Interplay of Viral and Cellular Proteins along the Long Control Region of Human Papillomavirus Type 18

ALEJANDRO GARCIA-CARRANCA,¹ FRANÇOISE THIERRY,² AND MOSHE YANIV^{2*}

Genética y Biología Molecular, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Apartado Postal 14-740, 07000 D.F. México,¹ and Unité des Virus Oncogènes, Département de Biologie Moléculaire, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France²

Received 20 May 1988/Accepted 8 August 1988

The long control region of human genital papillomavirus type 18 harbors transcription regulatory elements, such as the E6 promoter and a cell type-specific enhancer independent of the E2 protein. By performing DNase I footprint experiments *in vitro* with protein extracts from different cell lines and tissues we searched for cellular factors interacting with the totality of the control region. We detected a total of eight different protected sites; most of them were found with all the extracts. Two of these sites, one in the enhancer and the other in the sequences proximal to the E6 cap site, interact with a member of the activator protein 1 family. However, in the absence of fine mutational analysis, we cannot readily discern an exact functional role for the different binding sites. We also characterized two regions, which are protected only in the presence of E2, corresponding to the perfect palindromes ACCGN₄CGGT and present either 500 nucleotides or tandemly repeated about 70 nucleotides upstream of the E6 cap site. This last protected area covers a large part of the putative TATA box of the E6 promoter and could explain its repression by bovine papilloma virus type 1 E2, which could interfere with binding of the TFII D ubiquitous transcription factor to the TATA sequence of the promoter.

During the last few years, evidence has accumulated implicating different human papillomaviruses (HPV) as causative agents of human genital cancer (43). Among the human genital papillomaviruses, DNA of either type 16, 18, 31, 33, or 35 (3, 5, 13, 20) has been found in the majority of human cervical carcinomas so far analyzed.

In cervical carcinomas, as well as in cell lines derived from them, viral sequences are usually found integrated into the host genome (25, 42). This integration always disrupts the downstream early region which includes the E1 and E2 open reading frames (ORFs). In the case of established cell lines, it has been shown that transcripts from HPV type 16 (HPV16) and HPV18 map to the E6-E7 part of the early region (2, 29-31). These transcripts are also found in cervical carcinoma biopsies. Their persistence in tumor cells and neoplasias and the accumulating evidence on the transforming activity of the total genome or the E6-E7 region of HPV16 or HPV18 (4, 22, 26a, 28, 38, 41) strongly suggest that these gene products play an important role in the process of malignant transformation.

The genetic organization of the different papillomaviruses already sequenced is very similar and could be related to their biological properties (10). In the genomes of all papillomaviruses a transcription and replication regulatory region, usually called the long control region (LCR) (also known as upstream regulatory region or noncoding region), is located between the end of the late ORF and the beginning of the E6 early gene.

Functional analysis of the HPV18 LCR has provided evidence that it contains several important transcriptional regulatory sequences, including tissue-specific promoter and enhancer elements. The whole LCR can direct transcription from the E6 promoter in certain cell lines from epithelial origin (36). The central part of the LCR contains an E2-independent tissue-specific enhancer (15, 34, 35). Finally, the proximal end contains the early promoter (30, 36).

The E2 ORF from different papillomaviruses encodes a *trans*-activator of early gene expression (33). Its action is pleiotropic, affecting transformation and replication, and requires *cis*-acting sequences located in the LCRs (1, 32). This sequence is a conserved palindromic motif ACC N₆GGT, which is found several times in the LCRs of all sequenced papillomaviruses (11). An important feature of this *trans* activation by the E2 product is that it is able to affect expression from related papillomaviruses (9, 14, 17, 26, 35, 37). The products of the E2 ORF from bovine papillomavirus type 1 (BPV1), HPV1, and HPV11 produced in bacteria interact specifically *in vitro* with DNA sequences present in the LCR of these papillomaviruses (1, 17, 24; N. Dostatni, F. Thierry, and M. Yaniv, EMBO J., in press).

Unlike BPV1, the HPV18 enhancer is not dependent on the presence of the E2 gene product (15, 35). The functional E6 promoter that is active in the absence of E2 is strongly repressed by the intact E2 gene product of BPV1 (37). In contrast, in the context of BPV1 the intact E2 is a strong activator, and only the 3' domain of the E2 ORF encodes a transcriptional repressor (18, 30).

In the present study, we have searched for cellular or viral proteins interacting with the LCR of HPV18. By DNase I footprinting experiments, we assayed nuclear extracts from different cells and detected well-defined protected regions with extracts from cells such as HeLa, SW13 (human adrenocortical cells), or COS-1 (monkey epithelial cells carrying a defective simian virus 40 [SV40] genome). We found four protected domains in the distal part of the LCR, two in the constitutive enhancer element, and two in the proximal promoter. Two of the protected domains, one in the enhancer and the other in the proximal E6 promoter, contain the sequence 5'-TGACTAA-3', which binds a member of the activator protein 1 (AP1) transcription factor family (19, 27). We were not able to assign the other protections to known transcriptional factors.

To analyze the interaction of the E2 of BPV1 with its binding sites in the HPV18 LCR, we used a nuclear extract

* Corresponding author.

prepared after transfection of COS-1 cells with a construction expressing the E2 ORF of BPV1. The only regions of DNA that are protected by the BPV1 E2 gene product contain the conserved motif ACCGN₄CGGT. The fact that in HPV18 two perfect motifs are positioned just upstream from the TATA box suggests a possible mechanism for repression in this system, as opposed to the observed *trans* activation found with other LCRs, in which these elements are positioned further upstream from the cap site.

MATERIALS AND METHODS

Cell growth and transfection. All cells were grown in Dulbecco modified Eagle medium supplemented with 7% fetal calf serum. For transfection experiments, plasmid DNA was purified by two cesium chloride gradient centrifugations. Cells were seeded at a density of 10⁶ per 100-mm dish the day before transfection. Transfections were carried out as previously described (36).

