

Nuclear Matrix Attachment Regions of Human Papillomavirus Type 16 Repress or Activate the E6 Promoter, Depending on the Physical State of the Viral DNA

WALTER STÜNKEL, ZHONGHUI HUANG, SHYH-HAN TAN,[†] MARK J. O'CONNOR,[‡]
AND HANS-ULRICH BERNARD*

Institute of Molecular and Cell Biology, National University of Singapore, Singapore 117609, Republic of Singapore

Received 1 October 1999/Accepted 15 December 1999

Two nuclear matrix attachment regions (MARs) bracket a 550-bp segment of the long control region (LCR) containing the epithelial cell-specific enhancer and the E6 promoter of human papillomavirus type 16 (HPV-16). One of these MARs is located in the 5' third of the LCR (5'-LCR-MAR); the other lies within the E6 gene (E6-MAR). To study their function, we linked these MARs in various natural or artificial permutations to a chimeric gene consisting of the HPV-16 enhancer-promoter segment and a reporter gene. In transient transfections of HeLa cells, the presence of either of these two MARs strongly represses reporter gene expression. In contrast to this, but similar to the published behavior of cellular MARs, reporter gene expression is stimulated strongly by the E6-MAR and moderately by the 5'-LCR-MAR in stable transfectants of HeLa or C33A cells. To search for binding sites of soluble nuclear proteins which may be responsible for repression during transient transfections, we performed electrophoretic mobility shift assays (EMSAs) of overlapping oligonucleotides that represented all sequences of these two MARs. Both MARs contain multiple sites for two strongly binding proteins and weak binding sites for additional factors. The strongest complex, with at least five binding sites in each MAR, is generated by the CCAAT displacement factor (CDP)/Cut, as judged by biochemical purification, by EMSAs with competing oligonucleotides and with anti-CDP/Cut oligonucleotides, and by mutations. CDP/Cut, a repressor that is down-regulated during differentiation, apparently represses HPV-16 transcription in undifferentiated epithelial cells and in HeLa cells, which are rich in CDP/Cut. In analogy to poorly understood mechanisms acting on cellular MARs, activation after physical linkage to chromosomal DNA may result from competition between the nuclear matrix and CDP/Cut. Our observations show that *cis*-responsive elements that regulate the HPV-16 E6 promoter are tightly clustered over at least 1.3 kb and occur throughout the E6 gene. HPV-16 MARs are context dependent transcriptional enhancers, and activated expression of HPV-16 oncogenes dependent on chromosomal integration may positively select tumorigenic cells during the multistep etiology of cervical cancer.

Human papillomavirus type 16 (HPV-16) and several related HPV types cause cancer as a consequence of persistent infection of epithelial cells of the transformation zone of the uterine cervix. Three oncoproteins, products of the genes E5, E6, and E7, deregulate the cell cycle and intracellular signaling, induce immortalization, and increase mutation rates by forming complexes with cellular factors which include p53, E6-AP, E6-BP, Rb, and paxillin (for reviews, see references 30, 53 and 70, for recent references; 40). The expression of E6 and E7 is controlled by a promoter upstream of the E6 gene, which is called P97 in the case of HPV-16. It is generally believed that the strength of transcription from P97 will determine the concentration of the oncoproteins E6 and E7 *in situ* and that increases or decreases in the concentration of these proteins will favor or disfavor the carcinogenic process.

The activity of P97 and of the homologous E6 promoters of other genital HPV types is under the control of about a dozen different transcription factors (for a review, see reference 46)

and the chromatin structure surrounding P97 (61). These various factors modulate P97 in correlation with the identity of the infected epithelial cell (13, 15, 26, 27), the differentiation state of this cell (1, 22, 50), the physiology of the host (12, 51), and viral feedback loops (17, 63). An example of how enhanced expression of E6 and E7 may favor cellular transformation is the transcriptional stimulation of P97 by progesterone (12), which leads to increased transformation in cell culture (51) and may be the molecular mechanism underlying the epidemiological observation that the number of parturitions a woman has undergone positively correlates with the likelihood of developing cervical cancer (56). Another example is that P97 is repressed by a negative feedback loop exerted by the viral E2 protein. In tumor cells, the HPV genomes have frequently recombined with cellular DNA downstream of the E7 gene. As such a rearrangement leads to a shutoff of E2 expression, disruption of this negative feedback is considered to be important for progression of neoplasia (58). Here, we report that two nuclear matrix attachment regions (MARs) also influence P97 activity in a manner relevant for carcinogenesis.

The cellular nucleus contains, in addition to the chromatin, a variety of substructures, which establish spatially defined compartments with specialized functions (8, 36, 60). One of these structural elements has been described as the nuclear matrix based on biochemical fractionation (7, 8), and it is apparently identical to a ribonucleoprotein containing network

* Corresponding author. Mailing address: Institute of Molecular and Cell Biology, National University of Singapore, 30 Medical Dr., Singapore 117609, Republic of Singapore. Phone: 65-778-8823. Fax: 65-779-1117. E-mail: mcbhub@imcb.nus.edu.sg.

[†] Present address: National Institutes of Health, Unit of Cell Cycle Regulation, Bethesda, MD 20892-5431.

[‡] Present address: KuDOS Pharmaceuticals Ltd., Cambridge CB4 4GW, United Kingdom.

of fibrils and granules identified by electron microscopy (45). One of the functions of the nuclear matrix is to serve as an anchor for the attachment of chromatin loops, and such attachments are thought to isolate a loop-internal gene from surrounding genetic elements. DNA segments with a high affinity for the nuclear matrix are called MARs. MARs have functions beyond the anchoring of chromatin loops, as they often occur close to transcriptional enhancers and promoters. Most enzymatic machineries that handle DNA and RNA associate with insoluble nuclear structures (see references 33 and 67 and references therein), and MARs seem to bring together *cis*-responsive elements, the nuclear matrix, and its attached enzymatic machineries, topological changes that eventually would result in transcriptional modulation.

HPV-16 P97 activity is modulated by transcriptional activators and repressors that bind a DNA segment with a size of about 550 bp, roughly between genomic positions 7450 and 97. This DNA segment is flanked by two MARs, one in the 5' third of the LCR, approximately from positions 7150 to 7450. The second MAR overlaps with the coding sequence of the E6 gene (62). Our research addresses the question of whether these MARs affect the function of the enhancer-promoter segment. In studies of the function of cellular MARs, it has been observed that enhancers and promoters linked to MARs can be transcriptionally repressed in transient transfections, where DNA mostly occurs in episomal form, but stimulated in stable clones, where the same constructs have become integrated into the chromosomal DNA of the recipient cell (9, 10, 14, 34, 57). We report similar regulatory phenomena for HPV-16 MARs, have begun to unravel the underlying molecular mechanisms, and discuss the implications of this regulation for cervical carcinogenesis.

MATERIALS AND METHODS

Plasmid constructs. For all cloning procedures we used *Escherichia coli* JM109 (*recA*). All HPV-16 subclones were derived from the HPV-16 reference genome, subcloned into the *Bam*HI site of pSP65 (Promega). Contiguous segments of the HPV-16 genome containing either the epithelial cell-specific enhancer plus promoter (positions 7450 to 100, EP segments) or the same segment plus the MAR in the 5' part of the LCR (5'-LCR-MAR) (positions 7150 to 100, L-EP segments) were amplified by PCR, subcloned into the *Srf*I site of pCR-ScriptT-SK(+) (Stratagene, San Diego, Calif.), and recloned in the form of *Kpn*I-*Sac*I fragments into the luciferase expression vector pGL3 basic (Promega). Additional MARs of HPV-16 were cloned in the form of *Bam*HI-*Sal*I fragments into the pGL3 basic derivatives downstream of the luciferase gene. This second MAR was either the E5 gene including the early late intergenic region (E5-MAR, genomic positions 3536 to 4337), the E6 gene (E6-MAR, positions 105 to 560), or the 5'-LCR (long control region)-MAR (positions 7150 to 7450). For generation of the construct E6EP-Luc-E6, the 5'-LCR-MAR was cut out with *Eco*RI and replaced by the E6 gene in the form of a PCR-generated *Eco*RI fragment. All constructs are schematically represented in Fig. 1A. Three deletion mutants of the E6-MAR (L-EP-Luc-delmut1 to 3) were generated by PCR (Fig. 1B). Descriptions of CCAAT displacement protein the CDP expression vector pMT2-CDP and the parental vector pMT2 (44) have been published elsewhere.

