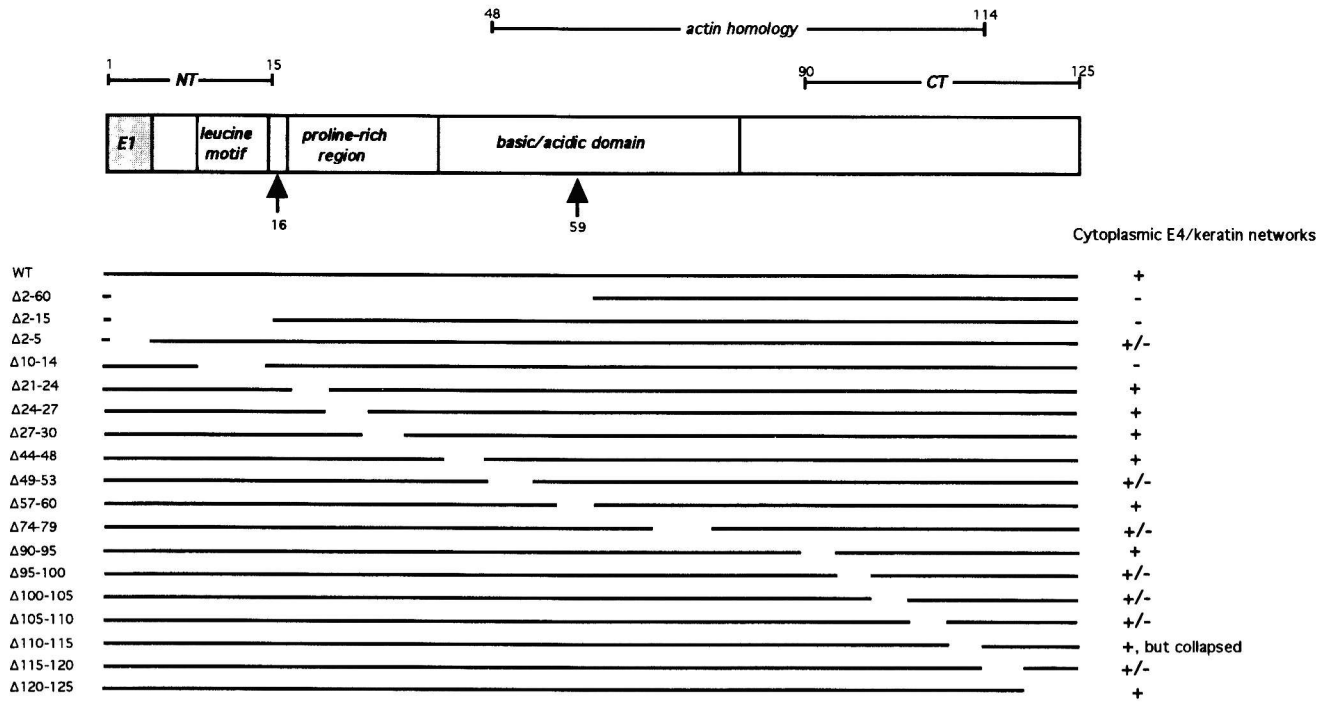


FIG. 2. Summary of HPV 1 E1^E4 deletion analysis. Schematic diagram of HPV 1 E1^E4 showing the E1 exon (shaded) and the position of the leucine motif, the proline-rich region, and the domain rich in charged (basic and acidic) amino acids. The extreme N terminus (NT) encodes the E1 exon and leucine motif which are conserved between E4 proteins (including HPV 16). The N termini of the smaller E4 polypeptides expressed in baculovirus-infected cells (35) have been mapped to tyrosine 16 (16 kDa) and alanine 59 (10/11 kDa) (their positions are indicated by arrows). The C-terminal region (CT) represents a conserved domain between cutaneous E4 proteins, and the E4 sequence from 48 to 114 has significant homology with actin. The structures of the HPV 1 E1^E4 deletion mutants are depicted below. The ability of these mutant proteins to form cytoplasmic E4-keratin networks when expressed in SVJD keratinocytes is indicated in the right-hand column. Symbols: +, formation of networks; −, no formation of networks; +/−, deletions which impair the ability of the mutant proteins to form networks.

(16, 11, and 10 kDa [4, 12]). It is generally assumed that they are derived from the 17-kDa E1^E4 protein by N-terminal proteolytic cleavage (16, 32), but this has not yet been confirmed, because their N termini are blocked and thus refractory to sequencing. In previous attempts to determine the primary structure of wart E4 proteins, an N-terminal sequence, commencing at glycine 61, was obtained from a 10/11-kDa protein preparation (35a). However, this species represented only a small fraction of the protein sample and was most probably a degradation product. On the basis of the reactivity of mapped anti-HPV 1 E4 MAbs (16), it was predicted that the E4 sequences of the smaller species must begin at approximately residue 16 (16 kDa) and between residues 58 and 70 (10/11 kDa) of the E1^E4 sequence. Therefore, it was judged that glycine 61 may be close to the position of the 10/11-kDa N termini. To determine whether a protein consisting of just the C-terminal half of the HPV 1 E4 protein could be expressed in epithelial cells, we constructed a deletion mutant which encoded residues 61 through to the C terminus [Δ2–60, Fig. 1A (i)]. More recently, we have observed that multiple HPV 1 E4 species with molecular masses similar to those detected in warts are expressed in insect cells infected with a recombinant E4-baculovirus (see Fig. 1 in reference 35). The purified baculovirus-produced E4 proteins were electroblotted onto an immobilization membrane, and the individual species were excised and sequenced (see Materials and Methods). No sequence was obtained from the 17-kDa E1^E4 species, indicating that the amino terminus is blocked. The sequences obtained from the 16-kDa (Tyr-Thr-Pro-Thr-Thr) and the 10/11-kDa (Ala-Asp-Gly-Leu-Thr) polypeptides correspond to residues 16 to 20 and 59 to 63, respectively, in the E1^E4 sequence [Fig. 1A (i)]. The N-terminal positions of tyrosine 16 (16 kDa) and alanine 59 (10/11 kDa) are in good agreement with the predicted N termini of the wart species. Therefore, in addition to the Δ2–60 mutant, we constructed a second deletion mutant which lacks residues 2 to 15 [Δ2–15; Fig. 1A (i)].

SV40-mutant E4 recombinant viruses were prepared as described in Materials and Methods and used to infect the SV40-transformed human adult skin keratinocyte cell line SVJD. Immunofluorescence microscopy with antibodies to the E4 and cytokeratin proteins monitored the expression of the mutant proteins and their effects on the keratin networks. More than one infection was performed for each mutant to ensure phenotypic consistency. The ability of HPV 1 E1^E4 deletion mutant proteins to form cytoplasmic filamentous networks and coalign with the cytokeratin filaments is summarized in Fig. 2. A more detailed description of the effects of all E4 mutations on cellular localization of the mutant E4 proteins is given in Table 1.

HPV 1 E4 N-terminal deletion mutants do not form cytoplasmic networks in epithelial cells. The Δ2–60 protein was not recognized by the 4.37 MAb (the 4.37 epitope has been mapped to approximately residues 55 to 63 [16]) and therefore was detected by using a rat anti-E4 polyclonal serum, p1p7. The mutants Δ2–15 and Δ2–60 produced similar but not identical immunofluorescence staining. Like the wild-type (*wt*) protein, both mutant proteins were detected in the nucleus and cytoplasm, but neither formed cytoplasmic filamentous networks (Fig. 3; compare Δ2–15 and Δ2–60 with WT). In the majority of positive cells, nuclear staining of both mutant

TABLE 1. Summary of the cellular distribution of mutant HPV E4 proteins expressed in SVJD cells[a]