Plasmids and cells. Three basic constructions containing restriction fragments from the HPV18 LCR were used for preparation of labeled fragments to perform footprinting assays. The two *RsaI*-*RsaI* fragments (nucleotides [nt] 7121 to 7509 and 7510 to 7738) and the *RsaI*-*Bam*HI fragment (nt 7739 to 124) were cloned in the polylinker of pSP64. For transfection experiments, the chloramphenicol acetyltransferase (CAT) expression plasmids containing the complete LCR or the central *RsaI*-*RsaI* subfragment, inserted 5' to the SV40 enhancer minus early promoter, have already been described (37). The upstream *RsaI*-*RsaI* subfragment (7121 to 7509) was inserted 5' of the deleted SV40 early promoter, while the *RsaI*-*Bam*HI subfragment (7739 to 124) was cloned 5' to the herpes simplex virus thymidine kinase promoter (at position -109) in CAT expression plasmids in both orientations. The pC59 BPV1 E2 expression plasmid (33) was a generous gift from P. Howley. Cell lines used in this study were previously described (36). All calcium phosphate transfections were done in the presence of a reference plasmid, pRSV- β gal. Both CAT and β -galactosidase activities were measured in extracts prepared after sonication, and the CAT conversion values obtained were corrected for the efficiency of transfection.

Nuclei isolation and preparation of protein extracts. Protein extracts were prepared from confluent dishes or 50 h after transfection for transiently transfected COS-1 cells. Cell nuclei were isolated either by the hypotonic procedure described previously (12) or by the detergent method as follows: cells were washed with phosphate-buffered saline suspended in buffer A (10 mM HEPES [*N*-2-hydroxyethyl-piperazine-*N'*-2-ethanesulfonic acid] [pH 7.9]–85 mM KCl–5.5% sucrose–0.5 mM spermidine–0.15 mM spermine–0.5 mM dithiothreitol–0.5 mM phenylmethylsulfonyl fluoride) and lysed by adding 1 volume of buffer A containing 1% Nonidet P-40. Nuclei were pelleted for 10 min at 2,000 rpm (model HN-S; International Equipment Co.) and washed with buffer A.

Nuclear extracts were prepared essentially as described previously (12) and concentrated as described previously (40). Nuclei were resuspended in buffer C (20 mM HEPES [pH 7.9]–25% glycerol–0.42 M NaCl–1.3 mM MgCl₂–0.2 mM EDTA–0.5 mM dithiothreitol–0.5 mM phenylmethylsulfonyl fluoride) to a final concentration of 3.3×10^7 nuclei per ml. The suspension was stirred for 30 min by gentle rocking at 4°C, and then the chromatin was sedimented at $25,000 \times g$ during 30 min. The soluble proteins were precipitated by slow addition of a 3.8 M (NH₄)₂SO₄ solution containing 20

mM HEPES (pH 7.9) at a ratio of 1.6/1. The mixture was rocked for 30 min at 4°C and centrifuged at 35,000 rpm in an SW56 rotor for 20 min. Proteins were suspended in 20 to 30 μ l per 10⁸ cells of buffer D (20 mM HEPES [pH 7.9]–60 mM KCl–20% glycerol–1 mM dithiothreitol–1 mM phenylmethylsulfonyl fluoride–0.2 mM EDTA) and dialyzed for 2 to 3 h at 4°C against 500 volumes of buffer D. Nuclear extracts were frozen in liquid nitrogen and stored either at –80°C or in liquid nitrogen. Protein concentrations were determined (6) by using bovine serum albumin as the standard.

DNase I footprinting assays. DNA end-labeling and binding reactions were carried out essentially as described previously (7), with the following modifications. End-labeled DNA fragments were separated in thin polyacrylamide gels and eluted by diffusion overnight at 37°C in a solution containing 0.1% sodium dodecyl sulfate–10 mM Tris hydrochloride (pH 7.9)–0.3 M NaCl. The solution was filtered and precipitated with 2-isopropanol. Labeled fragments were suspended in TE (10 mM Tris hydrochloride [pH 7.5], 0.1 mM EDTA) and reprecipitated in the presence of 2 M ammonium acetate.

Binding reactions were performed in a final volume of 10 μ l containing 10 mM HEPES (pH 7.9)–4 mM spermidine–4 mM MgCl₂–30 mM KCl–10% glycerol–0.5 mM dithiothreitol–0.5 mM phenylmethylsulfonyl fluoride–0.1 mM EDTA. Protein concentrations usually varied between 10 and 70 μ g. To prevent nonspecific DNA binding, protein extracts were preincubated for 10 min at 4°C with 250 to 500 ng of poly(dI-dC). DNase I (from a 2 mg/ml stock) was freshly diluted in 0.5 M sucrose–60 mM KCl–15 mM Tris hydrochloride (pH 7.5)–0.15 mM EDTA–0.15 mM EGTA [ethylene glycol-bis(β -aminoethyl ether)-*N,N,N',N'*-tetraacetic acid] plus 2.5 mM CaCl₂ and 5 mM MgCl₂. After DNase I was added, tubes were transferred to a 20°C bath and incubated for 60 s. The reaction was terminated by addition of 25 μ l of 50 mM EDTA–100 μ g of tRNA per ml–0.2% SDS–0.5 μ g proteinase K. After incubation at 42°C for 45 min, the DNA was phenol extracted, ethanol precipitated, and loaded on 7% polyacrylamide sequencing gels. Competition experiments were performed with double-stranded oligonucleotides containing either wild-type or mutated AP1- or E2-binding sites, which were as follows: 17-mer wild-type E2-binding site, 5'-CTAGACCGAAAACGGGTG; 15-mer mutated E2-binding site, 5'-CTAGACCGAAACGGT; wild-type AP1-binding site, 5'-TCGAGGAAGTGACTAAGT GAGCACAG; mutated AP1-binding site, 5'-TCGAGGAAG TAAGTAACTGACCGCAG.