Generation of stable cell lines. For analyses of MAR functions, we transfected HeLa and C33A cells. Both cell lines are derived from cervical cancers; HeLa contains chromosome internal copies of HPV-18 genomes, while C33A is free of HPV genomes. All cells were grown under standard conditions in Dulbecco's modified Eagle medium with 10% fetal calf serum. Plasmid pXJ418, which confers resistance against the neomycin homologue G418, served as a selection marker for stably transfected cells.

Stable transfectants were generated by electroporation of HeLa or C33A cells with 10 μ g of *Xmn*I-linearized HPV-16-luciferase test vectors mixed with 1 μ g of pXJ418, linearized with *Bam*HI. At 48 h after transfection, the cells were split and selection was started by adding G418 (500 μ g/ml) to the medium. The resistant colonies were pooled after 2 weeks of selection and subjected to analyses of the copy number and expression level of the reporter gene. To investigate stably transfected clones, individual G418-resistant colonies were picked and expanded. The copy number of the reporter gene was determined by Southern blotting after preparation of genomic DNA according to standard protocols. Five micrograms of genomic DNA was blotted onto a Hybond N membrane by slot blotting. The blot was dried at 80°C for 2 h and hybridized against a ³²P-random-primed luciferase gene.

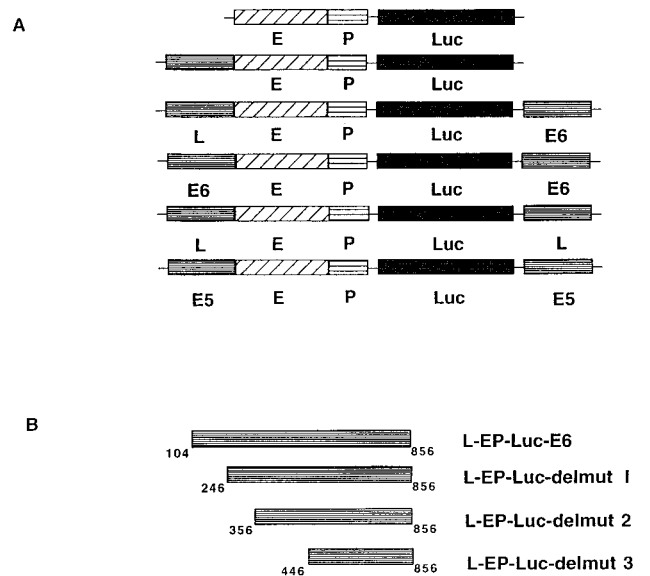


FIG. 1. Schematic structures of test vectors to examine transcription regulatory functions of three different MARs of HPV-16 (A) and deletions of the E6-MAR (B). The recombinant reporter gene (uppermost scheme) consisted of a contiguous genomic segment of the HPV-16 LCR with the epithelial cell-specific enhancer (E) and the E6 promoter P97 (P) fused to the luciferase gene (Luc). This reporter gene was linked in natural or artificial sequence to the MARs in the 5' part of the LCR (L), in the E6 gene (E6), or overlapping with the E5 gene and the early-late intergenic region (E5). Details of the constructs are described in Materials and Methods.

Transient transfection studies and luciferase assays. HeLa cells were transfected using the Lipofectamine reagent (Gibco-BRL) according to the instructions of the manufacturer. Cells (2×10^5) were seeded into six-well plates in a volume of 2 ml of Dulbecco's modified Eagle medium DMEM; 5 to 10 μ g of plasmid DNA was mixed with 15 μ l of Lipofectamine reagent and adjusted to a total volume of 200 μ l with serum-free medium. After complexes had formed during a 30-min incubation, 800 μ l of serum-free medium was added, and the samples were added to the cells. Eight hours later, the mixture was removed; after a wash with phosphate-buffered saline (PBS), new medium containing 10% fetal calf serum was added. The cells were harvested after 48 h by removal of the medium, washing with PBS, and lysis in 200 μ l of 1 \times cell lysis buffer (Promega). The plates were rocked for 20 min at room temperature, and the lysates were transferred to 1.5-ml Eppendorf tubes. The activity of the firefly luciferase was measured using a Turner TD-20/20 luminometer as instructed by the manufacturer (Promega); 10 μ l of cell lysate was mixed with 50 μ l of luciferase assay substrate (Promega), and luciferase activity was recorded as the mean of three independent transfections. As an internal standard, the cells were cotransfected with a plasmid encoding β -galactosidase, and luciferase activity was standardized against the activity of this enzyme.

Oligonucleotides and EMSAs. The sequences of all synthetic oligonucleotides are presented in Table 1. For use in electrophoretic mobility shift assay (EMSA), approximately 50 ng of each annealed oligonucleotide of the 5'-LCR-MAR was labeled with [α -³²P]dATP and [α -³²P]dCTP with Klenow polymerase (Boehringer), while E6-MAR oligonucleotides were labeled with [γ -³²P]ATP and polynucleotide kinase; 20,000 cpm of the labeled oligonucleotides was used in an EMSA in the presence of 1 μ g of poly(dI-dC) (Boehringer) (49). Samples were loaded onto 4% polyacrylamide gels and run in 0.25 \times Tris-borate-EDTA buffer for 2 h at 150 V; the gels were transferred onto Whatman blotting paper, dried under vacuum, and autoradiographed overnight at -80°C. For competitions, a 100-fold excess of either specific or unspecific nonradioactive oligonucleotide was incubated with the protein fraction prior to the addition of the labeled oligonucleotide.

Purification of proteins by column chromatography. To identify any factor(s) binding to oligonucleotide 10 (5'-AATACAACAAACCGTTGTGTGATTGTTAATTAGGTGAT-3') of the E6 gene (E6-10), HeLa nuclear extracts were separated by chromatographic methods, and activities were monitored by EMSA. Nuclear extracts were either prepared from HeLa cells (21) or purchased from the Computer Cell Culture Center (Brussels, Belgium). Ten milliliters of nuclear extract (protein content of 150 mg) was applied to a heparin-Sepharose column (Pharmacia), previously equilibrated with 5 volumes of buffer D (0.1 M KCl, 20 mM HEPES [pH 7.9], 20% [vol/vol] glycerol, 0.2 mM EDTA, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride). Subsequently, the column

TABLE 1. Overlapping oligonucleotides that completely represent the 5'-LCR-MAR and the E6-MAR^a

