JOURNAL OF VIROLOGY, Feb. 1994, p. 619–625
0022-538X/94/$04.00+0

Vol. 68, No. 2

Interaction of Human Papillomavirus (HPV) Type 16 Capsid Proteins with HPV DNA Requires an Intact L2 N-Terminal Sequence

JIAN ZHOU, XIAO-YI SUN, KYLIE LOUIS, AND IAN H. FRAZER*

Papillomavirus Research Unit, University of Queensland, Princess Alexandra Hospital, Woolloongabba, Queensland 4102, Australia

Received 23 August 1993/Accepted 20 October 1993

Encapsidation of papillomavirus DNA involves DNA-protein and protein-protein interactions. We sought to define the role of each human papillomavirus (HPV) capsid protein in HPV DNA encapsidation. HPV16 major (L1) and minor (L2) capsid proteins purified from recombinant vaccinia virus-infected cells were compared for their ability to bind nucleic acids. L2 protein, but not L1 protein, could bind HPV DNA. To map the DNA-binding region of L2, a series of truncated or point-mutated L2 protein open reading frames were used to show that only the N terminal of L2 was required for L2-DNA binding. This interaction depends critically on charged amino acids (Lys or Arg) in the first 12 amino acids of the N terminal of the protein. Several techniques were used to show that L2 interaction with DNA did not require specific DNA sequences. We propose that HPV L2 protein may play a major role in papillomavirus capsid assembly by introducing HPV DNA to the virus particles formed by the self assembly of the L1 major structural protein.

Human papillomaviruses (HPVs) are a family of small DNA viruses which induce benign hyperproliferative lesions of the cutaneous and mucosal epithelia. Of the 70 different virus types which have been identified, more than 20 are associated with anogenital lesions (9). Regulatory signals for HPV DNA replication and for transcription have been analyzed in great detail (4, 6). However, studies on papillomavirus (PV) virion assembly and PV DNA encapsidation are limited because no tissue culture system readily supports PV propagation and because lesions associated with infection with HPV16, the major genotype associated with anogenital malignancy (40), contain few PV particles and low levels of viral structural proteins.

PV capsids comprise two virally encoded structural proteins, designated L1 and L2, which are assembled onto a DNA-protein complex (13). A single virus capsid is a T=7d icosahedron composed of 72 pentameric capsomeres, each of which contains five molecules of the major capsid protein, L1 (1, 12). The minor capsid protein, L2, is present at approximately 1/10 the abundance of L1 (10) and has an unknown structural role. L1 protein is directed to the nucleus by a C-terminal nuclear localization signal (36); virus assembly occurs in the nucleus (28, 29). Recombinant L1 protein self-assembles into particles resembling virus capsids (37), but assembly is enhanced in the presence of L2 protein, which may be required for assembly of infectious virions (15, 39).

As a step toward the definition of the mechanism of PV assembly and of encapsidation of the viral minichromosome, we characterized the interaction of each of the capsid proteins of HPV16 with nucleic acid by using purified L1 and L2 proteins expressed from recombinant vaccinia viruses (rVV).

* Corresponding author. Mailing address: Papillomavirus Research Unit, University of Queensland, Princess Alexandra Hospital, Woolloongabba, Queensland 4102, Australia. Phone: 61 (7) 240-2251. Fax: 61 (7) 240-2048. Electronic mail address: frazer2@pahosp.uq.oz.au.

MATERIALS AND METHODS

Plasmid construction. For expression in VV, the open reading frame corresponding to the 474-amino-acid HPV16L2 coding region was amplified by PCR from a plasmid containing the HPV16 genome. The 5′ primer introduced a *Bam*HI site upstream from the L2 open reading frame ATG, and the 3′ primer introduced a *Sma*I site beyond the termination codon. The amplified L2 fragment was recovered by elution from an agarose gel, cut with *Bam*HI and *Sma*I, and ligated into plasmid RK19 (16), creating RK19/16L2, in which L2 expression is driven by the VV late promoter 4b. The HPV16L2 and 4b promoter were then transferred to the VV expression vector pSX3 (39) for rVV construction.

To create the simplified vector pUC18/4b16L2 and facilitate transfer of mutant L2 genes between vectors used for VV expression, a Klenow-blunted *Mlu*I-*Eco*RI fragment carrying the VV 4b promoter and the whole HPV16L2 open reading frame was cleaved from RK19/16L2 and inserted into pUC18. This plasmid was used as the DNA template for PCR amplifications with primers designed to create C-terminal truncations of L2. Mutants Δ374, Δ384, Δ394, Δ404, and Δ414, C-terminally truncated to residues 374, 384, 394, 404, and 414 of the L2 protein, respectively, were created by using a common 5′ primer (M13RSP) (36) and a panel of 3′ primers introducing stop codons (TAA) at codons 374, 384, 394, 404, and 414. To create N-terminal truncation and point mutations, we used a panel of 5′ primers. N-terminal deletions (Δ1–2, Δ1–3, Δ1–4, Δ1–5, Δ1–6, Δ1–7, Δ1–8, Δ1–9, Δ1–10, Δ1–11, Δ1–12, Δ1–13, Δ1–14, Δ1–15, Δ1–20, Δ1–40, Δ1–60, Δ1–80, and Δ1–100) were created by using the set of primers which introduced ATG codons at positions corresponding to amino acids 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 40, 60, 80, and 100, respectively. For N-terminal point mutations (designated H3P; K4P; 2,4,5,N; R5P; S6N; S6P; A7P; K8P; R9P; and 8,9N), amino acids 3, 4, 5, 6, 7, 8, and 9 were changed to either proline (Pro) or asparagine (Asn) by the mismatched-primer method (36). All mutations were confirmed by direct sequencing of the expression plasmids.

