JOURNAL OF VIROLOGY, Nov. 1986, p. 813–816
0022-538X/86/110813-04$02.00/0

Vol. 60, No. 2

The L2 Open Reading Frame of Human Papillomavirus Type 1a Encodes a Minor Structural Protein Carrying Type-Specific Antigens

CAROL A. KOMLY,[1]* FRANÇOISE BREITBURD,[2] ODILE CROISSANT,[2] AND ROLF E. STREECK[1]

Unité des Applications du Génie Génétique[1] and Unité des Papillomavirus,[2] Unité 190 de l'Institut National de la Santé et de la Recherche Médicale, Institut Pasteur, Paris 75724 Cedex 15, France

Received 25 March 1986/Accepted 7 July 1986

The proteins encoded by the open reading frames of papillomavirus genomes and the minor polypeptides detected in purified virions are still poorly defined. We show here by its expression in *Escherichia coli* that the open reading frame L2 of human papillomavirus type 1a codes for a minor structural protein of M_r 76,000. Antisera raised against a truncated L2–β-galactosidase fusion protein in which the conserved N-terminal region of L2 is missing are type specific for human papillomavirus type 1 virions and are reactive at high dilutions. Expression of the L2-encoded type-specific antigens thus provides a powerful new tool for the identification of papillomaviruses.

Papillomaviruses induce benign epithelial lesions in humans and in many animal species, some of which may undergo malignant conversion. Papillomaviruses are highly species and tissue specific. They replicate in vivo in terminally differentiating cells, and no cell culture system has as yet been developed for the multiplication of these viruses. More than 30 different types of human papillomaviruses (HPVs) have been distinguished on the basis of DNA cross-hybridization studies after molecular cloning of their genomes. It seems clear that the different types of HPV-associated lesions constitute distinct diseases caused by one or more specific types of HPV (for reviews, see references 14 and 16). Many HPVs are only known by their cloned genome. The genetic organization of their circular, ≈8-kilobase-long DNA is remarkably similar (3). From studies with bovine papillomavirus (BPV) and by analogy to papovaviruses, an early region and a late region have been distinguished (5, 10). Two open reading frames (ORFs), L1 and L2, have consistently been found in the late region of all papillomavirus genomes sequenced; both of these ORFs have the capacity of encoding proteins of M_r ≈55,000 (3).

Due to the low virus content of many lesions, only a few HPVs have been purified in quantities sufficient for analysis of the viral structural proteins and for the preparation of antisera against the virions (14). Analysis of dissociated virions by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) has revealed a major structural protein of M_r ≈54,000, as well as minor proteins, among which a 68,000- to 76,000-M_r protein is the most prominent, and histonelike proteins (14). Antisera raised against purified intact viral particles do not cross-react with papillomaviruses of a different type and have therefore been termed type specific. On the other hand, antisera raised against dissociated virions recognize antigens common to the papillomavirus group and have therefore been called group specific (8, 13–15).

We have chosen HPV type 1 (HPV1), subtype a, for further study of the relationship between viral structural proteins, ORFs, and type- and group-specific antigens. HPV1 can be purified in large quantities from deep palmoplantar warts, and type- and group-specific antisera have been prepared using HPV1 (13, 15, 20). HPV1a offers the additional advantage that its nucleotide sequence is known (4). By comparison with the late region of other papillomaviruses, the nucleotide sequence of L1 has been found to be highly conserved, whereas L2 differs considerably in different viruses (3). Here we report on the cloning and expression of L2 in *Escherichia coli*, the preparation of antisera against L2-encoded proteins, the identification of the viral structural protein corresponding to L2, and the characterization of the antigenic determinants carried by this protein.