Name	Genomic position	Nucleotide sequence
LCR-1	7132-7162	5'-gatccTAAACGCAAAAAACGTAAGCTGTAAGTATTGga-3'
LCR-2	7147-7177	5'-gatccTAAGCTGTAAGTATTGTATGTATGTTGAATTa-3'
LCR-3	7163-7193	5'-gatccTATGTATGTTGAATTAGTGTGTTTGTGTGa-3'
LCR-4	7178-7208	5'-gatccAGTGTGTTTGTGTTGTATATGTTTGTATGTa-3'
LCR-5	7194-7224	5'-gatccTATATGTTTGTATGTGCTTGTATGTGCTTGTa-3'
LCR-6	7209-7239	5'-gatccGCTTGTATGTGCTTGTAAATATTAAGTTGTa-3'
LCR-7	7225-7255	5'-gatccAAATATTAAGTTGTATGTGTGTTTGTATGTa-3'
LCR-8	7240-7270	5'-gatccTGTGTGTTTGTATGTATGGTATAATAAACACa-3'
LCR-9	7256-7286	5'-gatccTGGTATAATAAACACAGTGTGTATGTTTGTa-3'
LCR-10	7271-7301	5'-gatccGTGTGTATGTGTTTTTAAATGCTTGTGTAACa-3'
LCR-11	7287-7317	5'-gatccAAATGCTTGTGTAACCTATTGTGTCATGCAACa-3'
LCR-12	7302-7332	5'-gatccTATTGTGTGTCATGCAACATAAAATAAACTTATa-3'
LCR-13	7318-7348	5'-gatccATAAATAAACTTATTGTTTCAACACCTACTAa-3'
LCR-14	7333-7363	5'-gatccGTTTCAACACCTACTAATTGTGTTTGTGGTTa-3'
LCR-15	7349-7379	5'-gatccATTGTGTTGTGGTTATTCATTGTATATAAAACa-3'
LCR-16	7364-7394	5'-gatccTTCATTGTATATAAACTATATTTGCTACATCa-3'
LCR-17	7380-7410	5'-gatccTATATTTGCTACATCCTGTTTTTGTATTTa-3'
LCR-18	7395-7425	5'-gatccCTGTTTTTGTGTTTTATATATACTAATTTTGTa-3'
LCR-19	7411-7441	5'-gatccTATACTATATTTTGTAGCGCCAGCGGCCATTa-3'
LCR-20	7426-7456	5'-gatccAGCGCCAGCGGCCATTTTGTAGCTCAACCGa-3'
LCR-21	7442-7472	5'-gatccTTGTAGCTCAACCGAATTCGGTTGCATGCTa-3'
E6-1	104-143	5'-ATGTTTCAGGACCCACAGGAGCGAACCGAAAAGTTACCAC-3'
E6-2	134-173	5'-AAGTTACCACAGTTATGCACAGAGCTGCAAACAACCTATAC-3'
E6-3	164-203	5'-ACAACATAACATGATATAATATTAGAATGTGTGTACTGCA-3'
E6-4	194-233	5'-GTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATG-3'
E6-5	224-263	5'-AGGTATATGACTTGTGCTTTTCGGGATTTATGCATAGTATA-3'
E6-6	254-293	5'-GCATAGTATATAGAGATGGGAATCCATATGCTGTATGTGA-3'
E6-7	284-313	5'-CTGTATGTGATAAATGTTTAAAGTTTTATTCTAAAATTAG-3'
E6-8	314-343	5'-CTAAAATTAGTGAGTATAGACATTATTGTTATAGTTTGTa-3'
E6-9	344-373	5'-ATAGTTTGTATGGAACAACATTAGAACAACATAACAACAA-3'
E6-10	374-403	5'-AATACAACAAACCGTTGTGTGATTTGTTAATTAGGTGTAT-3'
E6-11	404-433	5'-TTAGGTGTATTAACCTGTCAAAAGCCACTGTGTCTCTGAAGA-3'
E6-12	434-463	5'-GTCCTGAAGAAAAGCAAAGACATCTGGACAAAAGCAAAG-3'
E6-13	464-493	5'-AAAAGCAAAGATTCCATAATATAAGGGTCCGGTGACCGG-3'
E6-14	494-523	5'-GGTGGACCGGTGCGATGTATGTCTTGTGTCAGATCATCAAG-3'
E6-15	524-553	5'-GATCATCAAGAACACGTAGAGAAACCCAGCTGTAATCATG-3'

^a These oligonucleotides span genomic positions 7132 to 7472 and 104 to 553, as indicated. The 5'-LCR-MAR is represented by 31-mers that overlap by 15 nucleotides, and the E6-MAR is represented by 40-mers with a 10-nucleotide overlap. In the case of the 5'-LCR-MAR, the nucleotides represented by lowercase letters were added to create artificial restriction sites for experiments unrelated to this study. For this research, they were used for radioactive labeling by the alpha deoxynucleotides and Klenow polymerase. Double-stranded E6 oligonucleotides were blunt ended and labeled with gamma deoxynucleotides and polynucleotide kinase.

was washed with 2 volumes of buffer D. Column-bound proteins were eluted with stepwise increases of KCl concentrations in buffer D. Band shift activity was found in the 0.3 M fraction. After dialysis of active fractions against buffer D, we further fractionated the preparation by ammonium sulfate precipitation. Activity was found in the 40% ammonium sulfate precipitate, as determined after overnight dialysis against buffer D. Active fractions (200 µl) were subjected to gel filtration (Sephacryl S400; Pharmacia) in the presence of buffer D, where the activity eluted near the void volume of the column with an apparent molecular mass of 200 kDa.

RESULTS

Three MARs of HPV-16 suppress transcription from the HPV-16 enhancer-promoter in transient transfections. Two short segments of the HPV-16 genome have strong affinity to the nuclear matrix: the 5' third of the LCR and the E5 gene together with the early-late intergenic region downstream of E5. A third segment, with moderately strong affinity, overlaps with the E6 gene. Computer-based sequence analysis makes it likely that MARs in similar positions are conserved in many or even all genital HPVs. This suggests that they play an essential role during the life cycle of genital HPVs (62). In this paper, we refer to these three nuclear MARs as 5'-LCR-MAR, E6-MAR, and E5-MAR; in the names of recombinant plasmids, we further abbreviate them as L, E6, and E5. Some other portions of the HPV-16 genome have low affinity to the nu-

clear matrix, and we have not included them in this study. MARs are typically identified through attachment to nuclear matrix preparations in vitro and in vivo, but investigations going beyond binding studies have measured functional correlates, e.g., influences on the transcription of linked genes. Several studies have reported diverging effects on the transcription during transient and stable transfections (9, 10, 14, 34, 57).

First, we addressed the question of whether the MARs of HPV-16 may influence the transcription by the HPV-16 enhancer and E6 promoter P97 in transient transfection studies. We compared the expression of the luciferase reporter gene directed by the HPV-16 enhancer and promoter with the expression from constructs containing in addition to the enhancer and promoter either one or two HPV-16 MARs, positioned upstream or downstream or on both ends of the enhancer-promoter-luciferase (EP-Luc) segment (e.g., we refer to a construct with the 5'-LCR-MAR upstream and the E6-MAR downstream of this chimeric reporter as L-EP-Luc-E6). Some of these constructs reflect the natural relative position of these MARs, while others are artificially permuted. All constructs are listed in Fig. 1A; deletion mutants (see below) are depicted in Fig. 1B.