Cells and virus. CV-1 cells were maintained in Dulbecco's

modified Eagle's medium (GIBCO) supplemented with 10% fetal or newborn bovine serum (CSL, Melbourne, Australia). Plaque-purified isolates of rVVs were propagated in CV-1 cells grown in Dulbecco's modified Eagle's medium supplemented with 2.5% fetal bovine serum.

rVV construction. We used previously described methods (39) for rVV construction. Briefly, plasmids including the HPV16L2 gene with various mutations driven from the VV late promoter 4b, the *Escherichia coli gpt* gene (7, 11) as a selectable marker, and flanking fragments of the VV B24R gene (18, 31) or thymidine kinase (TK) gene were transfected into VV WR strain-infected (0.05 PFU per cell) CV-1 cells by calcium phosphate precipitation. Virus plaques were purified twice in CV-1 cells in the presence of mycophenolic acid at a concentration of 25 μg/ml.

Immunoprecipitation of L1 and L2 proteins. CV-1 cells were infected with HPV16L1 rVV or HPV16L2 rVV at a multiplicity of infection of about 20 PFU per cell. At 48 h, 5×10^5 infected cells were lysed with RIPA buffer (150 mM NaCl, 1% Nonidet P-40, 0.5% deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 50 mM Tris [pH 8.0]). Lysed cells were centrifuged briefly at 12,000 × *g*, and the supernatant was used for immunoprecipitation. Immunoprecipitation were carried out with a 1:20 dilution of monoclonal anti-HPV16L1 antibody (24) or a 1:2,000 dilution of rabbit anti-HPV16L2 antibody (provided by D. A. Galloway). The precipitated L1 or L2 protein was collected with protein A-Sepharose beads and washed four times in RIPA buffer. Proteins were removed from protein A-Sepharose beads by boiling in polyacrylamide gel electrophoresis (PAGE) sample buffer and separated by SDS-PAGE for analysis.

Southwestern assays. The Southwestern (DNA-protein) assays were based on previously published procedures (23, 26). Immunoprecipitated HPV16L1 and HPV16L2 proteins were separated on a SDS–10% polyacrylamide gel and transferred to a nitrocellulose filter by electroblotting. Filters were blocked with blocking buffer (10 mM Tris [pH 7.5], 5% nonfat skim milk, 10% glycerol, 2.5% Nonidet P-40, 0.1 mM dithiothreitol [DTT], 150 mM NaCl) at 4°C for 12 h. The filters were then washed with binding buffer (10 mM Tris [pH 7.5], 40 mM NaCl, 1 mM EDTA, 1 mM DTT, 8% glycerol, 0.125% skim milk). ^{32}P-labeled probes were added to binding buffer, and incubation was continued for 4 h at 4°C. The filters were washed with five changes of binding buffer. After being air dried, they were wrapped and exposed to X-ray films. They were subsequently reprobed with anti-HPV16L1 or anti-HPV16L2 antiserum and ^{125}I-protein A to confirm protein transfer.

Binding specificity assay. DNA sequences binding specifically to L2 were selected from a pool of double-stranded, 76-mer oligonucleotides (R76) containing a central stretch of 26 random base pairs, flanked by two unique sequences of 25 bp each (33). The sequences of these oligonucleotides were

R76 5′-
CAGGTCAGATCAGCGGATCCTGTCG(N)$_{26}$GAGGCGAATTCAGTGCATGTGCAGC-3′
Forward primer 5′-GCTGCACATGCACTGAATTCGCCTC-3′
Back primer 5′-CAGGTCAGATCAGCGGATCCTGTCG-3′

Random oligonucleotide-binding assays were performed essentially as described previously (34). Briefly, L1 and L2 proteins purified by immunoprecipitation were electrophoresed on an SDS-PAGE gel (10% polyacrylamide) and transferred to a nitrocellulose filter. Filters were incubated for 4 h at 4°C in buffer A (20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES]-HCl [pH 7.9], 1 mM DTT, 10% glycerol, 0.01% Nonidet P-40) containing 50 mM KCl and 5% skim milk. The ^{32}P-labeled pool of random oligonucleotides was prepared by primed synthesis with the forward primer annealed to the random 76-mer oligonucleotide template, added to the filter, and incubated overnight at 4°C. Washes were performed at 4°C in buffer A adjusted to 100 mM KCl (two 10-min washes) and then in buffer A adjusted to 200 mM KCl (one 10-min wash). Filters were autoradiographed, the area of the filter corresponding to bound DNA was excised, and the DNA was eluted by heating at 100°C in water. Eluted DNA was amplified by 30 cycles of PCR with forward and backward primers and purified on 2% agarose gels. For subsequent rounds of binding, the 76-bp product was labeled by 18 cycles of PCR in the presence of ^{32}P-labeled nucleotide, as described previously (33). DNA from the fifth round of selection which remained bound to L2 after washing in buffer A was eluted and amplified as described above, cloned into pUC18, and sequenced.