RESULTS

Functional analysis of three subfragments from the LCR of HPV18. The *Bam*HI fragment covering the complete regulatory region of HPV18 acts as an enhancer of transcription independently of the presence of E2 (37) (Fig. 1). We were able to show that the sequence responsible for this constitutive enhancer activity is included in a central *RsaI*-*RsaI* fragment of about 230 base pairs (bp) (nt 7510 to 7738). This enhancer acts independently not only of E2 *trans*-activation, in accordance with the absence of any E2-binding sites, but also independently of any viral products, since it functions in some human cells not associated with papillomaviruses such as SW13 (35) or normal keratinocytes (B. Bernard et al., manuscript in preparation).

A long *RsaI* fragment of about 390 bp upstream of this constitutive enhancer contributes, to some extent, about 25% of the basal level to the promoter-enhancer activity

FIG. 1. Constitutive and inducible enhancers in the HPV18 LCR. Schematic representation of the different constructions containing the subfragments of the HPV18 LCR (shown at bottom) inserted 5' or 3' to the enhancerless SV40 promoter or herpes simplex virus thymidine kinase promoter in CAT expression plasmids. These plasmids were tested in cotransfection experiments in SW13, HeLa, or CV1 cells in the presence or absence of the BPV1 E2 expression plasmid (pC59). The different subfragments corresponding to the distal *RsaI*-*RsaI* fragment of 390 nt (▨), the central *RsaI*-*RsaI* fragment of 230 nt (■), the proximal *RsaI*-*Bam*HI fragment of 240 nt (□), and the distal *RsaI*-*Bam*HI fragment (▩) are indicated as well as E2-binding sites (vertical lines in the subfragments). The orientation relative to transcription from the E6 promoter is indicated by a horizontal arrow. Values at right are the mean (-fold) enhancement values obtained in two or three experiments (NT, not tested). Those labeled with asterisks are taken from reference 37. Total enhancement factor in the presence of E2 for the complete LCR is the multiple of the constitutive activity and the E2-dependent enhancer activity.

obtained with the complete region as shown by deletion analysis (35). In an attempt to demonstrate a direct enhancer activity of this fragment, we subcloned it 5' to the heterologous enhancerless SV40 early promoter. No enhancer activity could be detected with this fragment either in the presence or absence of E2, despite the presence of a perfect E2 recognition sequence in this fragment (nt 7458 to 7469). One of the questions we asked about putative protein-DNA interactions with this fragment is whether this site can in fact interact with E2 (see below).

The third fragment, proximal to the E6 ORF, has a very low promoter activity if cloned just in front of CAT as previously demonstrated by deletion experiments (35). We tested whether it could function as an enhancer either in the absence or in the presence of E2. As summarized in Fig. 1, no activation of transcription could be obtained with a construction containing this fragment 5' to the herpes simplex virus thymidine kinase promoter in either orientation. On the contrary, this construction is efficiently *trans* activated (six- to ninefold) in the presence of E2. This functional analysis shows that two distinct enhancers may coexist in the regulatory region of HPV18. One is responsible for the

high constitutive level of HPV18 transcription in certain cell lines, while the more proximal is responsible for its sensitivity to E2 *trans* activation when it is cloned in the enhancer configuration (see below).

Characterization of cellular proteins interacting with DNA sequences within the HPV18 LCR. In order to analyze cellular or viral protein-DNA interactions within the control region, we used the three fragments, described in the preceding section and shown in Fig. 1, which cover the complete LCR in HPV18.

We performed DNase I footprinting experiments, essentially as previously described (7), by incubation of asymmetrically end-labeled DNA fragments with nuclear extracts obtained from cultured cells or tissues followed by brief DNase I treatment. In order to reduce the binding of nonspecific proteins to the different labeled DNA fragments, the protein extracts were incubated with carrier double-stranded poly(dI-dC) before addition of the labeled DNA. After brief DNase I digestions, samples were treated with proteinase K, phenol extracted in order to free them from proteins, and then analyzed in polyacrylamide sequencing gels. We have prepared nuclear extracts from different cell

lines and tissues, including the human cell lines HeLa, SW13, and HepG2 and human liver tissue as well as monkey kidney CV1 and COS-1 cells.

(i) Interactions with the distal fragment. An asymmetrically end-labeled DNA fragment containing the 390-bp distal region of the HPV18 LCR was used for footprinting reactions. Figure 2 shows the footprint resulting from the interaction of human cellular factors with both strands of this fragment. There are four major protected regions in this fragment (I, II, III, and IV) found with cellular extracts from COS-1 or HeLa cells (Fig. 2A). The regions between nucleotides 7198 and 7227 (I) are clearly well protected on both strands. The two protected regions lying between nt 7260 to 7277 (II) and 7309 to 7330 (III) are slightly less prominent on the coding strand (Fig. 2A) than on the noncoding strand (Fig. 2B, lanes a and b). The last region, between nt 7418 and 7453 (IV), is strongly protected only with HeLa cell nuclear extract (Fig. 2A, lane e).

(ii) Interactions with the constitutive enhancer fragment. A DNA fragment containing the 230-nt enhancer sequences was used to carry out footprinting experiments as already described. We found two protected regions in this fragment (V and VI; Fig. 3), when incubated with human SW13 or HeLa extracts. The latter region, covering about 14 nt (between nt 7643 to 7656), is more prominent with HeLa cell extracts (Fig. 3, lane b).

Site V maps near a *Bst*NI site (nt 7600) and is protected in roughly identical positions along both strands (Fig. 3 and results not shown). Since the protected region contains the sequence 5'-TGACTAA, we wondered whether this protection could be due to the binding of AP1 transcription factor. In order to clarify this point, we performed footprinting reactions in the presence of different competitor oligonucleotides. We showed that the protection is clearly competed by an excess of unlabeled oligonucleotide containing the sequence 5'-TGACTAA-3', constituting the AP1 or its mouse homologous PEA1 site in the enhancers of polyomavirus and SV40 (27) (Fig. 3, lane e) (see Materials and Methods for a description of the oligonucleotides used).