Figure 2A shows luciferase expression after transient trans-

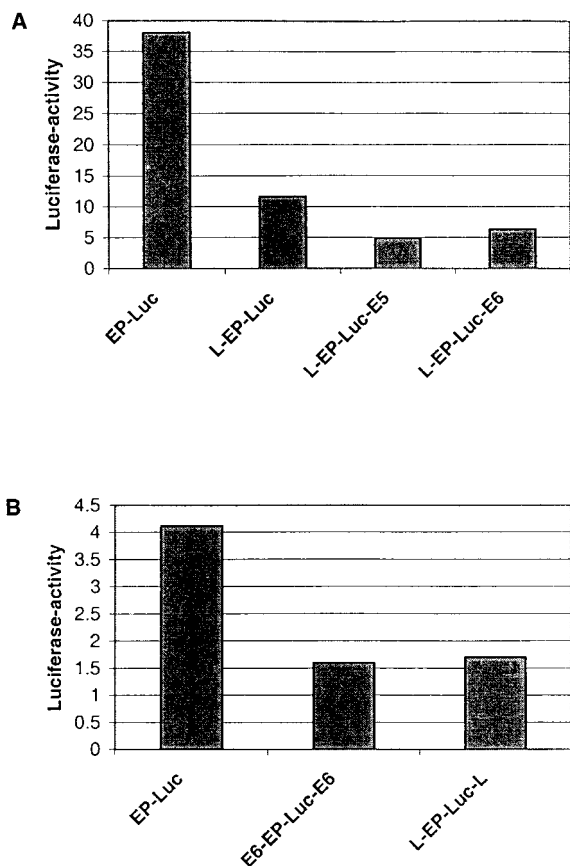


FIG. 2. MARs of HPV-16 in the 5' LCR, the E6 gene, and the E5 gene repress transcription from the HPV-16 enhancer-promoter in transient transfections of HeLa cells. The presence of the MAR in the 5' part of the LCR in its natural position relative to the EP segments (L-EP-Luc) represses luciferase reporter gene expression by 68%, while various artificial constructs reduce enhancer-promoter activity by 60% (L-EP-Luc-L) to 87% (L-EP-Luc-E5). Schematic structures of the test vectors are shown in Fig. 1.

fection of six of these recombinant clones into HeLa cells. With L-EP-Luc, a vector that had the 5'-LCR-MAR positioned upstream of the enhancer, the natural organization in HPV-16, we observed expression 3.5-fold lower than that generated by the MAR-free vector EP-Luc. Complementation of this clone with E6-MAR or E5-MAR downstream of the luciferase gene further suppressed luciferase expression. Suppression was also profound when, in an unnatural alignment, two identical MARs were present upstream and downstream of the EP-Luc segment (Fig. 2B). We conclude that MARs act in transient transfections on the HPV-16 enhancer-promoter as *cis*-responsive repressors of transcription.

Strong and moderate stimulation of transcription from the HPV-16 enhancer-promoter in stable transfectants by the E6-MAR and 5'-LCR-MAR, respectively. It is difficult to examine details of the molecular organization of transiently transfected DNA. While during such a short experiment most of the transfected DNA will exist episomally rather than recombined with host chromosomal DNA, it can only be speculated whether, in the case of a virus, such an episomal state is mimicking the natural episomal maintenance and allows for the natural transcription biology. On the other hand, in stably transfected clones, the transfected DNA has become integrated into the host's chromosomal DNA, just as it occurs with HPV DNA in the majority of malignant cervical lesions. As MARs are known

to modulate linked enhancers and promoters after chromosomal recombination in a manner differing from that during transient transfection, we studied HPV-16 MAR-EP-Luc recombinants in stably transfected HeLa and C33A cells.

C33A cells are derived from a cervical cancer but do not carry endogenous HPV genomic copies, while HeLa cells carry endogenous HPV-18 genomes. Both cells were investigated to exclude potential influences of integrated HPV genomes or its gene products on transfected vectors. To obtain cell lines harboring stably integrated copies of the EP-Luc constructs with and without MARs, we transfected these cells by electroporation with linearized DNA in the presence of a G418 selectable vector. Linear DNA was chosen to ensure integration through the DNA termini without disruption of the fusion genes during the process of integration. G418-resistant colonies were harvested, pooled, and prepared for luciferase assays. Dot blots confirmed that different pools of transfectants contained similar numbers of genomes (e.g., top of Fig. 3). Because of this, and because of the randomness of the integration in pools of transfectants, differences in luciferase expression must originate from the nature of the recombinant molecule rather than from a gene dosage effect.

Compared with transient transfections, luciferase activity of the basic construct was found low due to low copy numbers (EP-Luc constructs in Fig. 3A to D). Complementation of EP-Luc with the 5'-LCR-MAR or the E5-MAR led to an approximately threefold stimulation in C33A cells (Fig. 3C) but had no strong effect in HeLa cells. There was no alteration by supplying two copies of the 5'-LCR-MAR or the E5-MAR on either side of the reporter construct. In contrast, constructs having the E6-MAR downstream or on either side of the EP-Luc segment were stimulated by about 2 orders of magnitude in HeLa cells and by a factor of 5 to 10 in C33A cells. This activity does not require the complete E6 gene but only sequence elements 3' of position 246, as a deletion mutant of sequences between positions 104 and 246 leads to a small increase, but further deletion to position 356 leads to a complete loss of function (Fig. 1 and 3E). These observations identify the E6-MAR as strong and the 5'-LCR-MAR and E5-MAR as weak *cis*-responsive activators whose functions depend on the physical organization of the DNA as part of cellular chromosomes. This observation was intriguing, as we had not anticipated a regulatory element within an HPV gene, nor did the moderate nuclear matrix affinity of the E6-MAR, in contrast to the strong affinity of the other two MARs, suggest the possibility of such a function.

Stimulation of reporter gene expression depends on the E6-MAR but not on the site of chromosomal integration. The recombination of transfected vectors normally occurs at random sites throughout various chromosomes. In the previous experiments, we had used pools of stable transfectants to average out the effect that the site of chromosomal integration may have on expression of the reporter gene. To further study the contributions of the constructs and of the site of insertion, we picked individual G418-resistant clones from a similar transfection of HeLa cells and determined the luciferase activity for each of these clones separately (Fig. 4). Four clones which contained EP-Luc showed barely detectable luciferase activity. Four additional clones with the 5'-LCR-MAR in the natural position (L-EP-Luc) had luciferase activities significantly above background. Eight clones, having the 5'-LCR-MAR upstream and the E6-MAR downstream of the E6 gene, had activities much greater than those of the EP-Luc clones and even 10- to 20-fold above those of the L-EP-Luc clones, although they differed by up to a factor 3 among one another. We conclude that the increased activity of the individual clones