Two plasmids were constructed by the usual methods (11) to obtain expression of the L2 ORF of HPV1a in *E. coli*: pHPL2 and pHPL2 · βgal (Fig. 1). For the construction of pHPL2, a 1206-base-pair *Xho*II-*Xmn*I fragment (Fig. 1a) encoding all but the 106 N-terminal amino acid residues of L2 was inserted into plasmid pCQV2 (19) (Fig. 1b), replacing a nonessential *Bam*HI-*Pvu*II fragment downstream of the bacteriophage λp_R promoter and the *cro* ribosomal binding site present in pCQV2. To construct pHPL2 · βgal, we used plasmid pMC1403 (2), which carries the *lac* operon deleted for the *lac* promoter and the first 22 base pairs of *lacZ*. After elimination of its translational stop codon by BAL 31 nuclease treatment, L2 was fused to *lacZ*, together with all of the λ sequences upstream of L2 present in pHPL2 (Fig. 1b). In-phase L2-βgal fusions expressed under *c*I857/p_R control were selected by virtue of their *lac*$^+$ phenotype at 41°C and *lac*$^-$ phenotype at 30°C on MacConkey agar plates supplemented with lactose (21). One of about 120 clones obtained was chosen for further study. DNA sequence analysis of the L2-*lacZ* junction showed that only the last two C-terminal amino acid residues of L2 were missing in the fusion protein encoded by this plasmid (Fig. 1c).

pHPL2 and pHPL2 · βgal were transformed into *E. coli* CAG1139 (*Δlon100*) cells (7). To allow expression of the L2-encoded proteins, bacterial cells were grown at 30°C to A_{600} values of 0.5 to 0.9 and were then shifted to 41°C for 90 min. After centrifugation, bacteria were lysed by boiling for 10 min in 60 mM Tris hydrochloride (pH 6.8)–5% 2-mercaptoethanol–2% SDS–26% glycerol–0.03% bromophenol blue. Proteins were subjected to 10% SDS-PAGE (9), and L2-encoded proteins were identified by Western blotting (1) by using polyclonal antisera against intact or disrupted HPV1 virions (13, 15). These antisera detect the major (54,000 M_r) and the minor (76,000 M_r) structural proteins of

* Corresponding author.

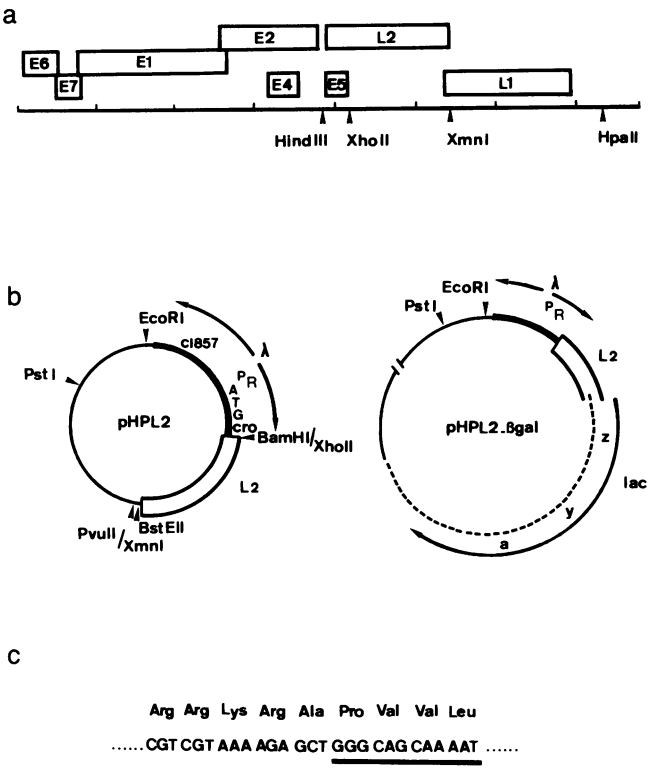

FIG. 1. HPV1a genome and plasmids for expression of L2. (a) Distribution of ORFs in HPV1a DNA (from reference 4). Restriction sites used for subcloning are indicated. (b) Expression vectors. L2-encoded segments are represented as an open box, phage λ sequences upstream of L2 are represented as heavy lines, pBR322 sequences are represented as thin lines, and the *lac* operon region from pMC1403 is represented as a dotted line. Arrows indicate directions of transcription. (c) Nucleotide and derived amino acid sequences at the L2-*lacZ* fusion site. Nucleotides from pMC1403 are underlined.