(iii) Interactions with the promoter fragment. When we analyzed the interactions between cellular factors and the *Rsa*I-*Bam*HI promoter fragment, we found a clear protection over a region located between nt 7781 and 7803 (VII) as shown with COS-1, SW13, or liver extracts (Fig. 4). We noticed again that this protection is centered over the sequence 5'-TGACTAA-3' and wondered whether it could also be due to the presence of AP1 factor in our extracts. To explore this possibility, we carried out footprinting competition experiments with different oligonucleotides containing known binding sites for this factor (27) (see Materials and Methods).

As shown in Fig. 4B, an oligonucleotide containing the polyoma-binding site for PEA1 (AP1) competes for the binding of this factor. No competition is observed with a mutated oligonucleotide unable to bind this factor, nor with an oligonucleotide containing the E2-binding site (not shown). We verified that all oligonucleotides known to bind AP1 competed for the binding to the two TGACTAA sequences in HPV18 (not shown), thus indicating clearly that the cellular factor interacting with these sequences is AP1.

We also found a protected region (VIII) covering about 19 nt, including nt 7809 to 7827, that is clearly protected with human liver extracts (Fig. 4A, lane f). Other extracts, such as those from HeLa cells, gave the same protection, although to a lesser extent.

Identification of the sites in the HPV18 LCR interacting with

FIG. 2. DNase I footprinting of cellular and viral protein interactions within the distal 390 nt of the HPV18 LCR. (A) Coding strand. Lane d, free DNA. Lanes a to c, footprints with 32.5 μ g of the protein extract from COS-1 cells transfected with pC59 as described in Materials and Methods. Before DNase I treatment, 10 ng of unlabeled 15-mer double-stranded oligonucleotide containing a mutated E2-binding site (lane a) or 10 ng of the 17-mer double-stranded oligonucleotide of the wild-type sequence of the E2-binding site (lane b) were added. Lane e, footprint with 45 μ g of HeLa extract; lane f, sequence ladder of chemical cleavage at purines. Numbers alongside the sequence ladder indicate nt positions on the HPV18 genome (8); bars indicate protected sites. (B) Noncoding strand. Lanes: a and b, footprints obtained with transfected COS-1 cell extracts in the presence of 10 ng of unlabeled oligonucleotides containing the sequence of either the wild-type (lane a) or mutated (lane b) E2-binding site; c and d, DNA incubated without proteins; e, sequence ladder of chemical cleavages at purines. Numbers indicate nt positions in the HPV18 genome. Positions of the different protected sites (I to IV and E2) are indicated by empty bars alongside the autoradiograph.

the BPV1 E2 protein. We have previously found that the BPV1 E2 gene product represses transcription from the E6 promoter in the HPV18 LCR and that the target of this repression is contained within 240 nt upstream from this cap site (35, 37). On the other hand, as indicated in a previous section and in Fig. 1, the same fragment is able to activate

FIG. 3. DNase I footprint analysis of cellular protein interactions along the central region of the HPV18 LCR. The *RsaI-RsaI* fragment of 230 nt corresponding to the cell-specific constitutive enhancer was end labeled, and the DNA-protein interactions were analyzed on the coding strand. Lanes: a, chemical cleavage at purine residues; b, footprint in the presence of 45 μ g of protein extract from HeLa cells; c, DNA incubated with no proteins; d and e, footprints in the presence of 30 μ g of protein extract from SW13 cells and in the absence (lane d) or the presence (lane e) of 10 ng of an unlabeled oligonucleotide containing the sequence of an AP1-binding site. The difference in the intensity of cleavage around site VI between lanes d and e is due to variations in the concentrations of the different sub-bands rather than to genuine competition by the AP1 oligonucleotide. Numbers indicate nt positions in the HPV18 genome; bars indicate protected sites.

transcription in the presence of E2 if cloned upstream of a heterologous promoter like the enhancerless SV40 early promoter or the herpes simplex virus thymidine kinase promoter. It was of great interest to search for E2 interactions directly with this fragment in order to help us to understand these two opposite functions.

To this end, we used a nuclear extract of COS-1 cells transiently transfected with pC59, a cloned cDNA coding for the BPV1 E2 intact protein. Since pC59 contains the origin of replication of SV40 and COS-1 cells constitutively express the SV40 large T antigen, we expected a sufficient production of E2 by replication of the expression plasmid even in a transient assay. Comparative DNase I footprintings were done with either transfected or nontransfected COS-1 cell

FIG. 4. Analysis of cellular and viral protein-DNA binding in the proximal region of the HPV18 LCR. The *RsaI-BamHI* proximal fragment was end labeled and used to perform footprint analyses on the noncoding strand. (A) Lanes: a, footprints with 30 μ g of a protein extract from untransfected COS-1 cells; b and c, footprint with 32.5 μ g of protein extract from COS-1 cells transfected with pC59; d and g, DNA incubated with no protein; e, chemical cleavage at purines; f, footprint obtained with 30 μ g of protein from a human liver nuclear extract. Lane c was treated with less DNase than lane b was. Arrowheads indicate palindromic sequences. (B) Competition experiment with oligonucleotides containing sequences of the AP1-binding site. Lanes: a and e, DNA treated with DNase I in the absence of nuclear proteins; b, c, and d, footprints with 70 μ g of SW13 cell nuclear proteins in the absence (lane b) or the presence (lane c) of an excess of an oligonucleotide containing the sequence of the AP1-binding site or in the presence of a mutated version of this sequence (lane d). In both panels, numbers indicate nt positions in the HPV18 genome and bars indicate protected sequences.

extracts in order to identify the HPV18 DNA sequences that interact with the product of BPV1 E2 ORF. The *RsaI-BamHI* fragment from the HPV18 LCR (covering the E6 promoter) was asymmetrically end labeled on both strands and incubated with nuclear extracts. Figure 4 shows the protected regions resulting from interactions between dif-

ferent protein extracts and the HPV18 promoter noncoding strand (nt 7738 to 124). We found that the COS-1-transfected cells (COS-pC59) exhibit a marked protection corresponding to a region extending roughly from nt 37 to 77 of the viral genome, covering about 42 nt that partially overlap the TATA box (see also Fig. 5B). The protected region contains two perfect palindromes with the sequence ACCGN₄CGGT. We found no similar protection when we used nontransfected COS-1 cell extracts (Fig. 4A, lane a) or extracts from different nontransfected cell lines, including HeLa and SW13 (results not shown).