Immuno-DNA binding assay. Extracts from HPV16L2 rVV-infected cells were prepared and immunoprecipitated as described above. Immune complexes attached to protein A-Sepharose beads were incubated with *Bam*HI-*Pst*I-digested HPV16 DNA for 2 h at 4°C in DNA-binding buffer (10 mM Tris [pH 7.4], 100 mM NaCl, 1 mM $MgCl_2$, 1 mM EDTA, 8% glycerol, 1 mM DTT, 5% skim milk). Following incubation, the beads were washed five times with the same buffer at room temperature. Protein-DNA complexes were eluted in 1% SDS–5 mM EDTA at 65°C. DNA was extracted with phenol twice and precipitated with ethanol. Samples were run on a 1.5% agarose gel and blotted onto nylon membranes. The DNA bound to the membranes was detected by Southern blotting with ^{32}P-labeled HPV16 DNA digested with *Pst*I-*Bam*HI.

DNA sequencing. Dideoxy DNA sequencing was performed with the Sequenase version 2.0 kit (U.S. Biochemicals, Cleveland, Ohio). To denature double-stranded DNA for sequencing, we incubated 2 μg of DNA for 30 min at 37°C in 200 mM sodium hydroxide. DNA was ethanol precipitated, and sequencing carried out as specified by the manufacturer.

RESULTS

HPV16L2 protein binds DNA. To investigate the DNA-binding activity of PV structural proteins, we expressed HPV16L1 or HPV16L2 in eukaryotic cells by using rVV. L1 and L2 proteins were purified from cell lysates by immunoprecipitation, separated by SDS-PAGE (10% polyacrylamide), and transferred to nitrocellulose filters. Binding of these proteins to DNA was investigated by Southwestern blotting with ^{32}P-labeled PV genomic DNA and bacteriophage λ DNA in a buffer containing 40 mM NaCl. HPV16L2 protein bound both HPV genomic DNA and λ DNA (Fig. 1B, lanes 2 to 5). HPV16L1 protein, in contrast, failed to bind any labeled DNA (Fig. 1A, lanes 2 to 5), although the purified L1 protein was detectable on the filters by using an anti-HPV16L1 monoclonal antibody (Fig. 1A lanes 1 and 6). Further experiments in buffers containing up to 150 mM NaCl showed no binding of DNA to HPV16L1 (data not shown). These results suggest that HPV16L2 protein, but not HPV16L1 protein, contains DNA-binding sequences and that the recognition of DNA by HPV16L2 may not be sequence specific.

HPV16L2 N terminus is important for DNA binding. To identify the protein sequence responsible for binding of L2 to DNA, we first checked for DNA-binding motifs in the predicted amino acid sequence of HPV16L2. HPV16L2 has two highly charged regions, rich in lysine and arginine. The first makes up the N terminus of the protein (MRHKRSAKRTKR)

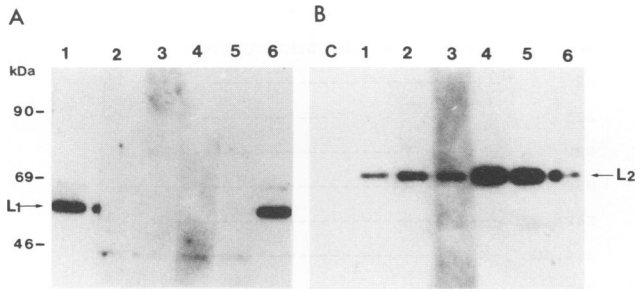

FIG. 1. Characterization of HPV16L1 (A) and HPV16L2 (B) DNA-binding activity. HPV16L1 and HPV16L2 proteins, from rVV-infected CV-1 cells, were immunoprecipitated by L1- or L2-specific antibodies, separated by SDS-PAGE, and transferred to nitrocellulose. Proteins were renatured in blocking buffer containing DTT and incubated with ^{32}P-labeled DNA from HPV16 (lane 2), HPV6b (lane 3), HPV11 (lane 4), or phage λ (lane 5). Lane C is the immunoprecipitate from wild-type VV-infected cells. Unbound DNA was removed, and DNA-binding proteins were detected by autoradiography. The positions of the L1 and L2 proteins, determined by immunoblotting, are shown (lanes 1 and 6).

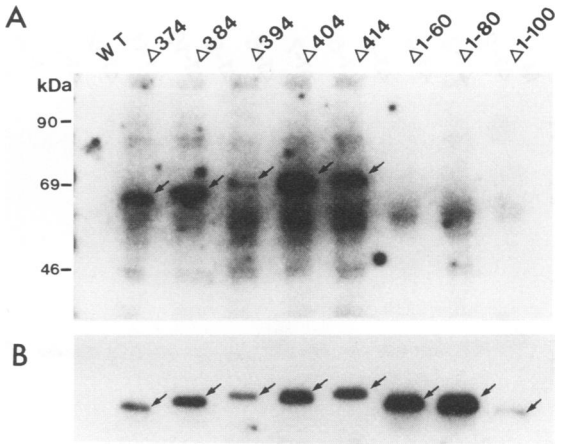

FIG. 2. Definition of the HPV16L2 DNA-binding region. L2 proteins were separated by SDS-PAGE, transferred to a nitrocellulose filter, and probed with ^{32}P-labeled HPV16 DNA (A) or L2-specific antiserum (B). C-terminal amino acid deletions are designated as follows: amino acids 374 to 474 as Δ374, 384 to 474 as Δ384, 394 to 474 as Δ394, 404 to 474 as Δ404, and 414 to 474 as Δ414. Removal of amino acids up to 100 from the C-terminal end had no effect on DNA binding. N-terminal amino acid deletions were designated as follows: amino acids 1 to 60 as Δ1–60, 1 to 80 as Δ1–80, and 1 to 100 as Δ1–100. Removal of any N-terminal sequence diminished DNA binding markedly. The L2 bands are indicated by arrows.

from amino acids 1 to 12; the second lies within the C terminus (RKRRKR) from amino acids 456 to 461.