Mutant E4 protein	Immunofluorescent staining patterns of mutant E4 proteins
HPV1	
wt[b] Y16F, Y16S, Δ21–24, Δ24–27, Δ27–30, Δ44–48, E74RD75R, K78E, C115A	Cytoplasmic and nuclear. Cytoplasmic filamentous networks, multiple E4-stained inclusions with bright annular fluorescence (*wt*[c]). Coincident cytokeratin staining.
R54EK55E, Δ57–60, Δ90–95, Δ120–125	Similar to the *wt* phenotype, except network formation was often associated with strong cytoplasmic diffuse staining.
Δ2–5	Exclusively cytoplasmic in ~70% of cells. Diffuse cytoplasmic staining, also localized to faint filamentous networks and inclusions (*wt*), coincident with cytokeratins. In ~20% of cells Δ2–5 localized exclusively to cytoplasmic inclusions (*wt*) and spots.
Δ2–15, Δ10–14,[d] L10SL11S,[d] L13SL14S,[d] Δ2–60	Cytoplasmic and nuclear. Cytoplasmic staining predominantly diffuse or patchy. Nuclear staining often bright. No coincident cytokeratin staining.
G12L	Predominantly cytoplasmic. Cytoplasmic staining localized to bright "spots" or large granules. Granules are clusters of smaller inclusions (*wt*). No coincident cytokeratin staining. No nuclear staining in majority of cells with cytoplasmic granules.
G12K	Cytoplasmic and nuclear. Cytoplasmic filamentous networks, coincident with cytokeratins, often associated with diffuse cytoplasmic and bright nuclear staining. Inclusions (*wt*) rare.
Δ49–53, Δ74–79, DEE-KKK	Predominantly nuclear, inclusions (some *wt*). Cytoplasmic staining localized to faint filamentous networks (Δ49–53, 20–30% of cells; Δ74–79, 70–90% of cells; DEE-KKK, 1–5% of cells) and inclusions, coincident with cytokeratins. Aberrant DEE-KKK/cytokeratin networks also detected in ~10% of cells. Δ74–79 also localized to multiple cytoplasmic spots which decorated the Δ74–79 filaments. Intranuclear localization of Δ49–53 inclusions confirmed by confocal microscopy.
Δ75–77, E76KE77K	Cytoplasmic and nuclear. Cytoplasmic filamentous networks and inclusions (*wt*) coincident with the cytokeratins. Filaments formed by the mutant protein were sometimes decorated with spots of staining.
Δ95–100, Δ100–105, Δ105–110, Δ115–120	Cytoplasmic and nuclear. Predominantly a punctate or diffuse (sometimes filamentous) cytoplasmic staining. In ~5–15% of cells the mutant proteins were almost exclusively localized to distinct cytoplasmic networks, coincident with cytokeratins.
Δ110–115	Cytoplasmic and nuclear. Diffuse or distinct cytoplasmic filamentous networks coincident with cytokeratins. Δ110–115/cytokeratin networks progressively collapsed to form fibrous perinuclear coils or clumps. Multiple cytoplasmic spots and inclusions (*wt*) detected in some cells. Exclusively cytoplasmic in some cells with distinct networks.
HPV16	
wt[b] 16Δ23–28, 16Δ27–32, 16Δ31–36	Extensive cytoplasmic filamentous networks coincident with the cytokeratins. E4/cytokeratin networks progressively collapsed to form fibrous inclusion bodies. Nuclear staining in cells with collapsed structures.
16Δ12–16	Exclusively intranuclear in >80% of cells. Cytoplasmic staining localized to short filaments or aggregates. Cytokeratin networks (K18 and K5) intact, some colocalization with cytoplasmic 16Δ12–16 structures.

[a] The mutant E4 proteins were broadly divided into groups based on similar immunofluorescent staining.
[b] The immunofluorescent staining of the *wt* HPV1 and 16 E4 proteins has been described previously (34).
[c] Abbreviation for the E4-stained inclusions detected in *wt*-expressing cells.
[d] Faint cytoplasmic aggregates were observed in some cells, but their detection varied between infection experiments.

proteins was bright, although cells (<1%) with only a punctate cytoplasmic Δ2–60 staining were detected (data not shown). Cytoplasmic staining appeared diffuse or patchy and, in some cells, faint. Occasionally (<5% of positive cells), the Δ2–15 protein was also localized to small bright cytoplasmic "spots" (see the lower cell shown in Fig. 3, Δ2–15). The lack of Δ2–15 filamentous networks was not due to masking of the 4.37 epitope, because 4.37 and p1p7 produced identical immunofluorescence staining (data not shown). The cytokeratin networks in these cells were normal (data not shown).

The leucine-rich motif conserved at the N terminus of E4 proteins was essential for the formation of cytoplasmic filamentous networks. Deletion of HPV 1 E4 residues 2 to 15 (Δ2–15), which removed the E1 sequences and the leucine-rich motif (LLGLL), abrogated the formation of cytoplasmic filamentous E4 structures and indicated that the extreme N-terminal sequences are important in HPV 1 E4 network formation. Further mutational analysis of this region showed that the leucine motif plays a crucial role in the formation of these structures. Deletion of the motif (Δ10–14) or substitution of the leucine amino acids (L10SL11S and L13SL14S) abolished the formation of cytoplasmic E4 networks in keratinocytes. The immunofluorescence staining patterns produced by these mutant proteins were identical (Table 1) and were similar to Δ2–15 staining (shown in Fig. 3), although in some cells faint E4-stained aggregates were detected (data not shown). Also, a glycine 12-to-leucine substitution prevented the mutant protein (G12L) from forming filamentous networks but produced strikingly different phenotypes (Fig. 4A and B). Cytoplasmic G12L staining was localized predominantly to multiple spots (Fig. 4A) or large brightly stained granules (Fig. 4B). At a higher magnification the granules appeared as clusters of small inclusions which show a bright annular fluorescence (data not shown). The staining of these small inclusions is similar to the E4 staining of phase-dense inclusions (arrowed in Fig. 3 WT) observed in some cells expressing the *wt* HPV 1 E4 protein and which also can be coincident with cytokeratin staining (34, 35a). In this study, inclusions with this characteristic staining pattern will be referred to as *wt* inclusions. Cytokeratin staining was not coincident with G12L staining (data not shown). Not all cells, and in particular not those expressing the large aggregates, expressed nuclear G12L.

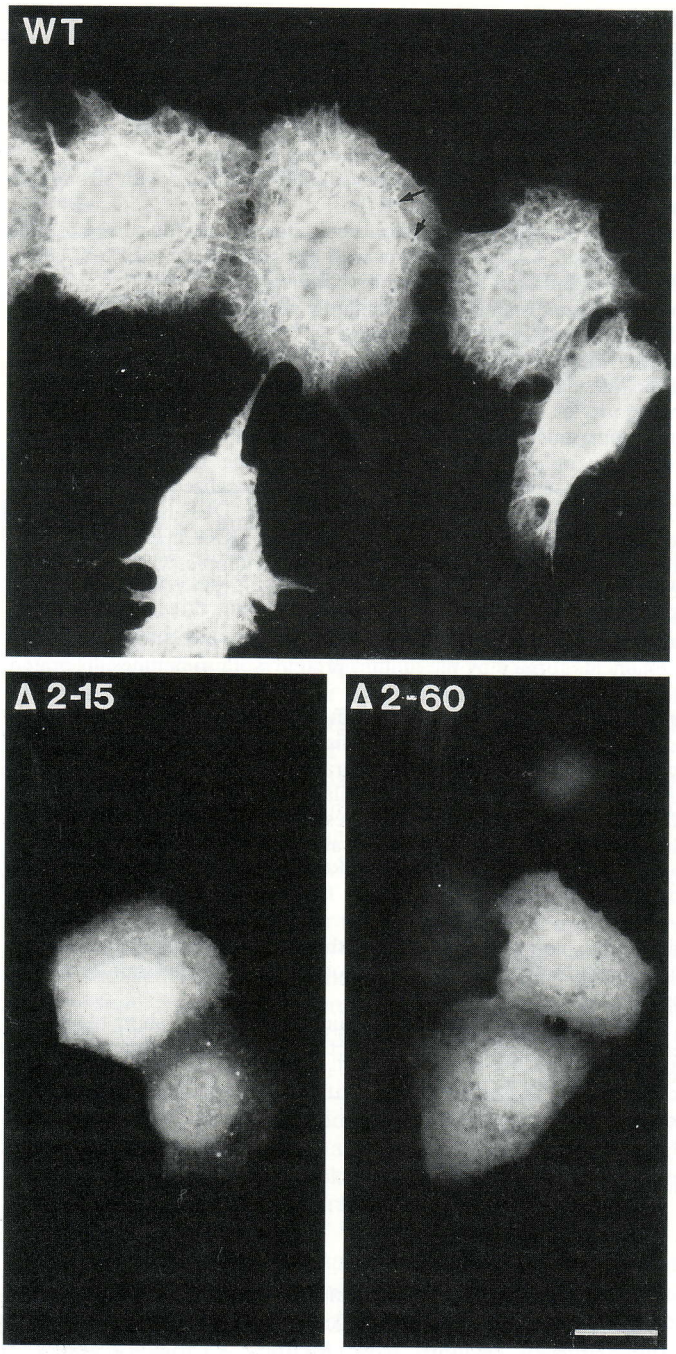


FIG. 3. Immunofluorescence staining of HPV 1 E4 N-terminal deletion mutants Δ2–15 and Δ2–60 expressed in SVJD cells. The phenotype of the *wt* E4 protein is also shown (WT). Both WT and Δ2–15 proteins were detected with an anti-E4 MAb (4.37), and Δ2–60 was detected with a rat polyclonal serum (p1p7) (see Materials and Methods for details). The cytoplasmic filamentous networks and small inclusions (indicated by arrows in WT) formed by the *wt* protein are not formed by Δ2–15 or Δ2–60 mutant proteins. Bar, 18 μm.