We performed competition footprints in order to conclusively establish that the interaction found with COS-pC59 extracts is due to the presence of E2 from BPV1. We used two different oligonucleotides, one containing a perfect E2-binding palindrome from the HPV18 proximal region and a second carrying a mutated site (see Materials and Methods for description of oligonucleotides). Only the perfectly matched oligonucleotide competed for the binding of E2 to the two palindromes. No competition was observed when a mutated E2 oligonucleotide or one containing the AP1-binding site was used (results not shown). When we looked for protein-DNA interactions on the coding strand, we found a similar protected region spanning roughly from nt 37 to 79 (results not shown) only in extracts prepared with COS-1-pC59-transfected cells.

Surprisingly, no protection of the imperfect palindrome ACCGN₄AGGT found around 140 bp upstream from the E6 cap site was observed with the E2-containing extract (Fig. 4A, lanes b and c). The absence of a high affinity of E2 for this site was confirmed by gel retardation assays with bacterially made E2 proteins of BPV1 or HPV18. A fragment containing this site was retarded 10- to 20-fold less efficiently than that containing the TATA proximal sites or an oligonucleotide containing a single ACCGN₄CGGT motif (N. Dostatni and F. Thierry, unpublished observations).

In contrast to the site centered around nt 7830, the E2-binding site present at nt 7458 to 7469 upstream in the LCR and containing a perfect palindrome ACCGN₄CGGT is protected by the transfected COS-1 extracts. The protected region covers about 22 nt and is centered on the perfect palindrome, as can be seen in Fig. 2A. This protection was specifically competed by an excess of the oligonucleotide containing the E2-binding site. This interaction and competition experiments are clearly depicted in Fig. 2A for the coding strand.

When we closely examined the protection of the E2 protein over its binding site, we found a very interesting pattern. We noticed a DNase I cut next to the second C of the palindrome 5'-ACCGN₄CGGT-3' on both the single and the double sites (Fig. 5). This C is not protected from nuclease attack by the binding protein, indicating that the protein contacts the helix on the other side. There is also a weak hypersensitive site at nt +35 found in the double site of the proximal promoter. The pattern found over the double perfect sites indicates that there are two symmetric protected regions, compared with the single protected region found in the perfect motif about 500 bp upstream from the cap site. These comparative data are shown in Fig. 5.

Figure 6 gives a summary of the data obtained by DNase I footprints. The protections observed with either cellular or viral proteins are indicated by horizontal lines, and induced hypersensitive sites are indicated by vertical arrows. Main regions of interaction correspond to the LCR distal region, the enhancer element, and the early promoter.

FIG. 5. Fine analysis of the interactions of the BPV1-E2 gene product with its binding sites in the HPV18 LCR. (A) Footprint obtained with nuclear extracts from COS-1 cells transfected with the E2-coding plasmid on the unique perfect E2-binding site situated about 500 nt upstream from the E6 cap site (7459 to 7470). Protected nt are shown by a striped box; stars indicate nt only partially protected by E2 against nuclease attack. Gel lanes (from top to bottom): DNase I digestion of free DNA, DNase I digestion in the presence of E2, and purine ladder. (B) Footprint obtained on the tandemly repeated sites localized about 55 nt upstream from the cap site. Protected nt are shown by a striped box; stars indicate nt which are either weakly protected (thin) or unprotected (thick). A weak hypersensitive site (position 35) is indicated by the vertical arrow. Gel lanes (from top to bottom): DNase I digestion in the presence of E2, digestion of free DNA, and purine ladder. In both panels, numbers indicate nt positions in the HPV18 genome and horizontal arrows indicate palindromic sequences.

DISCUSSION

In this study, we have analyzed protein-DNA interactions within the LCR of HPV18 by DNase I footprinting using different cell lines and tissue extracts. The HPV18 LCR consists of an 825-bp fragment (positioned 5' of the early region) which has been shown to contain *cis* regulatory elements for early transcription including the E6 promoter (15, 29, 34, 36).

Functional analysis of three subfragments from the LCR has shown that the distal fragment, relative to the E6 ORF, accounts for about 25% of the total activity of the E6 promoter. The central part contains an enhancer that is active in cell lines of epithelial origin independently of the presence of E2. The proximal region containing the E6 mRNA cap site is highly repressed in the presence of the BPV1 E2 gene product (37). Finally, the same fragment, if cloned 5' to a heterologous promoter, acts as an E2-dependent enhancer. We have used these three subfragments to study DNA-protein interactions with cellular as well as viral proteins present in nuclear extracts. We performed DNase I footprintings to precisely localize protein-binding sites. Four regions of interaction between cellular proteins and the distal part of the HPV18 LCR were characterized. These sites (I, II, III and IV) are located about 60, 130, 170, and 280 nt from the end of the L1 ORF, respectively. We are not yet able to assign these protections to known transcription factors. The interactions in the constitutive enhancer region are composed of two distinct protections (V and VI). By performing

oligonucleotide competition footprints, we found that the cellular factor interacting with domain V, about 350 bp from the E6 cap site, is a member of the AP1 family. In the proximal part of the LCR, we found at least two regions, designated VII and VIII, which interact with cellular factors. Region VII corresponds to a second AP1 site located about 175 bp upstream of the E6 cap site.