To determine whether either charged region was involved in DNA binding to L2, we made a series of deletion mutants with mutations in HPV16L2. Constructs encoding various C-terminal or N-terminal mutations of HPV16L2 protein were inserted into plasmid pSX3, which had been previously shown to efficiently direct the synthesis of the HPV16L1 proteins in rVVs (36). CV-1 cells were infected with rVV containing each mutant L2 gene, and cell lysate was analyzed by immunoprecipitation with a rabbit anti-HPV16L2 antibody. The expected relative size of each mutated protein was confirmed by comparing the electrophoretic mobility of truncated proteins with wild-type L2 protein (data not shown).

The deletion mutants were examined for binding to ^{32}P-labeled HPV16 genomic DNA by the Southwestern procedure under conditions associated with binding of native L2 protein. All C-terminal deletion mutants tested, including Δ374, which had the longest deletion, bound HPV DNA in proportion to the amount of immunoreactive L2 protein present (Fig. 2, lane Δ374). To delineate the contribution of the N terminus of L2 to DNA binding, we constructed N-terminal deletion mutants of L2. Three mutants, in which the first 60, 80, and 100 amino acids of L2 were deleted, each failed to bind HPV DNA (Fig. 2, lanes Δ1–60 through Δ1–100). To further characterize the N-terminal DNA-binding region, we tested a series of smaller deletions between amino acids 1 and 15 for their DNA-binding activity. Each of these, including the smallest deletion (which was missing only the arginine residue at position 2), failed to bind DNA (Fig. 3A, lanes Δ1–2 to Δ1–15). These results suggested that critical DNA-binding sequences were in the N terminus of L2 and that binding was dependent on charged amino acids including the arginine at position 2.

L2 protein uses an arginine-rich motif for DNA binding. The amino acid sequence of the N terminus of HPV16L2, starting from position 1, is MRHKRSAKRTKR (one-letter code with charged amino acids underlined). The role of the N terminus of L2 protein in DNA binding was further assessed by site-specific mutagenesis. Ten substitution mutants with mutations of the first 9 amino acids were expressed by rVV, and their DNA-binding activities were examined by Southwestern blotting with ^{32}P-labeled HPV16 genomic DNA. Similar amounts of the various mutant L2 proteins constructed were available for DNA binding, as determined by analysis of immunoreactive L2 protein on the blots, with the exception of K8P (Fig. 3B, lower panel). Mutation of some charged amino acid residues (Lys-4, Lys-8, Arg-9) to Pro or Asn (K4P and 8,9N) abolished DNA binding (Fig. 3B, lanes K4P and 8,9N), while substitution of Arg-5 with Pro (R5P) reduced binding activity (lane R5P). In contrast, substitution of Arg-9 with Pro (R9P) had no effect on DNA binding (lane R9P). Mutation of the neutral amino acids between the Lys-Arg clusters had less effect on DNA-binding activity. Mutations termed H3P, S6P, and A7P, in which substitution of His-3, Ser-6, and Ala-7 for Pro had been produced, showed binding of DNA comparable to that with wild-type L2 (lanes H3P, S6P, and A7P); in contrast, changing Ser-6 to Asn (S6N) abolished DNA binding (lane S6N). These results suggest that the four charged amino acid clusters are important for DNA binding. In each of these charged amino acid clusters, retention of at least one charged amino acid appears necessary for DNA binding. A flexible secondary structure might also be important for L2-DNA interaction because substitution of Ser-6 with Pro and of Arg-5 and Arg-9 did not abolish the DNA binding, whereas the substitution of Ser-6 with Asn removed the L2 DNA-binding function. A summary of L2-DNA interaction results is given in Fig. 4.

L2-DNA interaction has no DNA sequence specificity. We used a library of oligonucleotides and bound these to purified HPV16L2 protein to select for any high-affinity target DNA sequences. The oligonucleotides were random at 26 positions and were flanked by primer and cloning sequences. After incubation of L2 with a pool of these oligonucleotides, oligonucleotides bound to L2 protein were eluted and amplified by PCR for subsequent rounds of selection. Selection and amplification were carried out six times. DNA clones recovered after the selection rounds 5 and 6 were sequenced. Probability theory predicts that among 52 random 26-mer sequences, any