HPV1 (13, 20) (Fig. 2, lane d). Major bands corresponding to proteins of M_r ≈70,000 and ≈170,000 were detected in lysates of *E. coli* harboring pHPL2 and pHPL2 · βgal, respectively, when the bacteria were grown at 41°C (lanes b and c). These proteins were not detected when the bacteria did not carry a plasmid (lane a) or were grown at 30°C (data not shown). They were also not found with a polyclonal antiserum against dissociated BVP type 1 (BPV1) particles which, under these conditions, only detects the major (M_r 54,000) structural protein (data not shown).

The 170,000-M_r protein is in the range expected for the L2-βgal fusion, but the M_r of 70,000 determined by electrophoretic mobility for the pHPL2-encoded protein exceeds the 51,200 M_r calculated from the nucleotide sequence. The migration of this protein was not affected by treating the cell lysate with high concentrations of urea and SDS. We found that a readthrough of the L2 stop codon into pBR322 sequences could also not account for a protein of M_r 70,000.

For further characterization of the antigenic and immunogenic properties of the L2-encoded protein, the L2-βgal fusion protein obtained in *E. coli* was purified in one step by affinity column chromatography (22). Proteins binding to the column were eluted and analyzed by SDS-PAGE and Western blotting (data not shown). Three major proteins with M_rs of ca. 170,000, 110,000, and 60,000 were detected, of which only the 170,000-M_r protein reacted with anti-HPV1 antiserum. The 110,000-M_r protein comigrated with authentic β-galactosidase. The 60,000-M_r protein was probably a fortuitous contaminant not encoded by pHPL2 · βgal (see below). Antisera were raised against the mixture of purified proteins eluted from the affinity column by immunization of female Hartley guinea pigs (Cesal, Mont-Medy, France) with subcutaneous injections of about 150 μg of protein each in complete Freund adjuvant (Difco Laboratories, Detroit, Mich.) at intervals of 2 to 3 weeks. Serum was collected 1 week after the third injection. The antisera obtained were able to recognize in Western blotting experiments the same 70,000- and 170,000-M_r proteins in lysates of *E. coli* harboring pHPL2 and pHPL2 · βgal, respectively (Fig. 3, lanes b and c), that had been detected with antiserum to HPV1 virions (Fig. 2, lanes b and c).

A major protein band found in both lysates and in *E. coli* lacking any of the plasmids (Fig. 3, lanes a, b, and c) was likely to correspond to the 60,000-M_r *E. coli* protein fortuitously copurified with the L2-βgal fusion proteins. The only protein detected with the L2-βgal antiserum in dissociated HPV1-induced warts (lane d) and in purified HPV1 particles (lane e) was the 76,000-M_r minor structural protein of HPV1 virions.

The specificity of the L2-βgal antisera was further characterized by an indirect immunofluorescence technique by using sections of fixed and paraffin-embedded warts induced by different types of HPVs (Fig. 4). The L2-βgal antiserum stained the nuclei of differentiating cells of warts induced by HPV1 (Fig. 4a) in the same way as an HPV1 type-specific antiserum (Fig. 4b). The optimal dilutions were the same for both sera (1:500), but the intensity of fluorescence was lower for the L2-βgal antisera. No reaction was observed with sections of HPV2-induced warts (Fig. 4c), which showed an intense nuclear staining with an HPV2 type-specific antiserum (Fig. 4d) in adjacent sections. In addition, no reaction was observed with lesions of epidermodysplasia verruciformis, which are associated with several different HPV types. The genomes of epidermodysplasia verruciformis-specific HPVs have been shown to exhibit partial sequence