A second substitution at this position (glycine → lysine) converted the HPV 1 leucine motif (LLGLL) to the HPV 16 sequence (LLKLL). In contrast to G12L, this mutant protein (G12K) was localized to cytoplasmic filamentous networks which were associated with bright nuclear staining and often a diffuse cytoplasmic staining (Fig. 4C).

Mutation of tyrosine 16 (Y16S and Y16F) [Fig. 1A (i)] did not affect the formation of filamentous E4 networks (Table 1). Both proteins produced staining patterns indistinguishable from that of the *wt* protein (data not shown).

To determine whether the HPV 1 E1^E4 N-terminal amino acids encoded by the E1 exon played a role in the formation of E4-cytokeratin structures in SVJD cells, we constructed the deletion mutant Δ2–5 [Fig. 1A (i)]. Unlike the leucine motif mutations, deletion of the E1 residues did not abrogate the formation of filamentous structures. However, Δ2–5 networks were not distinct and were associated with a strong cytoplasmic diffuse staining (Fig. 4E and F), although filaments were clearly visible running across the top of the nucleus of cells which did not express nuclear Δ2–5 protein (Fig. 4G). In fact, the majority of positive cells (~70%) did not express nuclear Δ2–5 protein. The Δ2–5 protein was also localized to cytoplasmic *wt* inclusions (indicated by an arrow in Fig. 4E), and, in some cells (~20%), the mutant protein was exclusively localized to these inclusions or bright cytoplasmic "spots" (Fig. 4D).

Deletion of the leucine motif from the HPV 16 E1^E4 protein (LLKLL, 16Δ12–16) also had a deleterious effect on E4 network formation. In the majority of positive cells (>80%) the mutant protein was exclusively intranuclear (Fig. 5B). Intranuclear localization was confirmed by confocal microscopy (data not shown). Cytoplasmic 16Δ12–16 staining was less common and appeared as short filaments or bright aggregates (Fig. 5C and D, respectively). In SVJD cells expressing *wt* HPV 16 E4 protein, the E4-cytokeratin networks progressively collapsed to form a fibrous cytoplasmic inclusion body (34) (Fig. 5A). However, the morphology of the cytoplasmic 16Δ12–16 aggregates was unlike the fibrous clumps of collapsed *wt* E4-cytokeratin proteins. Moreover, dual staining of these cells with E4 and cytokeratin antibodies showed that the cytokeratin networks had not collapsed, although there was some coincident staining between the two proteins (Fig. 5D and E).

In summary, these studies have shown that the clusters of leucine residues found near the N terminus of the HPV 1 and 16 E4 proteins, which also represent a conserved motif of the E4 proteins, played a crucial role in the formation of cytoplasmic filamentous E4 networks in SVJD cells.

Proline-rich sequences are not important for the formation of HPV 1 and 16 E4-cytokeratin structures. A series of small overlapping deletions was constructed, covering the HPV 1 and 16 proline-rich sequences which lie C-terminal to the leucine clusters (HPV 1 residues 21 to 30; HPV 16 residues 23 to 36) (see Fig. 1A for details of the mutations). Following expression in SVJD cells, all deletions caused immunofluorescence staining indistinguishable from that of their respective *wt* proteins (Fig. 2; Table 1; data not shown), indicating that these proline-rich regions, a relatively conserved feature of E4 proteins, do not have an important role in the formation of HPV 1 and 16 E4-cytokeratin structures.

Mutation of HPV 1 E4 sequences rich in charged amino acids alters the cellular distribution of the mutant E4 proteins. Several of the mutations constructed in the charged region (residues 42 to 80) of the HPV 1 E4 protein affected the normal cellular distribution of the mutant protein (Table 1; Fig. 6). A notable feature of these mutations was a predominantly nuclear localization of the mutant proteins. Also, nuclear and/or cytoplasmic immunofluorescence staining was characterized by multiple bright "spots" and/or inclusions (of which some were *wt*). However, faint cytoplasmic filamentous staining of the mutants was detected. These phenotypes were particularly associated with the deletion mutations Δ49–53 (SDKDS; Fig. 6A) and Δ74–79 (EDEEKE; Fig. 6B) and the triple substitution of residues 75 to 77 (DEE-KKK [Fig. 6C and D, inset]). The cytoplasmic spots were a dominant feature of Δ74–79 (Fig. 6B) and were also formed by the mutant proteins,