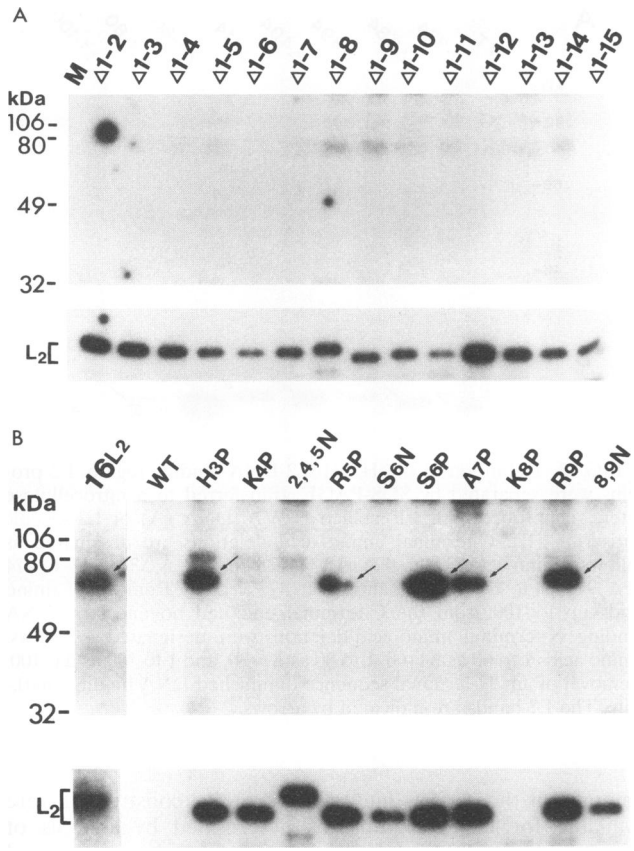

FIG. 3. HPV DNA L2 protein interactions defined by using mutants with L2 N-terminal mutations. (A) N terminus-truncated L2 proteins, with deletions indicated above each lane, were separated by SDS-PAGE, transferred to nitrocellulose membranes, and incubated with ^{32}P-labeled HPV16 DNA (upper panel). The ^{32}P-labeled DNA was removed, and the filter was reprobed with a rabbit anti-HPV16L2 antibody for quantitation of the L2 protein (lower panel). L2 bands are indicated by arrows, and molecular mass markers are indicated on the left. (B) Substitution mutants of L2 proteins, as indicated above each lane, were incubated with ^{32}P-labeled HPV16 genomic DNA (upper panel) or with rabbit anti-HPV16L2 antiserum (lower panel). L2 bands are indicated by arrows, and the molecular mass markers are shown on the left. Substitutions are coded as follows: WT, wild-type VV; H3P, His-3 to Pro; K4P, Lys-4 to Pro; 2,4,5N, Arg-2 to Asn, Lys-4 to Asn, Arg-5 to Asn; R5P, Arg-5 to Pro; S6N, Ser-6 to Asn; S6P, Ser-6 to Pro; A7P, Ala-7 to Pro; K8P, Lys-8 to Pro; R9P, Arg-9 to Pro; 8,9N, Lys-8 to Asn, Arg-9 to Asn.

trinucleotide would be found in 20 clones (95% confidence interval, 15 to 26) and any specificed sequence of 4 nucleotides would be found in 5 of the 26-mers (95% confidence interval, 0 to 9). Among 52 26-mer clones we observed that the most commonly observed trinucleotide (GGG) was present in 24 clones (twice in 5, three times in 1), whereas the most common series of 4 bp (GGGG) was observed in 8 clones. Thus there was no evidence that any short nucleotide sequence was represented among these clones more frequently than would be expected by chance alone. Further, no more complex conserved nucleotide patterns were observed by using standard sequence alignment programs. These results suggested that high-affinity binding between L2 and DNA is a DNA sequence-independent process. Two additional experiments confirmed this observation. First, extract from cells infected with

	^ ^ ^ ^	DNA binding
HPV16L2	MRHKRSAKRTKRASATQLYKTCKQAGTCPP	+
Δ2	----------	-
Δ15	----------	-
R2P	-P----------	-
H3P	--P----------	+
K4P	---P----------	-
2,4,5N	-N-NN----------	-
R5P	----P----------	+
S6N	-----N----------	-
S6P	-----P----------	+
A7P	------P----------	+
K8P	-------P----------	-
R9P	--------P----------	+
8,9N	-------NN----------	-
6,10N	-----N---N----------	-
11,12N	----------NN----------	-
6,10,14N	-----N---N---N----------	-

FIG. 4. Binding of mutant L2 proteins to HPV DNA. For each mutant, the sequence of the protein is given (single-letter code, with conserved amino acids shown as dashes) and the binding of DNA to the protein by Southwestern blot analysis is indicated (+ or −).

HPV16L2 rVV was immunoprecipitated with anti-L2 antibody, and the precipitated protein-antibody complexes, attached to Sepharose beads, were allowed to bind to a mixture of restriction fragments from HPV16 genomic DNA. After unbound DNA was washed away, bound DNA fragments were resolved on a agarose gel and detected by Southern blotting with ^{32}P-labeled HPV16 DNA. An HPV16L2-containing extract, bound to Sepharose beads with anti-L2 antibody, retained each of the HPV16 DNA fragments (Fig. 5, lanes 3, 6, and 7), whereas none of these fragments were bound by an L1 rVV-infected cell extract bound to the bead with anti-L1 antibody (data not shown) or by wild-type VV-infected cell extract bound to the beads with L2-specific antibodies (Fig. 5, lane 2). Second, a fixed amount of L2 protein was incubated with ^{32}P-labeled HPV DNA fragments, in the presence of a 100-fold (lane 4) or 1,000-fold (lane 5) excess of unlabeled λ DNA, and L2, together with any bound DNA, was immunoprecipitated from the mixture by antibody to L2. Phage λ DNA was able to prevent binding of HPV DNA. We conclude that L2 does not interact with DNA in a DNA sequence-specific manner by each of these criteria.