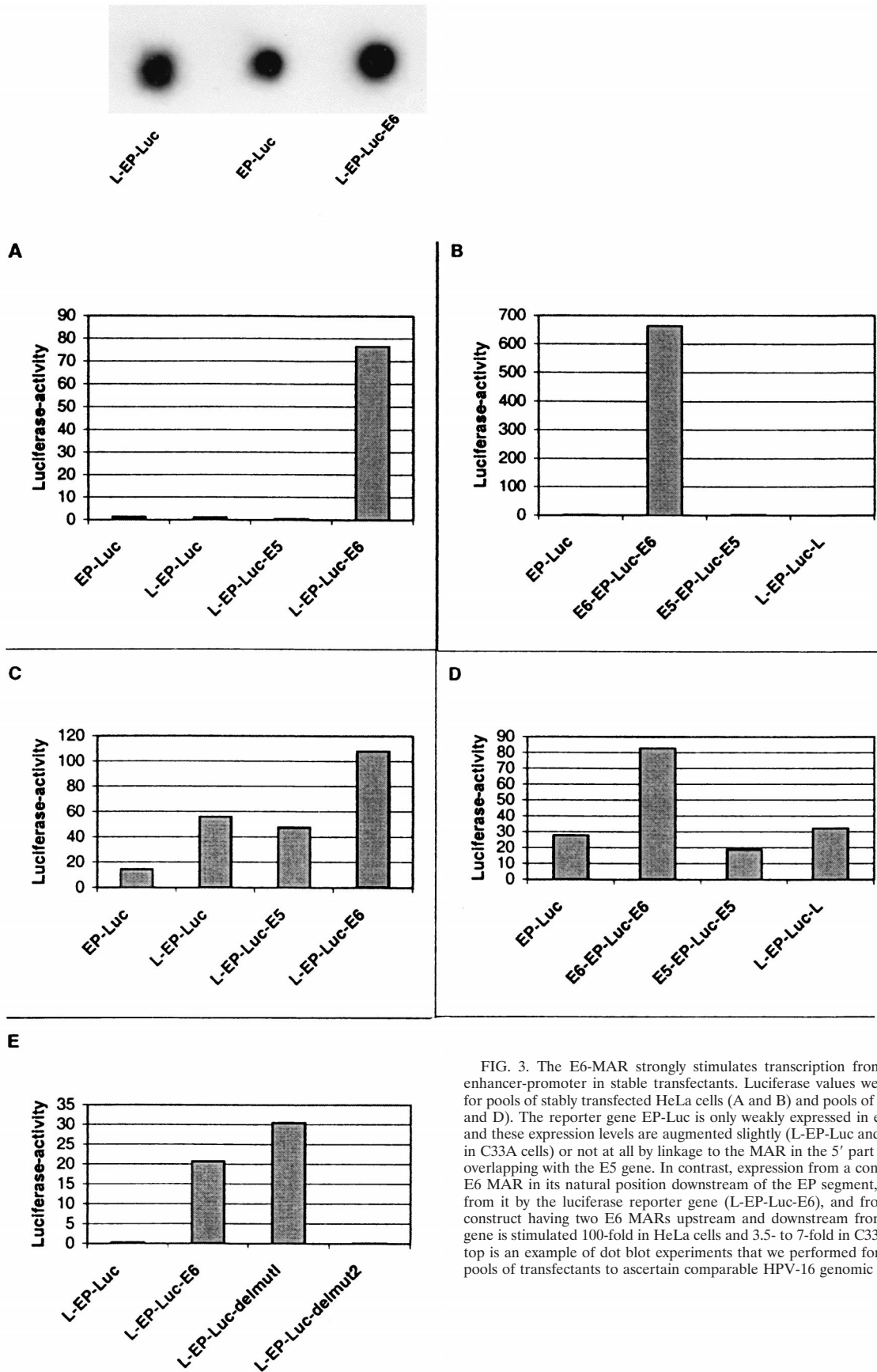


FIG. 3. The E6-MAR strongly stimulates transcription from the HPV-16 enhancer-promoter in stable transfectants. Luciferase values were determined for pools of stably transfected HeLa cells (A and B) and pools of C33A cells (C and D). The reporter gene EP-Luc is only weakly expressed in either cell line, and these expression levels are augmented slightly (L-EP-Luc and L-EP-Luc-E5 in C33A cells) or not at all by linkage to the MAR in the 5' part of the LCR or overlapping with the E5 gene. In contrast, expression from a construct with the E6 MAR in its natural position downstream of the EP segment, but separated from it by the luciferase reporter gene (L-EP-Luc-E6), and from an artificial construct having two E6 MARs upstream and downstream from the reporter gene is stimulated 100-fold in HeLa cells and 3.5- to 7-fold in C33A cells. At the top is an example of dot blot experiments that we performed for each of these pools of transfectants to ascertain comparable HPV-16 genomic copy numbers.

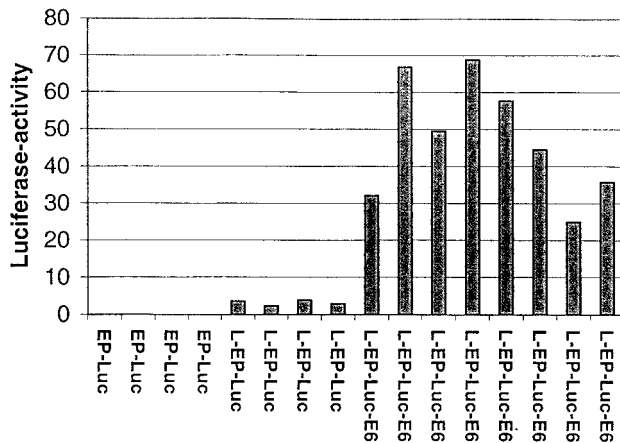


FIG. 4. Transcription of the luciferase reporter gene in individual stable transfectants is slightly stimulated by the MAR in the 5' LCR and strongly by the E6 MAR and does not significantly depend on the site of chromosomal integration. Four individually isolated stable transfectants of EP-Luc into HeLa cells showed nearly undetectable levels of luciferase expression, which is significantly stimulated by the presence of the MAR in the 5' LCR (L-EP-Luc) in four individual clones. Fusion of these constructs with the E6 MAR leads to a further 10- to 20-fold stimulation of luciferase expression in eight independent stable transfectants (L-EP-Luc-E6).

as well as of the pools of clones originates mostly from the MARs and is little influenced by the chromosomal environment. These experiments further identify the E6-MAR as strong and the 5'-LCR-MAR as weak physical context-dependent transcriptional enhancers.

An EMSA screen of oligonucleotides representing the 5'-LCR-MAR and the E6-MAR reveals numerous binding sites for soluble nuclear proteins. MARs are typically AT-rich regions with a high propensity for stable base unpairing under superhelical strain. They are bound by numerous proteins with little sequence specificity that are intrinsic components of the nuclear matrix, such as topoisomerase II, nucleolin, lamins, SAF-A, and even cell-type-specific factors such as SATB1 (19). These proteins are candidates for being responsible for the activation function of MARs. Unfortunately, it is technically difficult to map the interactions between MARs and these intrinsic matrix proteins.

There are only few reports about intrinsic matrix proteins with the properties of typical transcription factors, such as the B-cell-specific factor Bright (66). On the other hand, it has been noted that many sequence-specific transcription factors (for a review, see reference 11) as well as components of the general transcription machinery, including the RNA polymerase II holoenzyme (33), have high affinity to the nuclear matrix. By binding to their DNA target sequences, such proteins may generate some affinity of this DNA sequence to the nuclear matrix. Here, we decided to investigate whether the HPV-16 MARs can bind any of these soluble and sequence-specific factors.

Toward this end, we dissected the 5'-LCR-MAR and the E6-MAR into short segments, represented by overlapping oligonucleotides, and used these oligonucleotides in EMSAs with soluble proteins from nuclear extracts (21). Table 1 shows oligonucleotides with lengths of 31 bp representing the complete 5'-LCR-MAR, from the L1 gene on the 5' side to the epithelial specific enhancer on the 3' side (LCR-1 to LCR-21), and oligonucleotides with lengths of 40 bp representing the complete E6 gene (EP-1 to EP-15). Figure 5 shows the outcome of an EMSA analysis of these oligonucleotides with

HeLa nuclear extracts. As can be seen, there are at least seven bands, labeled A to F, which are apparently derived from specific protein-DNA interactions, as they appear only with some but not with other oligonucleotides.

We decided to study in detail complex C, as it occurs at least five times with high affinity in each of the two MARs (LCR-2, LCR-6 to LCR-9, and LCR-12 to LCR-15; E6-3, -5, -7, -10, and -13), and as these complexes include some of the strongest signals of this experiment. The weak formation of complex C with some additional oligonucleotides may stem from the high AT content of the HPV-16 genome, which appears to favor the formation of this complex (see below). To identify complex C, we used in this figure the terms C1 and C2 (the latter term designating the weak trailing band visible in some slots), as it became clear later during this research that this complex is identical to one that we detected binding HPV-16 sequences elsewhere (47, 48) (see below). Two complexes, A and B, form on one oligonucleotide each. In experiments not shown here (reference 49 and data not shown), we identified A and B as being derived from the binding of the transcription factors USF and YY1, respectively. Yet four other, often weak complexes, D to G, also occur with oligonucleotides derived from either MAR and multiple times in each MAR. The nature of these complexes is not known. Complex C also formed with some oligonucleotides representing the E5-MAR but not strongly with oligonucleotides representing the epithelial cell-specific enhancer (data not shown).