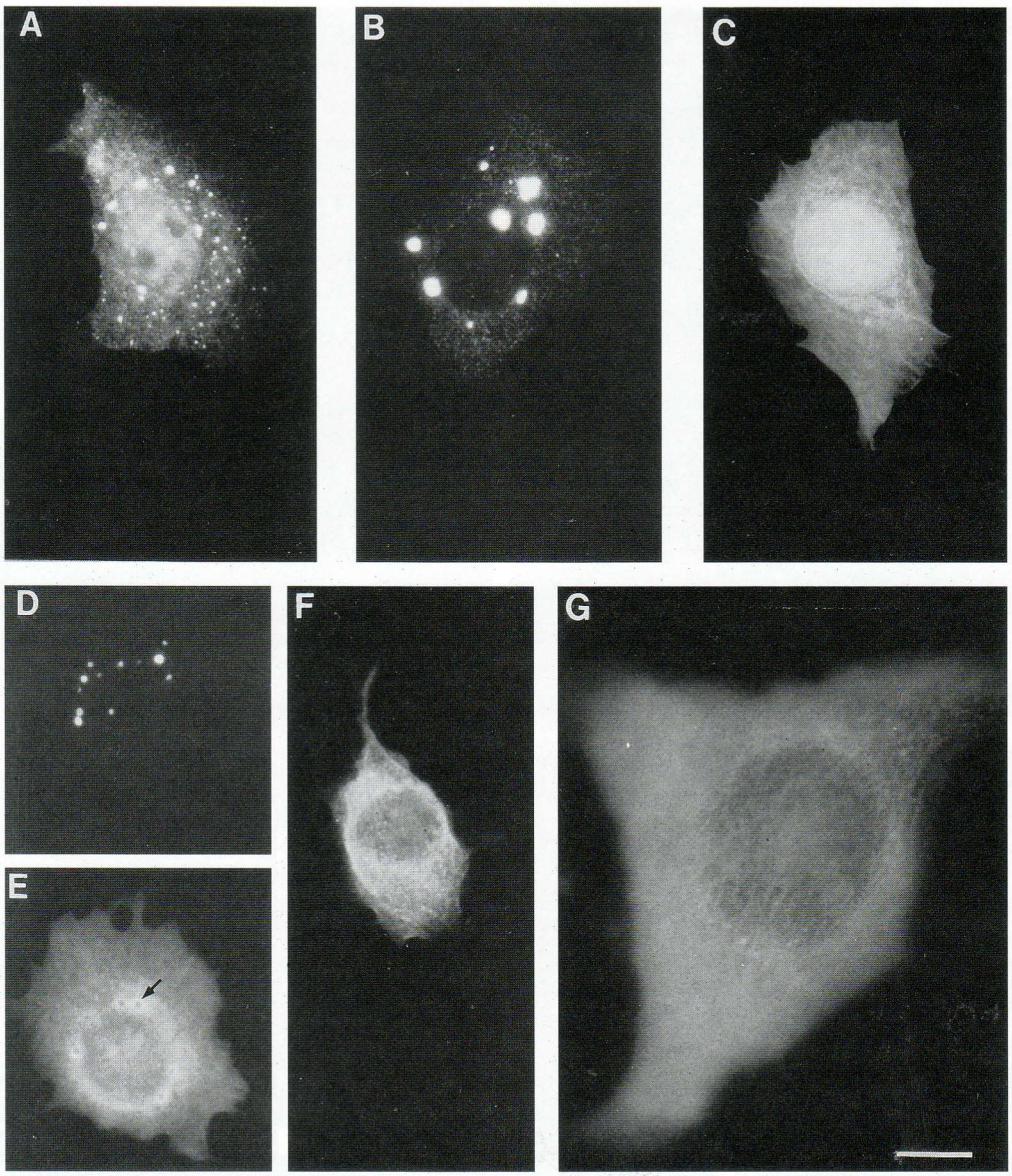


FIG. 4. Immunofluorescence staining patterns formed by HPV 1 E1^E4 N-terminal mutant proteins G12L (A and B), G12K (C), and Δ2–5 (D to G). Mutant proteins were detected with 4.37 (A and C to G) or p1p7 (B; similar staining is formed when using 4.37 [data not shown]). A G12L (A and B) but not a G12K (C) substitution in the HPV 1 E1^E4 LLGLL motif abrogates filament formation (A and B). A mutant protein which lacks the E1 amino acids (Δ2–5) formed inclusions (D and E, indicated by an arrow) and a diffuse cytoplasmic filamentous network (F and G). The cell in panel G has been focused over the top of the nucleus. Bar, 18 μm for panels A to F and 9 μm for panel G.

Δ75–77 and E76KE77K (Table 1). The spots of mutant protein decorated the filaments and formed a "beads-on-a-string" effect (e.g., E76KE77K [Fig. 6B, inset]). Interestingly, a unique phenotype was observed in ~10% of DEE-KKK-expressing cells (Fig. 6D). The protein had formed aberrant filamentous networks in which the filaments were thicker than those formed by the *wt* protein and the network had lost spatial distribution. Dual-immunofluorescence microscopy with anti-E4 and cytokeratin antibodies revealed coalignment between the two proteins (Fig. 6E and F). Cells expressing these aberrant networks did not always express nuclear mutant protein and often appeared to be in or had just undergone mitotic division.

Deletion of sequences near the HPV 1 E4 C-terminus induced the collapse of E4-cytokeratin networks. The cellular distribution of the HPV 1 E4 mutant proteins missing the six C-terminal amino acids (Δ120–125) or residues 90 to 95 (Δ90–95) was overall similar to that of the *wt* protein (Table 1). The deletion mutants covering residues 95 to 110 (Δ95–100, Δ100–105, and Δ105–110) and 115 to 120 (Δ115–120) formed identical staining patterns (Table 1). Cytoplasmic filamentous staining of these mutant proteins was predominantly diffuse, although distinct cytoplasmic networks were detected in ~5 to 15% of positive cells (e.g., Δ105–110 [Fig. 7, inset]). The deletion mutant Δ110–115 (DLDDFC) formed cytoplasmic filamentous networks coincident with cytokeratin networks, but, in contrast to the *wt* protein, the Δ110–115/cytokeratin networks collapsed to form a perinuclear fibrous coil or clump (a cell containing a collapsed structure is indicated by a large arrowhead in Fig. 7). Both K5 and K18 cytokeratin proteins

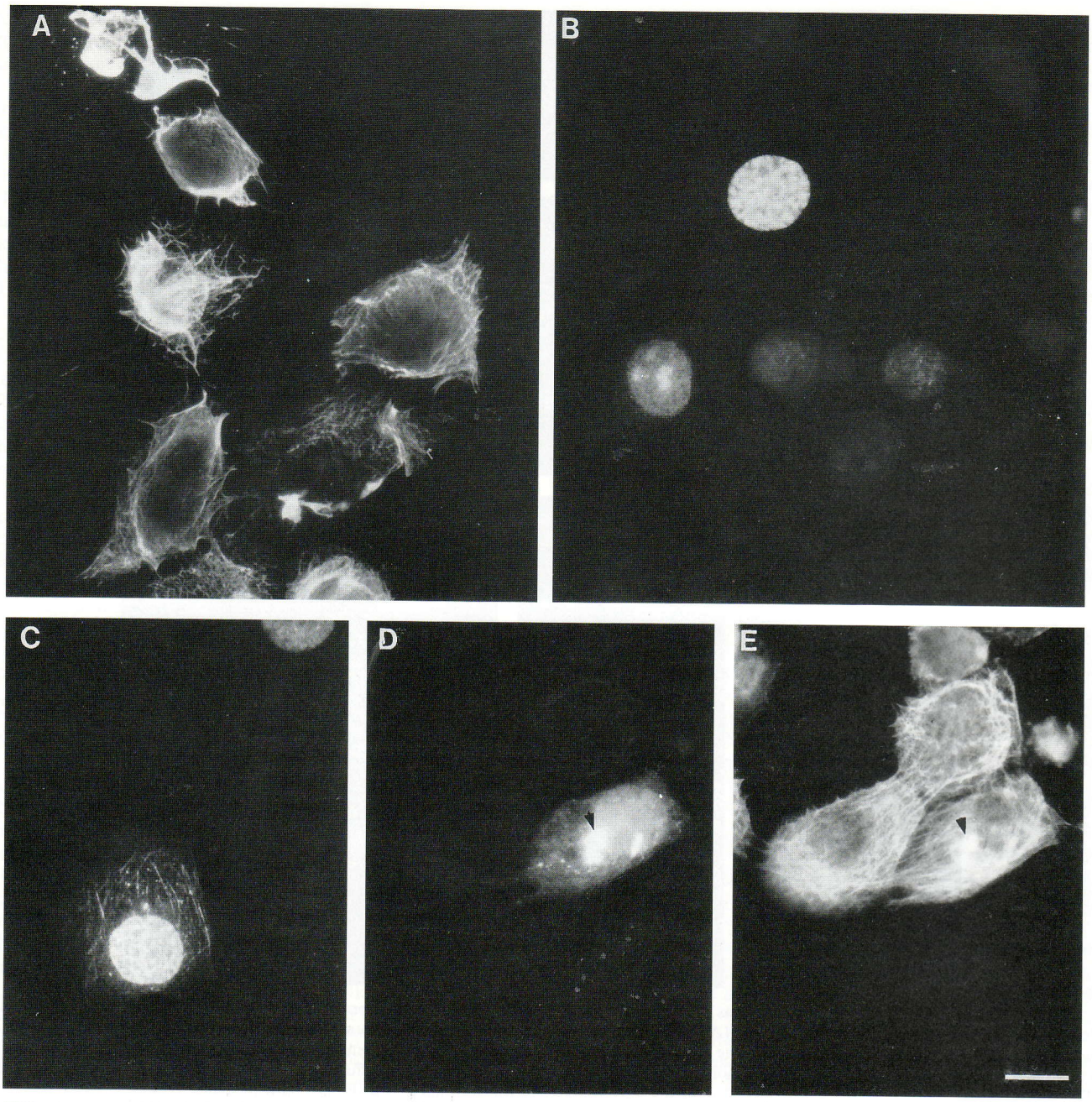
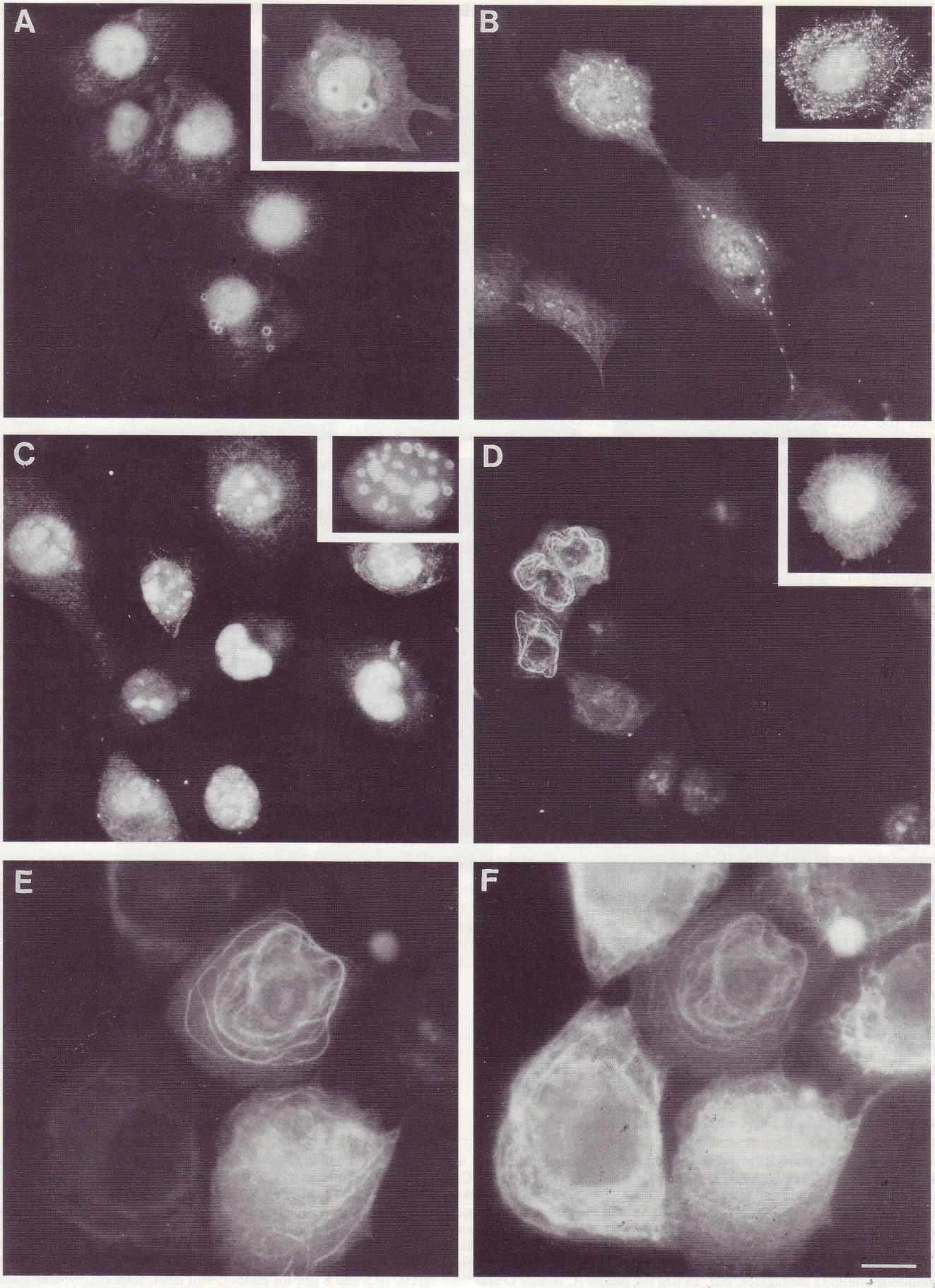