DISCUSSION

The predicted sequence of the carboxy terminus of the L1 and L2 proteins of all sequenced PVs includes several basic amino acids. We demonstrated that the carboxy terminus of the L1 protein of HPV16, which conforms to consensus nuclear localization sequences defined for eukaryotic cells, is necessary for nuclear localization of this protein (36). A basic sequence is also found at the amino terminus of L2, but this does not conform to the nuclear localization sequence motif and resembles the arginine-rich region which has been associated with protein-RNA binding (19). An arginine-rich motif sequence has previously been detected in several prokaryotic and eukaryotic RNA-binding proteins (19, 20), and the VP1

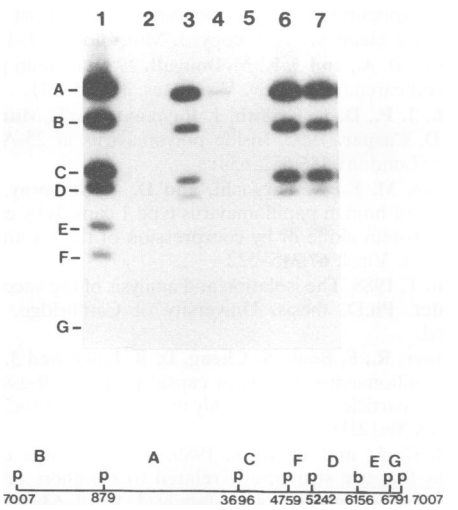

FIG. 5. DNA-binding assay for HPV16L2 proteins from rVVs. L2 protein was immunoprecipitated with anti-L2 antibody. Equal amounts of L2 protein were incubated with *Pst*I-*Bam*HI-cleaved HPV16 genomic DNA. The bound DNA fragments were eluted with 1% SDS and subjected to Southern blotting with ^{32}P-labeled HPV16 DNA (lanes 3, 6, and 7). In some experiments, a 100-fold (lane 4) or 1,000-fold (lane 5) molar excess of phage λ DNA was added to the initial incubation of HPV DNA with L2 protein. Mock assay of a control precipitate from wild-type VV-infected cells is also shown (lane 2). On the left the input DNA fragments are labeled from A to G. A linearized map of the HPV16 DNA is shown below. Restriction sites are *Pst*I (p) and *Bam*HI (b), and the corresponding fragments are labeled A to G according to size.

protein of polyomavirus has been shown to interact with DNA in a nonspecific fashion by using a similar sequence (26). Our data show that the arginine-rich region at the amino terminus of the L2 protein of HPV16 is critical for the binding of L2 to nucleic acid. The basic amino acid sequence at the carboxy terminus of L2 did not contribute to protein-DNA interaction and may, like its counterpart in L1, function as a nuclear localization sequence. The L1 protein of HPV16 did not bind DNA in our system. Polyomavirus uses the major structural protein for DNA interaction, and the minor (VP2 and VP3) proteins of polyomavirus do not seem to interact with DNA; however, other members of the family *Papovaviridiae*, like HPV16, interact with DNA via their minor capsid proteins and have similar basic amino-terminal sequences (8).

Our results suggest that DNA binding to L2 requires multiple basic amino acid sequences working in concert, while the amino acids separating the basic residues also contribute to DNA-binding activity. Similarities between the amino-terminal sequences of PV L2 proteins were identified in searches of the GenBank and EMBL data bases. The basic residues (Arg and/or Lys) are in each case located proximally in the amino-terminal sequence of L2, and their side chains could be positioned so that their positive charges interact with the backbone phosphates of the DNA helix.

To study the intrinsic DNA-binding activity of L1 and L2 proteins, it was necessary to exclude the contribution of cellular proteins to this reaction. We therefore examined binding of HPV DNA, produced in a prokaryotic system, to viral proteins isolated by immunoprecipitation and separated by SDS-PAGE for Southwestern analysis, because this avoided exposure of the HPV DNA to cellular proteins including histones. HPV16L2 protein associated with DNA under our assay conditions, whereas HPV16L1 protein did not. This latter observation is in contrast with previous work (22), which showed that prokaryotic L2 bound DNA and also that a prokaryotic N-terminally truncated L1 fusion protein, with 37% of the sequence from the N terminus removed, had DNA-binding activity. The L1 or L2 protein of HPV16 produced by VV are nonfused proteins which should have correct posttranslational modifications, including glycosylation (38) and phosphorylation (35), which could affect DNA-binding affinity (21, 32). Because the L1 protein we used in this study was a nonfusion protein expressed in eukaryotic cells, we believe that this protein should have correct posttranslational modifications. Further, the binding demonstrated in this assay occurred under physiologic conditions of pH and salt concentration. Thus, our L1-DNA-binding results are more likely to be representative of events within epithelial cells infected with HPV. However, the Southwestern assay system depends on prior immobilization of the test viral protein to positively charged nitrocellulose while the protein is denatured in SDS, and this would favor the demonstration of amino- or carboxy-terminal regions as DNA binding, so we cannot regard L1-DNA interactions under physiological conditions as excluded by our results.

We have previously shown that the minor capsid protein (L2) of HPV16 plays a role in the assembly of infectious PV virions: whereas viruslike particles assemble if L1 protein alone is expressed in eukaryotic cells (17), expression of both L1 and L2 together are necessary for the assembly of infectious virions (37), at least when the proteins are expressed by rVV in vitro. The findings reported here suggest that one role of L2 in the assembly of infectious PV virions may be to link the PV capsid to the viral minichromosome, a function made necessary by a lack of demonstrable interaction of HPV L1 protein with HPV DNA. Since one purpose of any virus capsid is to provide a protective container for the viral genome, capsid proteins are expected to exhibit some potential for direct (basic or other polar amino acid side chains) or indirect (encapsidated polyamines) interaction with nucleic acid (5). This interaction may be important in nucleation of capsid assembly but could also serve a structural purpose in defining the viral architecture for DNA viruses, as is now well recognized for RNA viruses, many of which will assemble virions only in the presence of appropriate viral RNA (2, 5, 25).