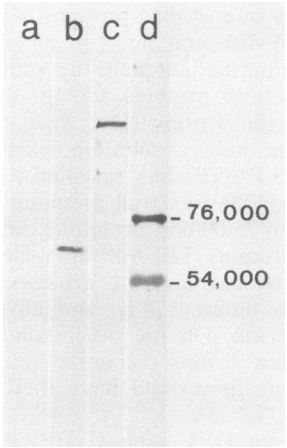

FIG. 2. Expression of L2-encoded proteins in *E. coli*. Bacterial cell lysates were separated by SDS-PAGE and analyzed by Western blotting by using a polyclonal antiserum against intact HPV1 virions. Lysates were prepared from *E. coli* CAG1139 alone (lane a) or harboring pHPL2 (lane b) or pHPL2 · βgal (lane c) induced at 41°C. Dissociated purified HPV1 virions were run in the same gel (lane d). Molecular weights of the structural proteins of the virion are indicated.

homology with short segments of the L2 region of HPV1 as determined by DNA heteroduplex studies (12). The L2-βgal antiserum showed no reaction with intact HPV1 particles by immunodiffusion techniques under conditions in which type-specific antisera showed a precipitation line (data not shown) and reacted only at a higher concentration (1:25 dilution) with acetone-fixed frozen sections of HPV1-induced warts in which antigens were less denatured than in fixed and paraffin-embedded warts (data not shown).

These data present evidence that the minor 76,000-M_r protein of HPV1 virions is encoded, at least in part, by the L2 ORF of the viral genome. The reason for the discrepancy between the 55,300 M_r deduced from the nucleotide sequence of L2 (4) and the determined M_r value (76,000) of the L2-encoded protein is not clear. It could be due to splicing of L2 to another viral ORF, as previously suggested (3, 20), to posttranslational modification, or to an unusually low electrophoretic mobility of the L2-encoded polypeptide. Since a similar discrepancy is found for the pHPL2-encoded truncated L2 protein expressed in *E. coli*, the aberrant molecular weight observed is probably due to an intrinsic property of the L2 polypeptide.

Since L2 is present in all of the papillomavirus genomes sequenced, it is tempting to propose that an L2-encoded minor protein is a genuine constituent of all virions. Given the exclusive reaction of the L2-βgal antiserum with the 76,000-M_r minor structural protein, the possibility that the 54,000-M_r major structural protein might be a mixture of L1- and L2-encoded proteins can now clearly be discarded. This finding strongly suggests that the major structural protein is encoded by L1. This conforms to unpublished data reported by Engel et al. (5) that transcripts from BPV1-induced papillomas including all of the late region (L1 and L2) directed the in vitro synthesis of this protein and is further supported by the demonstration that antisera against BPV1 particles precipitated an L1-encoded protein expressed in *E. coli* (18).

Our data show that the 76,000-M_r L2-encoded protein carries antigens which appear to be type specific. However, these antigens may not be readily accessible on the outside of the intact virion but instead may be somewhat masked in the capsid. The localization of type-specific antigens on the L2-encoded protein is well matched by the finding that L2 is among the least-conserved ORFs in papillomavirus genomes. The only sequence conserved to an appreciable extent in the L2 ORF codes for about 60 N-terminal amino acids of the corresponding protein (3). It is conceivable that group-specific antigens may be carried by this sequence, which is not present in our construction.

The association of certain types of HPV with some forms of human cancers suggests that patients infected with these HPV types are at risk. At present, HPVs found in lesions are typed by DNA hybridization techniques which are not yet easily adaptable to routine screening. Our results suggest the possibility of developing a new method for typing HPVs found in lesions, i.e., by using type-specific antisera against the L2-encoded protein. Since L2 is present in all papillomavirus genomes thus far sequenced, the approach should be generally applicable, even for those papillomaviruses for which virions have never been isolated and for which such type-specific antisera cannot be obtained otherwise. The further characterization of epitopes by this approach may lead to the identification of antigenic determinants common to high-risk virus groups and not present in low-risk groups.