FIG. 5. Cellular distribution of the mutant HPV 16 E4 protein, 16Δ12–16 (B to D). This mutant protein lacks the N-terminal LLKLL sequence. (A) Cellular localization of the *wt* protein, shown for comparison. *wt* and 16Δ12–16 proteins were detected with the TVG402 MAb (see Materials and Methods). The 16Δ12–16 mutant is almost exclusively intranuclear (B). Cytoplasmic 16Δ12–16 staining is shown in panels C and D. Dual-immunofluorescence staining with TVG402 (D) and an anti-K18 MAb, LE65 (E), showed that K18 networks were intact, although there was some colocalization between the two proteins (indicated by arrows in panels D and E). Immune complexes were visualized with anti-mouse immunoglobulin subclass-specific fluorescein (16Δ12–16) and rhodamine (K18) conjugates. Bar, 18 μm.

were associated with the collapsed Δ110–115 structures (Fig. 7 [K5] and data not shown [K18]). The collapse of the Δ110–115/cytokeratin filaments was remarkably similar to the HPV 16 E4-cytokeratin collapse in SVJD cells (34) (Fig. 5A). In some cells the Δ110–115 filamentous staining was very diffuse but coalignment with the cytokeratin filaments was observed (indicated by small arrowheads in Fig. 7). Interestingly, the mutant protein containing the cysteine 115-to-alanine substitution (C115A) formed a staining pattern indistinguishable from that of the *wt* protein (data not shown).

DISCUSSION

The identification of the HPV E4 protein as an IF-associated protein has been an important step in elucidation of its role(s) in the virus life cycle. As part of a study to characterize the mode of interaction between the E4 proteins and the cytokeratin IFs, we have undertaken a comprehensive mutational analysis of the HPV 1 E4 protein covering approximately 67% of the protein and identified specific sequences important in the interaction with the cytokeratin networks in cultured


A
B
C
D
E
F

keratinocytes. Furthermore, we have demonstrated that sequences conserved between E4 proteins play a crucial role in the formation of HPV 1 and 16 E4-cytokeratin structures.

Clusters of leucine residues (LLXLL or corrupted versions) are a conserved feature found near the N terminus of E4 proteins (11). Their importance is demonstrated by the observation that deletion of the HPV 1 (LLGLL) and 16 (LLKLL) motifs abrogated the formation of cytoplasmic E4 networks (Table 1; Fig. 5 [HPV 16]). Moreover, a point substitution in the HPV 1 sequence (glycine → leucine) also abolished network formation (Fig. 4A and B). Mutations elsewhere in the HPV 1 E4 protein did not produce such a deleterious effect (Fig. 2; Table 1). We therefore conclude that the leucine motifs are likely to play a central role in the interaction of the E4 proteins with the cytokeratin IF networks.

The N-terminal 5 amino acids of the HPV 1 E1^E4 protein are derived from the N terminus of the E1 open reading frame (7). Conservation of a spliced E1^E4 structure between other papillomavirus types (1, 6, 9, 17, 28, 39) does, in itself, suggest that the extreme N-terminal amino acids are likely to be important in E4 function. Furthermore, it has been noted that the HPV 1 E1^E4 N-terminal sequence (MADNKA) corresponds to a conserved motif (MADXXA), also found at the N terminus of other E4 proteins (37). The formation of distinct E4-cytokeratin structures in cultured keratinocytes by the HPV 1 E1^E4 protein and an E4 protein lacking the E1 amino acids constituted the first evidence that the N terminus may be important in the interaction of E4 proteins with IFs (36, 37). In this study, an equivalent N-terminal deletion protein (Δ2–5) formed cytoplasmic networks, but their morphology was distinct from that of the *wt* networks (compare Fig. 4F and G with Fig. 3 WT). These results support a role for the E1 residues in the interaction of E4 with the keratin cytoskeleton.

The HPV 1 E1^E4 protein is also associated with inclusion bodies (indicated by arrows in Fig. 3 WT), which can also coalign with cytokeratin staining (34, 35a). In this study the loss in ability of a mutant protein to form filamentous networks did not always correlate with a loss in the formation of inclusions (Table 1). While deletion of the leucine motif or substitution of the leucine pairs abolished the formation of both the networks and inclusions, the point substitution within the motif (G12L) abrogated the formation of the filamentous networks only, and inclusions were still formed. However, the inclusions had aggregated to form large granules (Fig. 4B), and no colocalization with the cytokeratin proteins could be detected. These results suggest that the formation of HPV 1 E4 filamentous networks and inclusions and their colocalization with the cytokeratin proteins involve similar but not identical E4 sequences. E4-stained inclusions, but not filamentous networks, were also detected in a rabbit keratinocyte cell line (VX2R) transiently expressing the HPV 1 E1^E4 protein (36). Ultrastructural and immunoelectron-microscopic studies showed that these inclusions are associated with tonofilament bundles and were similar to the eosinophilic E4 inclusions found in HPV 1-induced warts (37). Failure to detect E4-cytokeratin filamentous networks in VX2R cells may reflect differences in the posttranslational modification processes (34) and/or the level of E4 between the two expression systems.