Could some of the identified cellular factors be involved in the tissue specificity of the E6 promoter? The four protein interactions that we have found in the distal part of the LCR do not seem to play a role in this specificity, since a fragment deleted to the *Bgl*II site retains not only the activity but also the tissue specificity of this promoter (35). They could represent cellular factors involved in either DNA replication or transcription from upstream promoters. Among the proteins interacting with the enhancer and the promoter proximal regions, we can exclude factors V and VII as being tissue specific, because they belong to the AP1 family, which has been shown to be almost ubiquitous. The facts that factor VIII seems to be a general protein which is very abundant in human liver as well as in rat liver (not shown) and that the E6 promoter does not function in hepatoma cell lines (35) support the idea that this factor is also not involved in conferring tissue specificity to this promoter. The only interaction we have shown that could possibly play a role in tissue specificity involves factor VI. The facts that this factor was observed in HeLa cell extracts and that this cell line is from epithelial origin could support this notion. Much work is needed in order to clarify this assumption. It is also possible that as yet unidentified factors are responsible for tissue specificity, as there is still enough space for additional interactions in the enhancer region.

The fact that there are two binding sites for AP1 or transcription factors of this family in the LCR of HPV18, one around 7610 and the other around 7790, allows us to speculate that early genes can be stimulated by mitogens or by the products of certain oncogenes that mimic their action (26b, 39). We have shown that HPV18 transcription is repressed by adenovirus E1a gene product (36). Swift et al. (34) demonstrated that the target for this repression is the viral enhancer. E1a could interact with one of the factors binding to the *Rsa*I-*Rsa*I enhancer fragment, perhaps with the HeLa factor that gives footprint number VI. However, a detailed mutational analysis of the HPV18 control region will be required before we can discern the precise functional role of the different binding sites and the corresponding proteins.

We have recently shown that the HPV18 early promoter contained in the noncoding region is the target for E2-mediated repression in cotransfection experiments (37). It has been reported that the BPV1 E2 gene product interacts specifically with the conserved sequence 5'-ACCN₆GGT-3' (1, 17, 24). This sequence, found in the LCR of sequenced papillomaviruses, seems to be the target for transactivation in BPV1, HPV11, HPV16, and cottontail rabbit papillomavirus (9, 14, 17, 32).

In this study, we report that extracts prepared from COS-1 cells transfected with an E2 expression vector containing the SV40 origin of replication generate a clear protection pattern over the HPV18 early promoter. This protection covers about 42 nt and is composed of two contiguous protected regions centered on the conserved tracts ACCGN₄CGGT, found in the HPV18 LCR about 47 and 63 bp upstream from the E6 mRNA cap site. Surprisingly, we found no interaction of E2 with the imperfect palindrome ACCGN₄AGGT, situated about 140 nt upstream from the E6 cap site, under our experimental conditions. However, we found a clear protec-

		types
ACCGTTTTCGGT.....ACCGAAAAACGGTTCACCGAAAAACGGT (-143) (-68)		6b
ACCGTTTTCGGT.....ACCGAAAAACGGTTCACCGAAAAACGGT (-144) (-68)		11
ACCGATTTCGGG.....ACCGAAATCGGTTCGAAACCGAAACGGT (-148) (-69)		16
ACCGAAATAGGT.....ACCGAAAAACGGTCGGGACCGAAAAACGGT (-140) (-63)		18
ACCGTTTTCAGGT.....ACCGAAACGGTTCACCGAAAAACGGT (-153) (-70)		33

FIG. 7. Comparison of the sequences of E2 palindromes in the proximal part of the LCR of genital papillomaviruses (virus type indicated at right).

tion over the single perfect palindrome ACCGN₄CGGT located about 500 bp upstream from the cap site. We believe that the homologous HPV18 E2 will give similar footprints, since bacterially produced BPV1 or HPV18 E2 proteins bind to the E2 target palindrome with similar affinities (N. Dostatni and F. Thierry, unpublished observations).

The fact that E2 binds efficiently to the distal single perfect palindrome could have a functional significance for the *trans* activation process by cooperatively interacting with the double palindromes. This functional cooperativity between different E2 sites has been documented (15, 16, 32; B. Bernard et al., manuscript in preparation). As with E2, *trans* activation of transcription by AP1 requires an interaction between several sites for maximal effect. Our functional data point in this direction; when we deleted one AP1 binding site, the transcriptional activity from the E6 promoter decreased drastically (35).

It is important to notice that the E2 interactions we found over the E6 promoter were obtained with a full-length E2 from BPV1 and not with the 3' product from this gene, which functions as a repressor of BPV1 transcription (18). So it seems that the full-length E2 can function as a repressor depending on the location of its target, even though it has *trans* activating properties on the same region if cloned in an enhancer configuration.

When we compared the positions of the E2-binding sites in the LCRs from sequenced human genital papillomaviruses (6, 11, 16, 18, 31), we found that in all of them there are two adjacent perfectly conserved tracts positioned between the TATA and CAAT boxes (Fig. 7). Interestingly, we found that the motifs present about 140 bp from the E6 cap site in those viruses, which are highly associated with malignancies (16, 18, 31), are not perfect palindromes; this is opposed to the fact that in the viruses associated with benign lesions (6, 11), the upstream E2 target is a perfect palindrome. The putative role of the E2 target around -140 is not clear; it has been shown that this sequence in HPV11 binds E2 as efficiently as the two downstream well-conserved targets do (17). Certainly this difference in the interaction with E2 corresponds to different putative transcriptional regulations of the two groups of genital papillomaviruses. On the contrary, the most proximal E2 targets are efficient binding sites in HPV11 (11) as well as HPV18. We would like to suggest that the binding of E2 to these sites between the CAAT- and TATA-box regions could interfere with the binding of transcription factor TFIID to the TATA box or of RNA polymerase II to the E6 promoter, with consequent inhibition of

early transcription from this promoter. Previously, we did not observe repression of the HPV18 E6 promoter with the homologous E2 in HeLa or SW13 cells. However, recent transfection experiments with primary human keratinocytes showed that the intact HPV18 E2 acts as a repressor of the E6 promoter in a homologous system. (B. Bernard et al., manuscript in preparation). It is possible that in its natural context, binding of E2 to the E6 proximal site will enhance the activity of other potential viral promoters found downstream or upstream of the E6 cap site in a manner similar to that of the enhancement of the thymidine kinase promoter that we reported here.