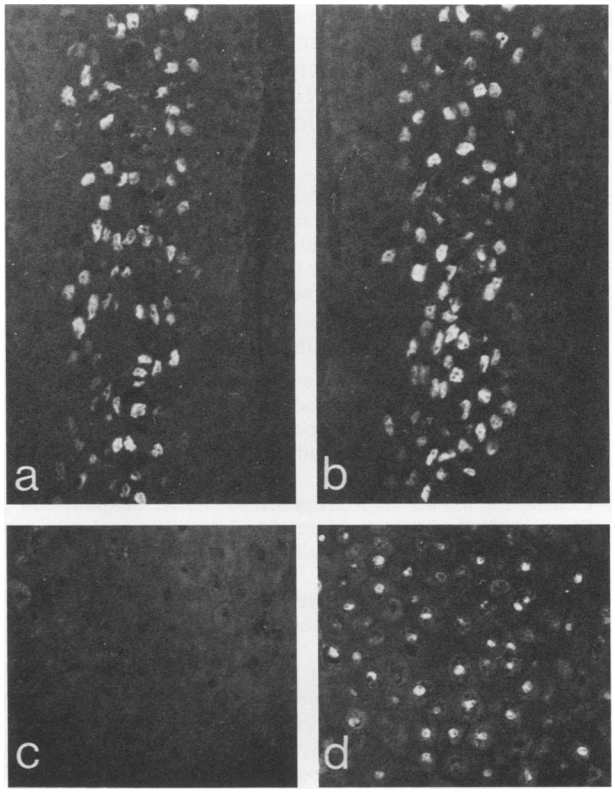

FIG. 4. Detection of antigen in HPV1-induced warts. Serial sections (7 μm thick) were obtained from specimens of HPV1-induced (lanes a and b) and HPV2-induced (panels c and d) warts fixed in Carnoy's fluid and embedded in paraffin. An indirect immunofluorescence test was performed as described previously (13, 15) by using guinea pig antisera raised against the L2-βgal fusion protein (a and c) or intact HPV1 (b) and HPV2 (d) particles. Fluorescent labeling is observed in the nuclei of the differentiating wart cells (13, 15) Magnification, ×130.

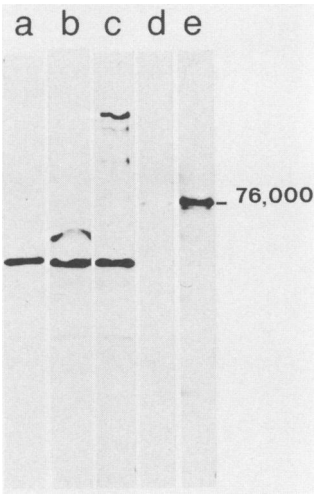

FIG. 3. Identification of proteins recognized by an L2-βgal antiserum. A Western blot with a polyclonal antiserum against the mixture of purified L2-βgal proteins is shown (lanes a to c). Shown are lysates of *E. coli* CAG1139 alone (lane a) or harboring pHPL2 (lane b) or pHPL2 · βgal (lane c), HPV1-induced warts (lane d), and purified HPV1 virions (lane e), dissociated as described previously (20). The molecular weight of the minor structural protein of the virion is indicated.

Antisera raised against proteins of the late region of BPV expressed in *E. coli* have been shown to be neutralizing in vitro (18). Expression of selected ORFs may therefore play an important role in the development of subunit vaccines against specific types of papillomaviruses.

We are particularly grateful to G. Orth for helpful suggestions and discussions. We thank O. Danos and M. Katinka for providing the cloned HPV1a genome, M. Favre for HPV1 particles, and M. Yaniv for his interest in the study. We also thank B. Lemercier for DNA sequence analysis, C. Prévost for help in guinea pig immunization, and C. Turcat and A. Galliot for secretarial work.