The HPV 1 E4 C-terminal region (residues 89 to 125) represents a conserved domain in E4 proteins. Unlike the N terminus, however, this conservation is restricted to cutaneous HPV E4 proteins (Fig. 1B). Deletions across part of this region in the HPV 1 E4 protein impaired but did not destroy the ability of the mutant proteins to form filamentous networks (Table 1). However, a surprising result was obtained with a deletion of residues 110 to 115 (Δ110–115). The Δ110–115/ cytokeratin networks collapsed in a manner analogous to the collapse of the cytokeratin networks caused by the HPV 16 E1^E4 protein (Fig. 7). The amino acid sequence deleted (DLDDFC) corresponds to a sequence highly conserved between the cutaneous E4s (Fig. 1B). Interestingly, mutation of the HPV 16 E4 protein has identified C-terminal sequences, conserved among E4 proteins of mucosa-specific HPV types, which have also been shown to be important in E4-cytokeratin collapse (33a). Therefore, the differences in behavior of the HPV 1 and 16 E1^E4 proteins in cultured keratinocytes, i.e., noncollapse versus collapse of the *wt* E4-cytokeratin structures (34), may be related to the C-terminal sequences of the proteins. The tissue-linked diversity of these sequences possibly reflects that their role in the association between E4 and the cytoskeleton may be governed by tissue-specific factors.

The HPV 1 E1^E4 protein expressed in cultured keratinocytes (34, 37) and in vivo (4) is both cytoplasmic and nuclear. The mode of nuclear translocation of the E4 protein is presently unknown, although it has been suggested that the extreme N-terminal amino acids of the E1^E4 protein could be involved (36, 37). Our findings also showed that a mutant HPV 1 E4 protein lacking the E1 amino acids (Δ2–5) is exclusively cytoplasmic in most SVJD cells (Fig. 4D and G). Rogel-Gaillard et al. (37) also noted that the HPV 1 E4 protein encodes a potential nuclear translocation signal (residues 33 to 48). However, deletion of part of this putative signal sequence (Δ44–48) did not alter the cellular distribution in SVJD cells of the mutant protein (Table 1). A notable effect of mutation of the charged domain of the HPV 1 E4 protein (residues 42 to 80) was a predominant nuclear localization of the mutant proteins (Fig. 6). However, the ability of the mutant proteins to form filamentous networks when localized to the cytoplasm was, overall, not impaired. A significant fraction of those mutant proteins unable to form both cytoplasmic networks and inclusions are retained in the cytoplasm (Fig. 3; Table 1), which suggests that the cytoplasmic occurrence of the HPV 1 E4 protein is not dictated simply by the ability to form E4-cytokeratin structures. Taken together, these observations may indicate that sequences in the charged domain have a role in the cytoplasmic retention of the HPV 1 E4 protein, possibly through interaction with cytoplasmic anchoring proteins. In contrast, the cytoplasmic localization of the HPV 16 E1^E4 protein did appear to be dependent on the formation of

FIG. 6. Mutation of HPV 1 E4 sequences rich in charged amino acids produced mutant proteins which have a predominantly nuclear localization in SVJD cells (Δ49–53, Δ74–79, and DEE-KKK [A, B, and C, respectively]). Δ49–53 and DEE-KKK formed inclusions (similar to *wt* inclusions), cytoplasmic and intranuclear (Δ49–53 [A and inset]) or predominantly intranuclear (DEE-KKK [C and inset]). Multiple cytoplasmic spots of Δ74–79- and E76KE77K-staining decorated the cytoplasmic filaments (B and inset, respectively). All formed faint cytoplasmic filamentous network (A and inset, B, and D and inset). Aberrant DEE-KKK networks were observed (D), and dual-immunofluorescence staining of DEE-KKK (E) and cytokeratins (F) showed colocalization between the two networks. The mutant E4 and cytokeratin proteins were detected with 4.37 MAb and a rabbit anti-K5 polyclonal serum. Immune complexes were visualized with fluorescein and rhodamine anti-mouse and anti-rabbit immunoglobulin G conjugates. Bar, 18 μm for panels A, B, and D and their insets, 16.5 and 8 μm for panel C and inset, respectively, and 7 μm for panels E and F.

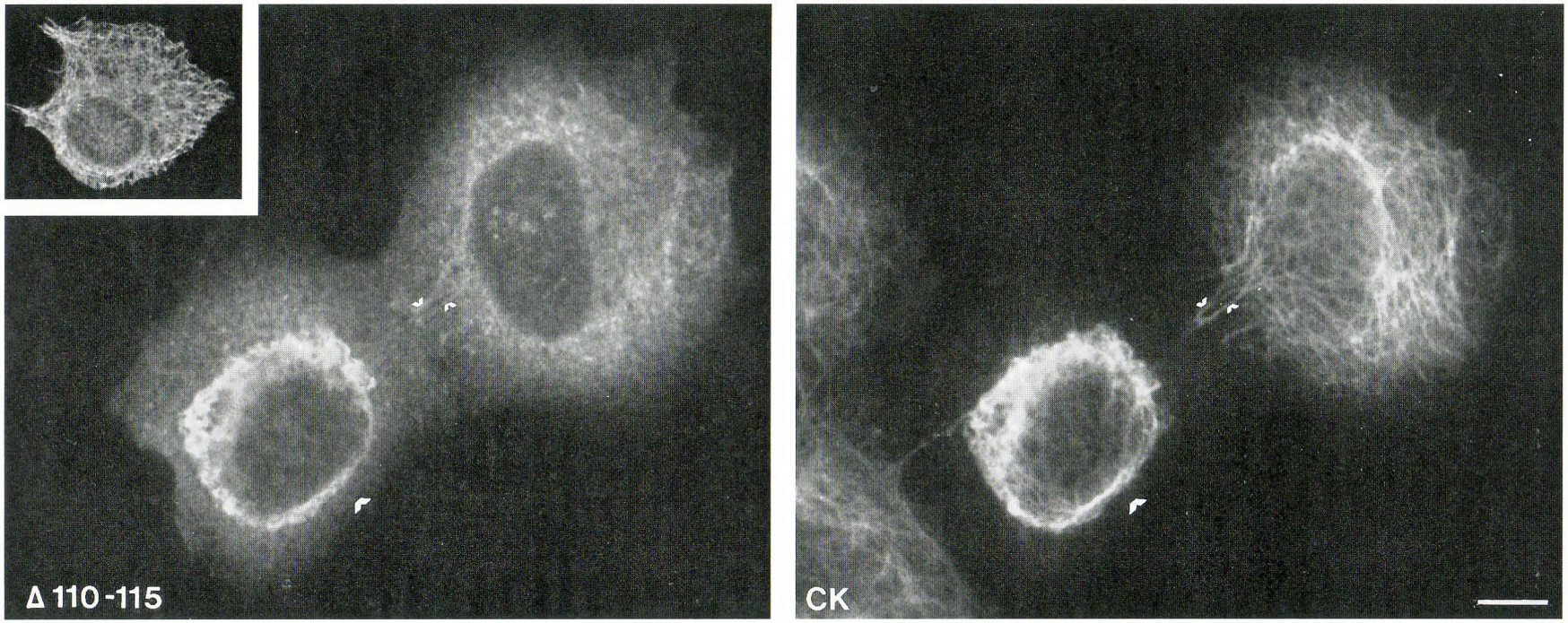


FIG. 7. Immunofluorescence localization of HPV 1 E4 mutant proteins containing deletions of C-terminal amino acids. Dual-immunofluorescence localization of Δ110–115 protein and cytokeratins (CK) is shown in SVJD cells detected with 4.37 MAb and a rabbit anti-K5 polyclonal serum. Antibody staining was followed by rhodamine- and fluorescein-conjugated anti-mouse and anti-rabbit immunoglobulin G second antibodies. In the cell indicated by a large arrowhead, the Δ110–115/CK5 filaments have collapsed to form a perinuclear fibrous clump. In the adjoining cell, CK5 filaments are intact and colocalization with Δ110–115 was observed (small arrowheads in Δ110–115 and CK indicate colocalized staining of filaments). Inset, cytoplasmic filamentous network formed by the Δ115–120 deletion mutant. Bar, 7 μm (18 μm for the inset).