Studies of RNA viruses have shown that specific encapsidation initiation sites are present on the viral genome (3). Such sites have also been identified for DNA viruses (27, 30). Our data do not support the hypothesis that packaging selectivity in PV could be achieved by sequence-specific recognition process between viral structural proteins and PV DNA. Lack of specific reactivity also makes a role for the L2-DNA interaction in regulation of transcription, replication, recombination, or viral integration less likely. This finding is perhaps not too surprising, because the expression of capsid proteins occurs in terminally differentiated cells already programmed to undergo apoptotic death.

A model has been proposed for the structure of polyomavirus, which has an identical T=7d icosahedral structure to PV (14). Polyomavirus VP1 protein, which has sequence homology with HPVL1 and will similarly self-assemble into viruslike particles, can interact with VP2 and VP3, which have homologies with PV L2, including an arginine-rich amino terminus. Approximately one molecule of VP2 and VP3 is associated with each VP1 pentamer, just as approximately one molecule of L2 is associated with each L1 pentamer in HPV. The crystal structure of polyomavirus has been partially solved, but the exact binding region of VP2 and VP3 with VP1 is unclear: it

has been proposed that VP2 and VP3, which have DNA-binding activity, facilitate correct and preferential assembly of complete polyomavirus particles by preventing VP1 assembly in the absence of DNA (8). Our data, taken together with this model, suggest that the amino terminus of L2 will be located deep within the viral particle in contact with the viral minichromosome, which is supercoiled round cellular histone proteins. The viral minichromosome will be held in place by an array of L2 proteins, each interacting via its amino terminus with any convenient stretch of viral DNA and, at a more carboxy location, with a pentameric L1 capsomere. In the terminally differentiating PV-infected keratinocyte, HPV DNA would be one of the most abundant short polynucleotides and nuclear DNA would generally be present in the form of condensed cellular chromatin (23). No specific protein-DNA interaction may therefore be required for PV particle assembly. Alternatively, L1 protein (36a) appear to localize in particular areas of the cell nucleus in the presence of L2 protein, at least in vitro. If some mechanism exists for local concentration of L1, L2, and viral DNA, correct viral assembly would follow. Use of L2 deletion mutants as described here should facilitate testing of these alternate hypotheses.

ACKNOWLEDGMENTS

We thank Dalia Tadros for synthesis of oligonucleotides and Denise Galloway for providing rabbit anti-HPV16L2 antibody.

This work was funded in part by NIH grant RO1 CA57789-01, the National Health and Medical Research Council of Australia, the Queensland Cancer Fund, the Mayne Bequest, and the Princess Alexandra Hospital Research and Development Foundation.