In carcinomas and cell lines derived from them, viral DNA is usually found integrated into the host chromosomes (2, 5, 25, 42). The fact that all human cervical carcinoma cell lines studied retain the E6 ORF indicates that expression of this sequence, probably with E7, may be required for continued maintenance of the transformed phenotype (30, 31, 41). It is intriguing that integration of HPV16 and HPV18 reproducibly results in interruption or deletion of the E1 and E2 ORFs. This can be interpreted by considering the facts that E1 is mainly implicated in episomal replication (21) and that E2 can act as a repressor of E6 gene expression. In an integrated state there would be no need for these functions, which are probably deleterious in this context; this would, in turn, leave an early gene unit transcriptionally active.

This active transcription is cell specific and requires cellular factors since it takes place in the absence of any viral functions. Among these general cellular transcription factors interacting with the HPV18 regulatory region, we were able to identify a member of the AP1 family. However, more work is needed in order to identify the other putative cellular factors and, among them, factors responsible for the cell specificity of the virus.

ACKNOWLEDGMENTS

We are grateful to S. Cereghini for continuous advice during this work. We thank F. Arnos for technical assistance, M. Blumenfeld for the gift of cell extracts and help in the preparation of some cellular extracts, J. Piette and N. Dostatni for the gift of AP1 and E2 oligonucleotides, and R. Sousa for valuable comments on the manuscript.

This work was supported by grants from the Centre Nationale de la Recherche Scientifique, the Institut National pour la Santé et la Recherche Médicale, the Association pour la Recherche sur le Cancer, the Ligue Nationale Française contre le Cancer, and the Fondation pour la Recherche Médicale Française. A.G.-C. held a fellowship from the French Ministry of Foreign Affairs.

ADDENDUM

The footprint seen between positions 7470 to 7490 only in HeLa cells (Fig. 2A) may be due to a keratinocyte-specific factor binding to the sequence TTTGGCTT present in several keratinocyte-specific genes (T. P. Cripe, T. H. Haugen, J. P. Turk, F. Tabatabai, P. G. Schmid, M. Dürst, L. Gissmann, A. Roman, and L. P. Turek, *EMBO J.* 6:3745–3753, 1987).

LITERATURE CITED

1. Androphy, E. J., D. R. Lowy, and J. T. Schiller. 1987. Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA. *Nature (London)* 325:70–73.
2. Baker, C. C., W. C. Phelps, V. Lindgren, M. J. Braun, M. A. Gonda, and P. M. Howley. 1987. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *J. Virol.* 61:962–971.
3. Beaudenon, S., D. Kremsdorf, O. Croissant, S. Jablonska, S. Wain-Hobson, and G. Orth. 1986. A novel type of human papillomavirus associated with genital neoplasias. *Nature (London)* 321:246–249.
4. Bedell, M. A., K. H. Jones, and L. A. Laimins. 1987. The E6 and E7 open reading frames of human papillomavirus type 18 are sufficient for transformation of NIH 3T3 and Rat-1 cells. *J. Virol.* 61:3635–3640.
5. Boshart, M., L. Gissmann, H. Ikenberg, A. Kleinheinz, W. Scheurlen, and H. zur Hausen. 1984. A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J.* 3:1151–1157.
6. Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* 72:248–254.
7. Cereghini, S., M. Raymondjean, A. Garcia-Carranca, P. Herbolmel, and M. Yaniv. 1987. Factors involved in control of tissue-specific expression of albumin gene. *Cell* 50:627–638.
8. Cole, S. T., and O. Danos. 1987. Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. *J. Mol. Biol.* 193:599–608.
9. Cripe, T. P., T. H. Haugen, J. P. Turk, F. Tabatabai, P. G. Schmid, M. Dürst, L. Gissmann, A. Roman, and L. P. Turek. 1987. Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer, and by viral E2 trans-activator and repressor gene products: implications for cervical carcinogenesis. *EMBO J.* 6:3745–3753.
10. Danos, O., I. Giri, F. Thierry, and M. Yaniv. 1984. Papillomavirus genomes: sequences and consequences. *J. Invest. Dermatol.* 83:7s–11s.
11. Dartmann, K., E. Schwarz, L. Gissmann, and H. zur Hausen. 1986. The nucleotide sequence and genome organization of human papillomavirus type 11. *Virology* 151:124–130.
12. Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11:1475–1489.
13. Dürst, M., L. Gissmann, H. Ikenberg, and H. zur Hausen. 1983. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc. Natl. Acad. Sci. USA* 80:3812–3815.
14. Giri, I., and M. Yaniv. 1988. Study of the E2 gene product of the cottontail rabbit papillomavirus reveals a common mechanism of transactivation among papillomaviruses. *J. Virol.* 62:1573–1581.
15. Gius, D., S. Grossman, M. A. Bedell, and L. A. Laimins. 1988. Inducible and constitutive enhancer domains in the noncoding region of human papillomavirus type 18. *J. Virol.* 62:665–672.
16. Hawley-Nelson, P., E. J. Androphy, D. R. Lowy, and J. T. Schiller. 1988. The specific DNA recognition sequence of the bovine papillomavirus E2 protein is an E2-dependent enhancer. *EMBO J.* 7:525–531.
17. Hirochika, H., R. Hirochika, T. R. Broker, and L. T. Chow. 1988. Functional mapping of the human papillomavirus type 11 transcriptional enhancer and its interaction with the trans-acting E2 proteins. *Genes Dev.* 2:54–67.
18. Lambert, P., B. Spalholz, and P. M. Howley. 1987. A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 transactivator. *Cell* 50:69–78.
19. Lee, W., A. Haslinger, M. Karin, and R. Tjian. 1987. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. *Nature (London)* 325:368–372.
20. Lorincz, A. T., W. D. Lancaster, and G. F. Temple. 1986. Cloning and characterization of DNA of a new human papillomavirus from a woman with dysplasia of the uterine cervix. *J. Virol.* 58:225–229.
21. Lusky, M., and M. R. Botchan. 1985. Genetic analysis of bovine papillomavirus type 1 *trans*-acting replication factors. *J. Virol.* 53:955–965.
22. Matlashewski, G., J. Schneider, L. Banks, N. Jones, A. Murray, and L. Crawford. 1987. Human papillomavirus type 16 DNA cooperates with activated ras in transforming primary cells.