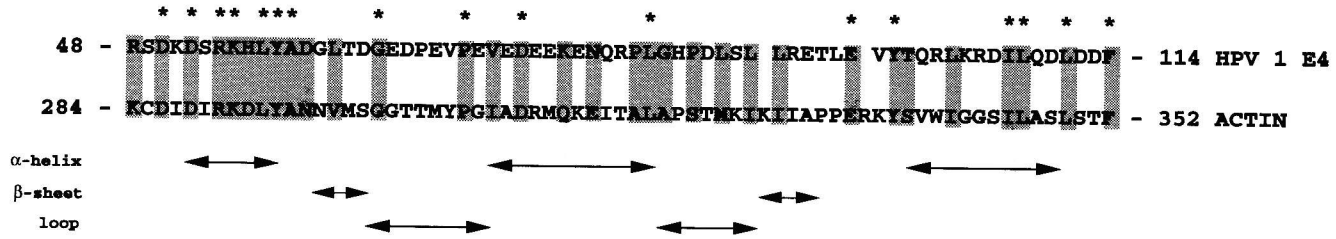


FIG. 8. Alignment of the HPV 1 E4 (residues 48 to 114) and human skeletal muscle actin (residues 284 to 352) protein sequences. The SD score of the alignment was 8.8 (see Materials and Methods for details of the alignment procedure). The average score expected for proteins that do not share similarity is 3.0. Significant alignments to proline-rich precursor proteins were also identified. This similarity is most probably due to the unusual composition of the HPV 1 E4 sequence rather than to any genuine homology. Conserved amino acids are shaded, and invariant residues are indicated by stars. The single-letter amino acid code is used. Arrows below the actin sequence denote the positions of α-helices, β-sheets, and loops in the actin structure according to Kabsch et al. (24).

E4-cytokeratin networks. The leucine motif deletion mutant (16Δ12–16) was almost exclusively intranuclear (Fig. 5B).

Some of the mutant E4 proteins displayed heterogeneous immunofluorescent staining patterns in SVJD cells (Table 1). For example, the mutant protein encoding the triple substitution, DEE-KKK, in the acidic region, was localized almost exclusively intranuclearly in the majority of cells (Fig. 6C) but formed aberrant cytoplasmic E4-cytokeratin networks in cells which appeared to be undergoing mitotic division (Fig. 6D). Therefore, the localization of the E4 proteins may be cell cycle regulated. It may be that changes in the organization of the cytokeratin networks during mitosis (reviewed by Stewart [41]) affect the interaction with the E4 proteins.

The association between E4 and the cytokeratin IF networks may occur through a direct interaction or be mediated by accessory proteins. Whatever the mode of interaction, we have shown that the HPV 1 E4 N-terminal sequences, and specifically the leucine cluster, have an important role in this association. Secondary-structure predictions suggest that the leucine cluster is likely to adopt a β-sheet conformation (16; our unpublished results). This may represent an important structural motif, possibly acting as a binding site or forming part of one for an as yet unidentified binding partner(s). The E1 residues, although important, appear to play a lesser role in the formation of the E4-cytokeratin networks (see above) and therefore may have an indirect role in the interaction. It is interesting that removal of the E1 sequences results in a protein which is less phosphorylated than the full-length E1^ E4 protein (37), and it was suggested that a conformational change may have masked potential phosphorylation sites. Assuming this to be the case, an altered conformation and/or phosphorylation status may result in a reduced affinity of the mutant protein for an interacting species. In contrast to the deleterious effect of mutation of N-terminal E4 sequences, deletion of HPV 1 and 16 E4 sequences which lie immediately C-terminal to the leucine clusters was not detrimental to the interaction with the cytokeratin IFs or their collapse (HPV 16 only). These sequences are relatively rich in proline amino acids and represent a conserved characteristic of the E4 proteins. Prolines are often associated with flexible segments of proteins (3), and it is therefore plausible, although we have no direct evidence, that these regions in the E4 proteins may act as a flexible hinge or linker between the important N-terminal domain and the rest of the E4 protein (Fig. 2). Structural constraints may not be high within a linker region and therefore would be relatively tolerant to mutation.

It is interesting that a search of a protein database by using a rigorous local alignment algorithm (2) identified significant homology (~48% homology and ~25% sequence identity) between a region of the HPV 1 E4 protein (residues 48 to 114) and actin (residues 284 to 352). The alignment and the regions of actin secondary structure, based on the three-dimensional structure (24), are shown in Fig. 8. This region of actin has been implicated in actin self-interactions (residues 286 to 305 [21]) and in the interaction with the myosin head (actin sequences 340 to 349 [33]). Whether this region of the HPV 1 E4 protein does have structural similarity with actin is being investigated by using computer modeling techniques. It should be noted that this alignment with actin is not found with other E4 types.

The N-terminal positions of the 16- and 10/11-kDa HPV 1 E4 polypeptides expressed in baculovirus-infected insect cells (35) were assigned to tyrosine 16 and alanine 59, respectively [Fig. 1A (i)]. These positions agree with the predicted starts of E4 sequences in the wart polypeptides (16) and provide strong evidence that these species are derived form the 17-kDa polypeptide by N-terminal cleavage. Interestingly, the 10- and 11-kDa proteins had the same N-terminal ends, and therefore the difference in molecular mass may be caused by variance in C-terminal sequences (16) and/or posttranslational modification. The processed E4 polypeptides would lack the N-terminal leucine motif of the E1^ E4 protein, which we have shown to be important in the interaction with the cytokeratin networks in cultured keratinocytes (Fig. 2). Indeed, HPV 1 E4 N-terminal deletion mutant proteins which correspond to, or are close to, the 16-kDa (Δ2–15) and 10/11-kDa (Δ2–60) polypeptides did not form cytoplasmic E4-cytokeratin networks in SVJD cells (Fig. 3). Together, these results support the postulate that the function of the E4 protein in vivo may be modulated by N-terminal processing (4, 16, 36). It is interesting that cytoplasmic E4-stained filamentous structures were detected in parabasal cells of an experimental HPV 1-induced wart (34). These may represent structures formed between the 17-kDa E1^ E4 protein, which predominates in these layers (4, 16), and the cytokeratins. Multiple E4 polypeptides have also been identified in lesions induced by HPV types 2, 4, 6, and 11 (5, 13, 43) and bovine papillomavirus type 1 (23), which could indicate that N-terminal processing of E4 proteins is common among other papillomavirus types. We have shown that the N-terminal sequences of the HPV 16 E1^ E4 protein are also important in the association with the cytokeratin IFs, but whether this E4 undergoes similar N-terminal processing is presently unknown. Only a single HPV 16 E4 protein of 10-kDa has been detected in cultured cells (15, 34).

Whether the Δ2–15 and Δ2–60 mutant proteins and the full-length E1^ E4 protein interact with other nuclear or cytoplasmic proteins in SVJD cells is at present unknown. In warts the different HPV 1 E4 polypeptides are expressed in a

progressive manner (4, 16), and the function of the different E4 proteins is likely to be related to their location of expression. It is possible that cellular proteins, important in the function of the processed E4 proteins in vivo, e.g., protein kinases (20) and/or cytokeratins (18), are not expressed in SVJD cells. We have recently identified the HPV 1 E4 protein as a zinc-binding protein (35). Mutation of the zinc-binding site did not abrogate the formation of cytoplasmic E4-cytokeratin networks in SVJD cells. However, zinc may have a role in an, as yet, unassayable E4 activity(ies).

The identification of specific E4 sequences important in the association with the keratin cytoskeleton in cultured keratinocytes is the first step toward the elucidation of the mode(s) of interaction of this diverse group of viral proteins. N-terminal proteolytic cleavage of the E1^E4 protein could be a mechanism by which this association may be modulated in vivo.

ACKNOWLEDGMENTS

We are indebted to Birgitte Lane, University of Dundee, for helpful discussions and for the kind gift of anti-cytokeratin antibodies and to John Doorbar, University of Cambridge, for the gift of the TVG402 antibody. We are also grateful to Julian Parkhill for help with the sequence alignments and to Sue Williams for photography.

This work was supported by the Cancer Research Campaign (CRC). P.H.G. is a CRC Gibb Fellow.

REFERENCES

1. **Baker, C. C., and P. M. Howley.** 1987. Differential promoter utilization by bovine papillomavirus in transformed cells and productively-infected wart tissue. EMBO J. **6:**1027–1037.
2. **Barton, G. J.** 1990. Protein multiple sequence alignment and flexible pattern matching. Methods Enzymol. **183:**403–428.

2a. **Barton, G. J.** Unpublished data.