REFERENCES

1. **Baker, T. S., W. W. Newcomb, N. H. Olson, L. M. Cowsert, C. Olson, and J. C. Brown.** 1991. Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. Biophys. J. **60:**1445–1456.
2. **Birnbaum, F., and M. Nassal.** 1990. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. J. Virol. **64:**3319–3330.
3. **Blumberg, B. M., C. Giorgi, and D. Kolakofsky.** 1988. N protein of vesicular stomatitis virus selectively encapsidates leader RNA in vitro. Cell **32:**559–567.
4. **Bream, G. L., C.-A. Ohmstede, and W. C. Phelps.** 1993. Characterization of human papillomavirus type 11 E1 and E2 proteins expressed in insect cells. J. Virol. **67:**2655–2663.
5. **Chen, Z. G., C. Stauffacher, Y. Li, T. Schmidt, W. Bomu, G. Kamer, M. Shanks, G. Lomonossoff, and J. E. Johnson.** 1989. Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution. Science **245:**154–159.
6. **Chiang, C.-M., M. Ustav, A. Stenlund, T. F. Ho, T. R. Broker, and L. T. Chow.** 1992. Viral E1 and E2 proteins support replication of homologous and heterologous papillomaviral origins. Proc. Natl. Acad. Sci. USA **89:**5799–5803.
7. **Coupar, B. E. H., M. E. Andrew, and D. B. Boyle.** 1988. A general method for the construction of recombinant vaccinia viruses expressing multiple foreign genes. Gene **68:**1–10.
8. **Delos, S. E., I. Montross, R. B. Moreland, and R. L. Garcea.** 1993. Expression of the polyomavirus VP2 and VP3 proteins in insect cells: coexpression with the major capsid protein VP1 alters VP2/VP3 subcellular localization. Virology **194:**393–398.
9. **de Villiers, E. M.** 1989. Heterogeneity of the human papillomavirus group. J. Virol. **63:**4898–4903.
10. **Doorbar, J., and P. H. Gallimore.** 1987. Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. J. Virol. **61:**2793–2799.
11. **Falkner, F. G., and B. Moss.** 1988. Escherichia coli *gpt* gene provides dominant selection for vaccinia virus open reading frame expression vectors. J. Virol. **62:**1849–1854.
12. **Finch, J. T., and A. Klug.** 1965. The structure of viruses of the papilloma-polyoma type. III. Structure of rabbit papillomavirus, with an appendix on the topography of contrast in negative staining for electron microscopy. J. Mol. Biol. **13:**1–12.
13. **Galloway, D. A., and J. K. McDougall.** 1989. Human papillomaviruses and carcinomas. Adv. Virus Res. **37:**125–171.
14. **Griffith, J. P., D. L. Griffith, I. Rayment, W. T. Murakami, and D. L. D. Caspar.** 1992. Inside polyomavirus at 25-A resolution. Nature (London) **355:**652–654.
15. **Hagensee, M. E., N. Yaegashi, and D. A. Galloway.** 1993. Self-assembly of human papillomavirus type 1 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins. J. Virol. **67:**315–322.
16. **Kent, R. T.** 1988. The isolation and analysis of the vaccinia virus 4b promoter. Ph.D. thesis. University of Cambridge, Cambridge, England.
17. **Kirnbauer, R., F. Booy, N. Cheng, D. R. Lowy, and J. T. Schiller.** 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proc. Natl. Acad. Sci. USA **89:**12180–12184.
18. **Kotwal, G. J., and B. Moss.** 1989. Vaccinia virus encodes two proteins that are structurally related to members of the plasma serine protease inhibitor superfamily. J. Virol. **63:**600–606.
19. **Lazinski, D., E. Grzadzielska, and A. Das.** 1989. Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. Cell **59:**207–218.
20. **Lee, C. Z., J. H. Lin, M. Chao, K. McKnight, and M. M. Lai.** 1993. RNA-binding activity of hepatitis delta antigen involves two arginine-rich motifs and is required for hepatitis delta virus RNA replication. J. Virol. **67:**2221–2227.
21. **Li, L., J. Zhou, G. James, R. Heller Harrison, M. P. Czech, and E. N. Olson.** 1992. FGF inactivates myogenic helix-loop-helix proteins through phosphorylation of a conserved protein kinase C site in their DNA-binding domains. Cell **71:**1181–1194.
22. **Mallon, R. G., D. Wojciechowicz, and V. Defendi.** 1987. DNA-binding activity of papillomavirus proteins. J. Virol. **61:**1655–1660.
23. **McCall, C. A., and J. J. Cohen.** 1991. Programmed cell death in terminally differentiating keratinocytes: role of endogenous endonuclease. J. Invest. Dermatol. **97:**111–114.
24. **McLean, C. S., M. J. Churcher, J. Meinke, G. L. Smith, G. Higgins, M. Stanley, and A. C. Minson.** 1990. Production and characterisation of a monoclonal antibody to human papillomavirus type 16 using recombinant vaccinia virus. J. Clin. Pathol. **43:**488–492.
25. **Miskimins, W. K., M. P. Roberts, A. McClelland, and F. H. Ruddle.** 1985. Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognise the promoter region of the transferrin receptor gene. Proc. Natl. Acad. Sci. USA **82:**6741–6744.
26. **Moreland, R. B., I. Montross, and R. L. Garcea.** 1991. Characterization of the DNA-binding properties of the polyomavirus capsid protein VP1. J. Virol. **65:**1168–1176.
27. **Oppenheim, A., Z. Sandalon, A. Peleg, O. Shaul, S. Nicolis, and S. Ottolenghi.** 1992. A *cis*-acting DNA signal for encapsidation of simian virus 40. J. Virol. **66:**5320–5328.
28. **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.
29. **Pfister, H., and E. Fuchs.** 1987. Papillomaviruses: particles, genome organisation and proteins, p. 1–18. *In* K. Syrjanen, L. Gissmann, and L. G. Koss (ed.), Papillomaviruses and human disease. Springer-Verlag KG, Berlin.
30. **Saragosti, S., S. Cereghini, and M. Yaniv.** 1982. Fine structure of the regulatory region of simian virus 40 minichromosomes revealed by DNAse I digestion. J. Mol. Biol. **160:**133–146.
31. **Smith, G. L., S. T. Howard, and Y. S. Chan.** 1989. Vaccinia virus encodes a family of genes with homology to serine proteinase inhibitors. J. Gen. Virol. **70:**2333–2343.
32. **Sobczyk, A., G. Schyns, N. Tandeau de Marsac, and J. Houmard.** 1993. Transduction of the light signal during complementary chromatic adaptation in the cyanobacterium Calothrix sp. PCC 7601: DNA-binding proteins and modulation by phosphorylation. EMBO J. **12:**997–1004.
33. **Sorger, P. K., P. G. Stockley, and S. C. Harrison.** 1986. Structure

and assembly of turnip crinkle virus. II. Mechanism of reassembly in vitro. J. Mol. Biol. **191**:639–658.

34. **Treacy, M. N., X. He, and M. G. Rosenfeld.** 1991. I-POU: a POU-domain protein that inhibits neuron-specific gene activation. Nature (London) **350**:577–584.
35. **Xi, S.-Z., and L. M. Banks.** 1991. Baculovirus expression of the human papillomavirus type 16 capsid proteins: detection of L1-L2 protein complexes. J. Gen. Virol. **72**:2981–2988.
36. **Zhou, J., J. Doorbar, X. Y. Sun, L. V. Crawford, C. S. McLean, and I. H. Frazer.** 1991. Identification of the nuclear localization signal of human papillomavirus type 16 L1 protein. Virology **185**:625–632.

36a. **Zhou, J., X.-Y. Sun, L. Crawford, and I. H. Frazer.** Co-transportation of HPV L1 + L2 proteins to the nucleus requires a conserved L1 sequence. Submitted for publication.

37. **Zhou, J., D. J. Stenzel, X.-Y. Sun, and I. H. Frazer.** 1993. Synthesis and assembly of infectious bovine papillomavirus particles *in vitro*. J. Gen. Virol. **74**:763–768.
38. **Zhou, J., X. Y. Sun, and I. H. Frazer.** 1993. Glycosylation of human papillomavirus type 16 L1 protein. Virology **194**:210–218.
39. **Zhou, J., X. Y. Sun, D. J. Stenzel, and I. H. Frazer.** 1991. Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. Virology **185**:251–257.
40. **zur Hausen, H.** 1987. Papillomaviruses in human cancer. Cancer **59**:1692–1696.