Biochemical purification, EMSA competition, and supershift analyses identify a principal factor binding to both HPV-16 MARs as CDP/Cut. To identify the nature of the factor in complex C1, we used the E6-10 as a reference for all 10 binding sites of the C1 factor in both MARs. HeLa nuclear extract was loaded onto a heparin-Sepharose column, and fractions with EMSA activity on E6-10 were eluted at a salt concentration of 0.3 M (Fig. 6A). These fractions were further processed by a fractionated precipitation with ammonium sulfate. The band shift activity was retained in the 20 to 40% fraction (Fig. 6B). To estimate the molecular mass of the protein, the ammonium sulfate precipitate was dissolved and loaded onto a Sephacryl S400 gel filtration column. The band shift activity eluted shortly after the void volume with an estimated molecular mass of around 200 kDa (Fig. 6C). A lower-molecular-mass protein that also bound to E6-10 may be a degradation product of the C1 factor, as it was not detectable after ammonium sulfate fractionation.

One of the few known transcription factors with such a high molecular mass is CDP/Cut. CDP/Cut has a molecular mass of 180 kDa and, reminiscent of the repressor function of these MARs in transient transfections, is one of the few well-characterized mammalian repressors of transcription (44). Also, CDP/Cut has been found to bind the 5' part of the LCR of HPV-6 as well as sites in early genes of this virus (1, 52). Against this background, we performed EMSA competitions and supershift analyses to determine whether CDP/Cut could be the C1 factor. Figure 7A shows that the band shift generated with E6-10 is abolished by competition with two oligonucleotides harboring known binding sites for CDP/Cut, namely, one of the gp91^{phox} gene (59) (slots 6 and 7) and one of the sea urchin sperm histone H2B-1 gene (6) (slots 6 to 9). It is also eliminated by an oligonucleotide representing the HPV-16 papillomavirus silencing motif (PSM) (slots 2 and 3), which we found to bind and functionally depend on CDP/Cut (47). In contrast, competition with an E6-MAR oligonucleotide which did not form complex C1 (slots 10 and 11) or with a mutant version of PSM (slots 4 and 5) did not affect complex formation. Complex C1 is also eliminated by antiserum raised against

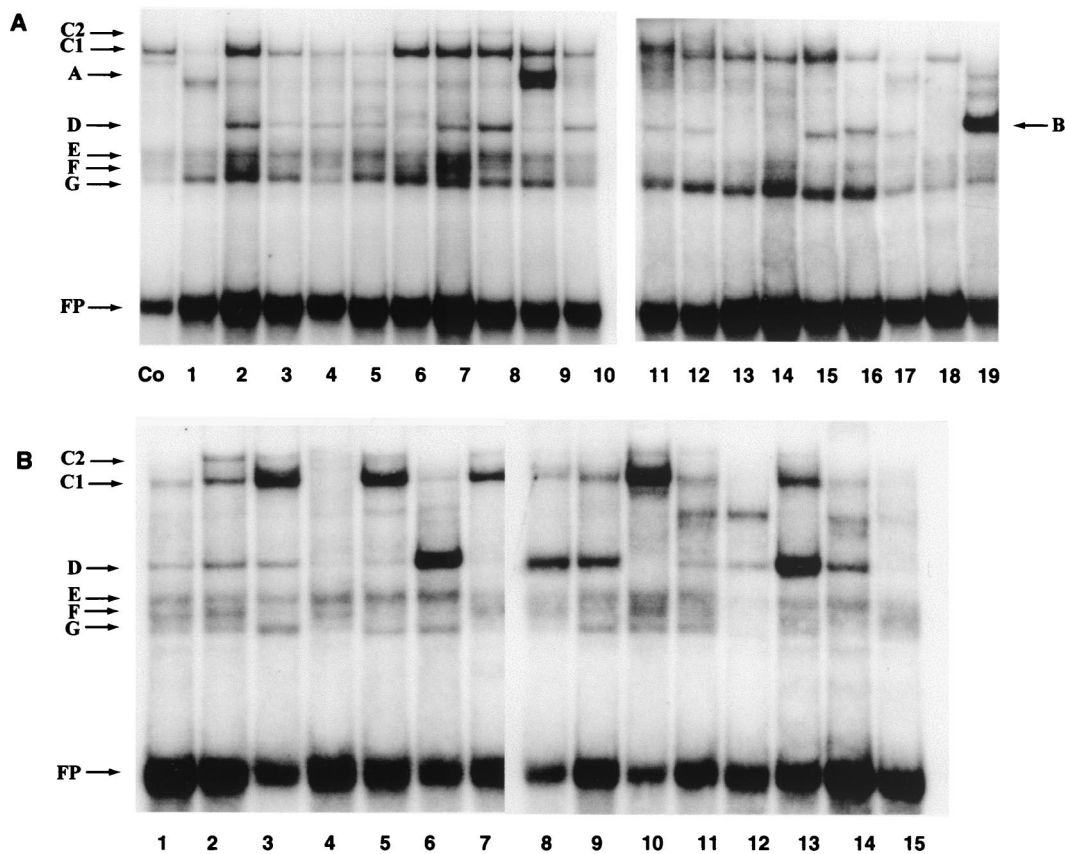


FIG. 5. EMSA screen of oligonucleotides derived from the 5'-LCR-MAR and the E6-MAR, showing strong binding sites for several nuclear proteins, some of which are common to both MARs. (A) Band shifts obtained with HeLa nuclear extracts and overlapping 31-mer oligonucleotides representing the 5' third of the HPV-16 LCR from the L1 gene to the E2 binding site most distal from the E6 promoter; (B) band shifts with 40-mer oligonucleotides representing the complete E6 gene. For details, see the footnote to Table 1 and Materials and Methods. The control (Co) is band shift of an oligonucleotide of the CDP/Cut binding site of the *gp91^{phox}* gene promoter (59). The slowest-mobility complex was termed C1, as it became clear later in this study that it is identical to the complex forming on the PSM, which we termed C1 in a previous study (48). The weak complex C2 that migrated even slower than C1 and is visible only in some slots may be a dimer of the protein giving rise to the C1 complex but could also be a heteromer. It forms efficiently on PSM, which has two flanking binding sites for C1, but only weakly on nonrepeated binding sites. Band B represents binding of YY1 to a previously described site (49), and band A represents binding of the transcription factor USF (data not shown). This study did not make an attempt to identify the proteins giving rise to bands D to G. FP, free probe.

CDP/Cut but not by preimmune serum (Fig. 7B). We conclude from this that the transcriptional repressor CDP/Cut is the factor that gives rise to complex C1 with the E6-10.

To investigate whether the slowly migrating complexes seen with other oligonucleotides of the 5'-LCR-MAR and the E6-MAR that comigrated with E6-10 complex C1 are also caused by CDP/Cut, we also analyzed the behavior of these C1 complexes. Figure 7C is an example of several EMSAs, and shows competitions with LCR-6. As can be seen, the oligonucleotide derived from the *gp91^{phox}* promoter and that from the HPV-16 PSM site are both able to compete for CDP/Cut binding, while a PSM mutant oligonucleotide has no effect on complex formation (compare lanes 2 to 5 with lanes 6 and 7). From this observation and similar data for other oligonucleotides (data not shown), we conclude that all C1 complexes of the 5'-LCR-MAR and the E6-MAR are generated by the binding of CDP/Cut and that CDP/Cut binds strongly to at least five sites within each of the two MARs.