- EMBO J. 6:1741-1746.
23. McBride, A. A., R. Shlegel, and P. M. Howley. 1988. The carboxy-terminal domain shared by the bovine papillomavirus E2 transactivator and repressor proteins contains a specific DNA binding domain. EMBO J. 6:533-539.
 24. Moskaluk, C., and D. Bastia. 1987. The E2 "gene" of bovine papillomavirus encodes an enhancer binding protein. Proc. Natl. Acad. Sci. USA 84:1215-1218.
 25. Pater, M. M., and A. Pater. 1985. Human papillomavirus types 16 and 18 sequences in carcinoma cell lines of the cervix. Virology 145:313-318.
 26. Phelps, W. C., and P. M. Howley. 1987. Transcriptional transactivation by the human papillomavirus type 16 E2 gene product. J. Virol. 60:626-634.
 - 26a. Phelps, W. C., C. L. Yee, K. Munger, and P. M. Howley. 1988. The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to adenovirus E1A. Cell 53:539-547.
 - 26b. Piette, J., S. Hirai, and M. Yaniv. 1988. Constitutive synthesis of activator protein 1 transcription factor after viral transformation of mouse fibroblasts. Proc. Natl. Acad. Sci. USA 85:3401-3405.
 27. Piette, J., and M. Yaniv. 1987. Two different factors bind to the a-domain of the polyoma virus enhancer, one of which also interacts with the SV40 and c-fos enhancers. EMBO J. 6:1331-1337.
 28. Pirisi, L., S. Yasumoto, M. Feller, J. Doniger, and J. A. Di Paolo. 1987. Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA. J. Virol. 61:1061-1066.
 29. Schneider-Gadick, A., and E. Schwarz. 1986. Different human cervical carcinoma cell lines show similar transcription patterns of human papillomavirus type 18 early genes. EMBO J. 5:2285-2292.
 30. Schwarz, E., U. K. Freese, L. Gissmann, W. Mayer, B. Roggenbuck, A. Stremlau, and H. zur Hausen. 1985. Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. Nature (London) 314:111-114.
 31. Smotkin, D., and F. O. Wettstein. 1986. Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein. Proc. Natl. Acad. Sci. USA 83:4680-4684.
 32. Spalholz, B. A., P. F. Lambert, C. L. Lee, and P. M. Howley. 1987. Localization of the E2 responsive elements of the long control region. J. Virol. 61:2128-2137.
 33. Spalholz, B. A., Y. C. Yang, and P. M. Howley. 1985. Transactivation of a bovine papillomavirus transcriptional regulatory element by the E2 gene product. Cell 42:183-191.
 34. Swift, F. V., K. Bhat, H. Youngusband, and H. Hamada. 1987. Characterization of a cell-specific enhancer found in the human papillomavirus type 18 genome. EMBO J. 6:1339-1344.
 35. Thierry, F., A. Garcia-Carranca, and M. Yaniv. 1987. Elements that control the transcription of genital papillomavirus HPV18. Cancer Cells 5:23-32.
 36. Thierry, F., J. M. Heard, K. Dartmann, and M. Yaniv. 1987. Characterization of a transcriptional promoter of human papillomavirus 18 and modulation of its expression by simian virus 40 and adenovirus early antigens. J. Virol. 61:134-142.
 37. Thierry, F., and M. Yaniv. 1987. The BPV1-E2 trans-acting protein can be either an activator or a repressor of the HPV18 regulatory region. EMBO J. 6:3391-3397.
 38. Tsunokawa, Y., N. Takebe, T. Kasamatsu, M. Terada, and T. Sugimura. 1986. Transforming activity of human papillomavirus type 16 DNA sequences in a cervical cancer. Proc. Natl. Acad. Sci. USA 83:2200-2203.
 39. Wasylyk, C., J. L. Imler, J. Perez-Motul, and B. Wasylyk. 1987. The c-Ha-ras oncogene and a tumor promoter activate the polyoma virus enhancer. Cell 48:525-534.
 40. Wildeman, A. G., P. Sassone-Corsi, T. Grundstrom, M. Zenke, and P. Chambon. 1984. Stimulation of *in vitro* transcription from the SV40 early promoter by the enhancer involves a specific trans-acting factor. EMBO J. 3:3129-3133.
 41. Yasumoto, S., A. L. Burkhard, J. Doniger, and J. A. Di Paolo. 1986. Human papillomavirus type 16 DNA induced malignant transformation of NIH 3T3 cells. J. Virol. 57:572-577.
 42. Yee, C., I. Krishnan-Hewlett, C. C. Baker, R. Schlegel, and P. M. Howley. 1985. Presence and expression of human papillomavirus sequences in human cervical carcinoma cell lines. Am. J. Pathol. 119:361-366.
 43. zur Hausen, H., and A. Schneider. 1987. The role of papillomaviruses in human anogenital cancer, p. 245-263. In N. P. Salzman and P. M. Howley (ed.), The papovaviridae, vol. 2. Plenum Publishing Corp., New York.