3. **Bhandari, D. G., B. A. Levine, I. P. Trayer, and M. E. Yeadon.** 1986. ^{1}H-NMR study of mobility and conformational constraints within the proline-rich N-terminal of the LC1 alkali light chain of skeletal myosin. Correlation with similar segments in other protein systems. Eur. J. Biochem. **160:**349–356.
4. **Breitburd, F., O. Croissant, and G. Orth.** 1987. Expression of human papillomavirus type 1 E4 gene products in warts. Cancer Cells (Cold Spring Harbor) **5:**115–122.
5. **Brown, D. R., J. Bryan, M. Rodriguez, R. C. Rose, and D. G. Strike.** 1991. Detection of human papillomavirus types 6 and 11 E4 gene products in condylomata acuminatum. J. Med. Virol. **34:**20–28.
6. **Chow, L. T., M. Nasseri, S. M. Wolinsky, and T. R. Broker.** 1987. Human papillomavirus types 6 and 11 mRNAs from genital condylomata acuminata. J. Virol. **61:**2581–2588.
7. **Chow, L. T., S. S. Reilly, T. R. Broker, and L. B. Taichman.** 1987. Identification and mapping of human papillomavirus type 1 RNA transcripts recovered from plantar warts and infected epithelial cell cultures. J. Virol. **61:**1913–1918.
8. **Crum, C. P., S. Barber, M. Symbula, K. Snyder, A. M. Saleh, and J. K. Roche.** 1990. Coexpression of the human papillomavirus type 16 E4 and L1 open reading frames in early cervical neoplasia. Virology **178:**238–246.
9. **Danos, O., R. C. Mulligan, and M. Yaniv.** 1986. Production of spliced DNA copies of the cottontail rabbit papillomavirus genome in a retroviral vector. CIBA Found. Symp. **120:**68–82.
10. **Dayhoff, M. O., R. M. Schwartz, and B. C. Orcutt.** 1978. A model of evolutionary change in proteins. Matrices for detecting distant relationships, p. 345–358. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure. National Biomedical Research Foundation, Washington, D.C.
11. **Doorbar, J.** 1991. An emerging function for E4. Papilloma Rep. **2:**145–147.
12. **Doorbar, J., D. Campbell, R. J. A. Grand, and P. H. Gallimore.** 1986. Identification of the human papilloma virus-1a E4 gene products. EMBO J. **5:**355–362.
13. **Doorbar, J., I. Coneron, and P. H. Gallimore.** 1989. Sequence divergence yet conserved physical characteristics among the E4 proteins of cutaneous human papillomaviruses. Virology **172:**51–62.
14. **Doorbar, J., S. Ely, N. Coleman, M. Hibma, D. H. Davies, and L. Crawford.** 1992. Epitope-mapped monoclonal antibodies against the HPV 16 E1^E4 protein. Virology **186:**353–359.
15. **Doorbar, J., S. Ely, J. Sterling, C. McClean, and L. Crawford.** 1991. Specific interaction between HPV 16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. Nature (London) **352:**824–827.
16. **Doorbar, J., H. S. Evans, I. Coneron, L. V. Crawford, and P. H. Gallimore.** 1988. Analysis of HPV-1 E4 gene expression using epitope-defined antibodies. EMBO J. **7:**825–833.
17. **Doorbar, J., A. Parton, K. Hartley, L. Banks, T. Crook, M. Stanley, and L. Crawford.** 1990. Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. Virology **178:**254–262.
18. **Fuchs, E.** 1990. Epidermal differentiation: the bare essentials. J. Cell Biol. **111:**2807–2814.
19. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23:**175–182.
20. **Grand, R. J. A., J. Doorbar, K. J. Smith, I. Coneron, and P. H. Gallimore.** 1989. Phosphorylation of the human papillomavirus type 1 E4 proteins *in vivo* and *in vitro*. Virology **170:**201–213.
21. **Holmes, K. C., D. Popp, W. Gebhard, and W. Kabsch.** 1990. Atomic model of the actin filament. Nature (London) **347:**44–49.
22. **Howley, P. M.** 1990. Papillomavirinae and their replication, p. 1625–1650. *In* B. N. Fields et al. (ed.), Virology, 2nd ed. Raven Press, New York.
23. **Jareborg, N., and S. Burnett.** 1991. Immunofluorescent detection of bovine papillomavirus E4 antigen in the cytoplasm of cells permissive *in vitro* for viral DNA amplification. J. Gen. Virol. **72:**2269–2274.
24. **Kabsch, W., H. G. Mannherz, D. Suck, E. F. Pai, and K. C. Holmes.** 1990. Atomic structure of the actin: DNase I complex. Nature (London) **347:**37–44.
25. **Kozak, M.** 1986. Bifunctional messenger RNAs in eukaryotes. Cell **59:**67–79.
26. **Lane, E. B.** 1982. Monoclonal antibodies provide specific intramolecular markers for the study of epithelial tonofilament organization. J. Cell Biol. **92:**665–673.
27. **Mitchison, H. M., R. J. A. Grand, P. J. Byrd, G. D. Johnson, A. Parton, and P. H. Gallimore.** 1990. The expression of the adenovirus 12 early region 1B 19K protein using a recombinant simian virus 40 system. J. Gen. Virol. **71:**1713–1722.
28. **Nasseri, M., R. Hirochika, T. R. Broker, and L. T. Chow.** 1987. A human papilloma virus type 11 transcript encoding an E1^E4 protein. Virology **159:**433–439.
29. **Neary, K., B. H. Horwitz, and D. DiMaio.** 1987. Mutational analysis of open reading frame E4 of bovine papillomavirus type 1. J. Virol. **61:**1248–1252.
30. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48:**443–453.
31. **Palefsky, J. M., B. Winkler, J. P. Rabanus, C. Clark, S. Chan, V. Nizet, and G. K. Schoolnik.** 1991. Characterization of *in vivo* expression of the human papillomavirus type 16 E4 protein in cervical biopsy tissues. J. Clin. Invest. **87:**2132–2141.
32. **Palermo-Dilts, D. A., T. R. Broker, and L. T. Chow.** 1990. Human papillomavirus type 1 produces redundant as well as polycistronic mRNAs in plantar warts. J. Virol. **64:**3144–3149.
33. **Rayment, I., H. M. Holden, M. Whittaker, C. B. Yohn, M. Lorenz, K. C. Holmes, and R. A. Milligan.** 1993. Structure of the actin-myosin complex and its implications for muscle contraction. Science **261:**58–65.

33a. **Roberts, S., I. Ashmole, and P. H. Gallimore.** Unpublished results.

34. **Roberts, S., I. Ashmole, G. D. Johnson, J. W. Kreider, and P. H. Gallimore.** 1993. Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells. Virology **197:**176–187.
35. **Roberts, S., I. Ashmole, T. M. T. Sheehan, A. H. Davies, and P. H. Gallimore.** 1994. Human papillomavirus type 1 E4 protein is a zinc-binding protein. Virology **202:**865–874.

35a. **Roberts, S., and P. H. Gallimore.** Unpublished data.

36. **Rogel-Gaillard, C., F. Breitburd, and G. Orth.** 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:**816–823.
37. **Rogel-Gaillard, C., G. Pehau-Arnaudet, F. Breitburd, and G. Orth.** 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.
38. **Shah, K. V., and P. M. Howley.** 1990. Papillomaviruses, p. 1651–1675. *In* B. N. Fields et al. (ed.), Virology, 2nd ed. Raven Press, New York.
39. **Stamps, A. C., and M. S. Campo.** 1988. Mapping of two novel transcripts of bovine papillomavirus type 4. J. Gen. Virol. **69:**3033–3045.
40. **Sterling, J. C., J. N. Skepper, and M. A. Stanley.** 1993. Immunoelectron microscopical localization of human papillomavirus type 16 L1 and E4 proteins in cervical keratinocytes cultured *in vivo*. J. Invest. Dermatol. **100:**154–158.
41. **Stewart, M.** 1993. Intermediate filament structure and assembly. Curr. Opin. Cell Biol. **5:**3–11.
42. **Taylor, J. W., J. Ott, and F. Eckstein.** 1985. The rapid generation of oligonucleotide-directed mutations at high frequency using phosphothioate-modified DNA. Nucleic Acids Res. **13:**8764–8785.
43. **Tomita, Y., A. Fuse, H. Sekine, H. Shirasawa, B. Simizu, M. Sugimoto, and S. Funahashi.** 1991. Human papillomavirus type 6 and 11 E4 gene products in condyloma acuminata. J. Gen. Virol. **72:**731–734.