LITERATURE CITED

1. **Burnette, W. N.** 1981. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal. Biochem. **112**:195–203.
2. **Casadaban, M. J., A. Martinez-Arias, S. K. Shapira, and J. Chou.** 1983. Beta-galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100**:293–308.
3. **Danos, O., I. Giri, F. Thierry, and M. Yaniv.** 1984. Papillomavirus genomes: sequences and consequences. J. Invest. Dermatol. **83**:7–11.
4. **Danos, O., M. Katinka, and M. Yaniv.** 1982. Human papillomavirus 1a complete sequence: a novel type of genome organization among *Papovaviridae*. EMBO J. **1**:231–236.
5. **Engel, L. W., C. A. Heilman, and P. M. Howley.** 1983. Transcriptional organization of bovine papillomavirus type 1. J. Virol. **47**:516–528.
6. **Favre, M., F. Breitburd, O. Croissant, and G. Orth.** 1977. Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses. J. Virol. **21**:1205–1209.
7. **Grossman, A. D., R. R. Burgess, W. Walter, and C. A. Gross.** 1983. Mutations in the *lon* gene of *E. coli* K12 phenotypically suppress a mutation in the sigma subunit of RNA polymerase. Cell **32**:151–159.
8. **Jenson, A. B., J. D. Rosenthal, C. Olson, F. Pass, W. D. Lancaster, and K. Shah.** 1980. Immunologic relatedness of papillomaviruses from different species. J. Natl. Cancer. Inst. **64**:516–528.
9. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
10. **Lowy, D. R., I. Dvoretzky, R. Shober, M. F. Law, L. Engel, and P. M. Howley.** 1980. In vitro tumorigenic transformation by a defined subgenomic fragment of bovine papillomavirus DNA. Nature (London) **287**:72–74.
11. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Orth, G.** 1986. Epidermodysplasia verruciformis: a model for understanding the oncogenicity of human papillomaviruses. CIBA Found. Symp. **120**:157–169.
13. **Orth, G., F. Breitburd, and M. Favre.** 1978. Evidence for antigenic determinants shared by the structural polypeptides of Shope rabbit papillomavirus and human papillomavirus type 1. Virology **91**:243–255.
14. **Orth, G., and M. Favre.** 1985. Human papillomaviruses: biochemical and biological properties. Clin. Dermatol. **3**:27–42.
15. **Orth, G., M. Favre, and O. Croissant.** 1977. Characterization of a new type of human papillomavirus that causes skin warts. J. Virol. **24**:108–120.
16. **Pfister, H.** 1984. Biology and biochemistry of papilloma-viruses. Rev. Physiol. Biochem. Pharmacol. **99**:112–181.
17. **Pfister, H., U. Linz, L. Gissmann, B. Huchthausen, D. Hoffmann, and H. Zur Hausen.** 1979. Partial characterization of a new type of bovine papillomavirus. Virology **96**:1–8.
18. **Pilacinski, W. P., D. L. Glassman, R. A. Krzyzek, P. L. Sadowski, and A. K. Robbins.** 1984. Cloning and expression in *Escherichia coli* of the bovine papillomavirus L1 and L2 open reading frames. Bio/Technology **2**:356–360.
19. **Queen, C.** 1983. A vector that uses phage signals for efficient synthesis in *Escherichia coli*. J. Mol. Appl. Genet. **2**:1–10.
20. **Roseto, A., P. Pothier, M. C. Guillemin, J. Peries, F. Breitburd, N. Bonneaud, and G. Orth.** 1984. Anti-HPV1 monoclonal antibodies. J. Gen. Virol. **65**:1319–1324.
21. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Ullmann, A.** 1984. One-step purification of hybrid proteins which have β-galactosidase activity. Gene **29**:27–31.