Mutations of an oligonucleotide derived from E6 gene sequences identify an AT-rich segment as CDP/Cut binding site. To identify a binding motif for CDP/Cut, we performed EMSAs with the wild-type version of E6-10 (Fig. 8B, slots 1 and 8) and with six different mutated derivatives thereof (Fig. 8A slots 2 to 7). As can be seen, binding of the factor that

generates C1 depends on an AT-rich stretch in the 3' part of the oligonucleotide. CDP/Cut is known not to have a long and highly conserved target sequence rather, it binds to AT-rich sequences which normally include the sequence 5'-TAAT-3' (3, 28). This behavior is reflected in our EMSA results. Alteration of the TAAT element in the 3' part of this oligonucleotide either eliminates (V and VI) or reduces (II and III) CDP/Cut binding, while alteration of two other AT-rich stretches (I and IV) has no effect on the CDP/Cut band shift. We take this as additional support for the identification of these 10 CDP/Cut binding sites. Similarly, among the oligonucleotides used to scan the 5'-MAR and the E6-MAR (Table 1; Fig. 5), all 11 oligonucleotides that gave particularly strong CDP/Cut band shifts were particularly AT rich. Nine of them contained the sequence 5'-TAAT-3' (LCR-2, -6, -7, -8, and -12; E6-3, -7, -10, and -13), while the remaining two (LCR-15 and E5-5) contained sequence elements deviating in one position from 5'-TAAT-3'.

Effect of CDP/Cut overexpression on transient and stable transfectants. To confirm that CDP/Cut represses HPV-16 transcription during transient transfection, we cotransfected HeLa cells with the construct E6-EP-Luc-E6 and two different concentrations of the CDP/Cut expression vector pMT2-CDP (Fig. 9A, columns 2 and 3) and the parental vector pMT2

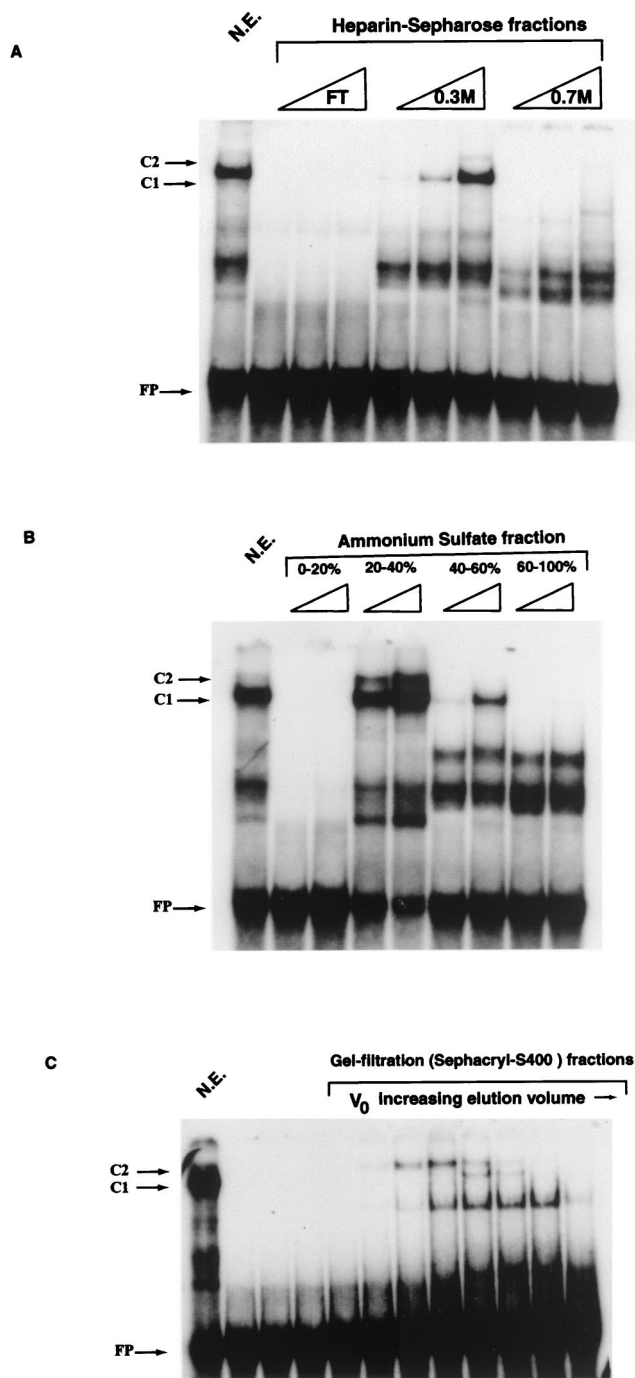


FIG. 6. Biochemical enrichment of a principal factor binding numerous sites in the 5'LCR-MAR and the E6-MAR. The purification profile of the factor binding E6-10, as monitored by EMSA, is similar to that of the differentiation-specific factor CDP/Cut. (A) Ion-exchange chromatography purification of HeLa nuclear extracts (N.E.) on a heparin-Sepharose column demonstrates that PSM binding protein activity is maximal in the 0.3 M KCl fraction. FT, flowthrough. (B) Ammonium sulfate precipitation of E6-10 binding fractions of the heparin-Sepharose chromatography defines the activity in the 20 to 40% fraction. (C) A gel filtration experiment, with the 20 to 40% ammonium sulfate fraction from the heparin-Sepharose column, indicates that the protein binding E6-10 is a very large protein of approximately 180 kDa (as defined by the β -amylase molecular weight marker). V_0 , void volume; FP, free probe. The purification profile presented here is consistent with that described for the differentiation-specific transcriptional repressor CDP/Cut (44).

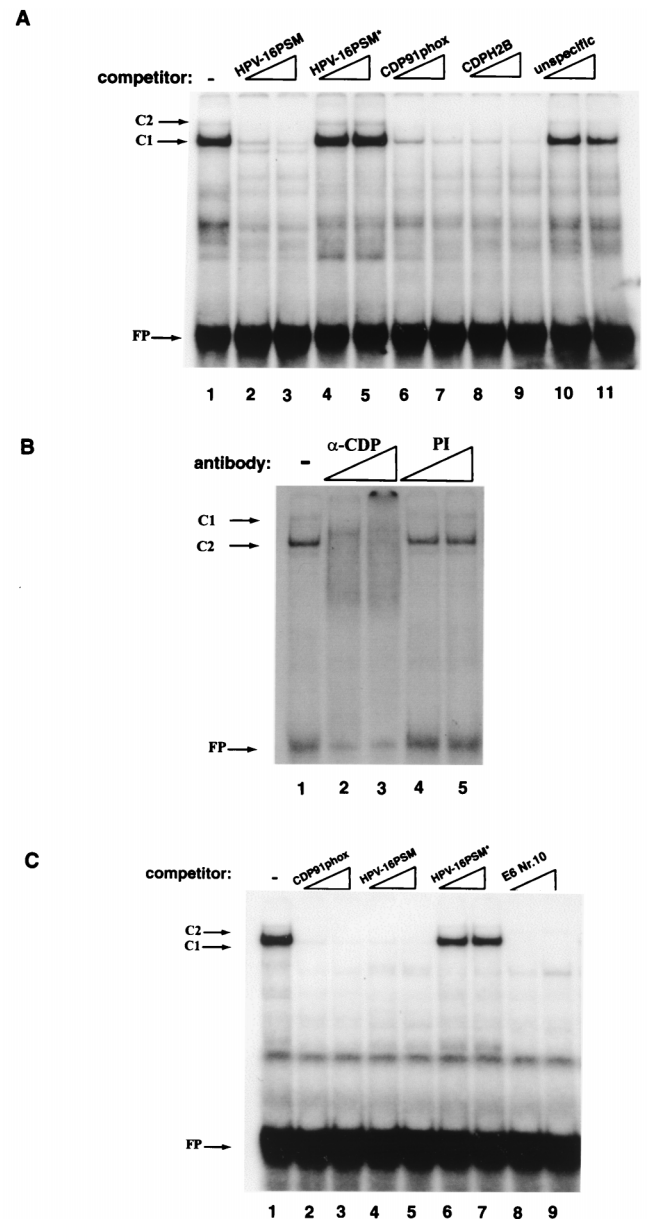


FIG. 7. Band shift competition and supershift analyses identify CDP/Cut as the factor that binds at least 10 sites in the 5'LCR-MAR and the E6-MAR. (A) E6-10 generates a strong C1 band shift as well as the weak trailing band C2 (lane 1). Formation of both complexes is eliminated by competition with an oligonucleotide representing the HPV-16 PSM (lanes 2 and 3) (47, 48) and the well-studied CDP/Cut binding sites of the promoters of the *gp91^{phox}* gene (lane 6 and 7) (59) or the histone H2B gene (lanes 8 to 9) (6) but not by a mutant PSM oligonucleotide (lane 4 and 5). As outlined by O'Connor et al. (48), we interpret complex C2 as reflection of a bound dimer binding. FP, free probe. (B) An anti-CDP antibody (α -CDP) abolishes the formation of the C1 and C2 complexes on E6-10. This effect is specific and is not seen upon the addition of a preimmune serum (PI) to the EMSA reaction. (C) LCR-6 produces C1 and C2 complexes similar to those seen with E6-10. The band shift is eliminated by the CDP/Cut binding sites of the *gp91^{phox}* promoter, the HPV-16 PSM, and E6-10 but not by a mutated PSM oligonucleotide (PSM^{*}). This experiment is representative of numerous crosswise competitions which identify at least five CDP/Cut binding sites in each of the two HPV-16 MARs.

(column 1); we observed more than 95% repression of luciferase activity with the higher amount of the CDP/Cut expression vector (Fig. 9C). This suggests that the high concentration of CDP/Cut factor in HeLa cells does not saturate the corre-

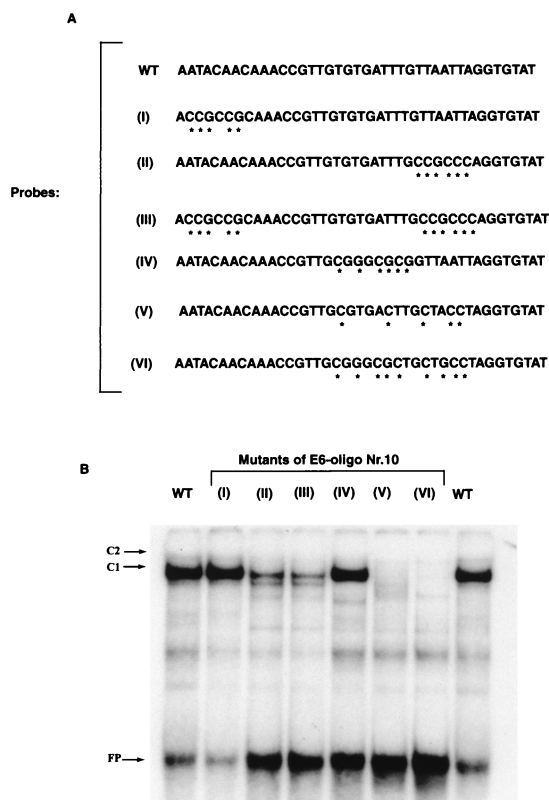


FIG. 8. Mutational analysis of one of the CDP/Cut binding sites within the E6 MAR of HPV-16. CDP/Cut does not bind to a particular highly conserved sequence but rather binds to AT-rich sequences that normally include the sequence 5'-TAAT-3' (3, 28). This behavior is confirmed by an EMSA analysis of mutations of the CDP/Cut binding site within E6-10. Alteration of the TAAT element in the 3' part of this oligonucleotide either eliminates (V and VI) or reduces (II and III) CDP/Cut binding, while alteration of two other AT-rich stretches (I and IV) has no effect on the CDP/Cut band shift.

sponding binding sites on HPV-16 and is in line with an earlier study where we failed to detect CDP/Cut binding in footprint experiments (27). We next examined whether the repression by CDP/Cut can still act on chromosomally integrated copies although it is overridden by the MAR-dependent stimulation. Under these conditions, CDP/Cut can still function as a weak repressor by reducing reporter gene expression by 35% (Fig. 9B and C) which is reminiscent of the ability of SATB1 to repress transcriptional stimulation by a chromosomally integrated MAR (35).

In both of these experiments, CDP/Cut may act through the five binding sites in the E6-MAR as well as through the two binding sites in the PSM. To separate these effects, we constructed CDP/Cut binding site mutants of the PSM (E*P-Luc), and tested them in transient transfections of constructs that contained or lacked the E6-MAR. Figure 9D shows that the mutant E*P-Luc with mutated CDP/Cut binding sites of the PSM and no MAR sequences is not affected by CDP/Cut overexpression. In contrast, CDP/Cut overexpression represses 30% of the activity of the mutant E*P-Luc-E6, which is void of the two CDP/Cut binding sites of PSM but contains the CDP/Cut sites within the E6-MAR.

DISCUSSION

CDP/Cut, a transcriptional repressor that is down-regulated during differentiation, binds conserved clusters of sites in most genital HPV types. Activity of the epithelial specific

enhancer and the E6 promoter of genital HPV types depends on a dozen different transcription factors, which induce epithelial specificity and couple transcription to physiological signals from the host. In the model systems HPV-6, -11, -16, -18, and -31 (for a review and references, see reference 46), and probably in all genital HPVs, most of the binding sites for these factors are positioned in a 550-bp segment between the E2 binding site most distal from the E6 promoter, at position 7450 in the case of HPV-16, and the transcription start site. In spite of sequence divergence among these HPVs, this genomic region is similar in composition to the most conspicuous transcription factor binding sites. Outside this 550-bp region, only few *cis*-responsive elements have been found upstream of the promoter-distal E2 binding site (2, 29, 32, 65); until recently, no *cis*-responsive element has been detected downstream of the E6 promoter.

This traditional view has been completely altered by our findings and by reports from Roman and colleagues (1, 52), as it emerges that large genomic segments outside the classical EP segment, including protein-encoding sequences, are bound by numerous nuclear proteins. These segments are *cis*-responsive elements and exert dramatic influences on HPV transcription. The most prominent factor among these nuclear proteins is, with about 10 binding sites, the factor CDP/Cut. Beyond this, our EMSA screens point to several other, not yet identified soluble factors, while interactions with the nuclear matrix suggest the binding of yet other little soluble proteins.

CDP/Cut is a repressor of transcription that acts by two alternative mechanisms, through displacement of activators (6, 42, 66, 69) and by binding the histone deacetylase HDAC1, whose activity changes nucleosomes such that it becomes difficult for the transcriptional machinery to access the DNA (38). CDP/Cut is regulated during the differentiation of a variety of cell types, having high activity in stem cells that decreases during differentiation to the target cell type. Consequently, genes that are repressed in stem cells become derepressed during differentiation (42, 44). This behavior of CDP/Cut has been extended to epithelia by studies of Roman and colleagues (1), who showed that CDP/Cut is abundant in undifferentiated and down-regulated in differentiated epithelial cells. As a consequence, CDP/Cut represses in undifferentiated epithelial cells three HPV-6 promoters by binding to promoter-flanking regions, and this repression is released in differentiating cells. One of these promoters is the E6 promoter, whose homologue we studied in HPV-16, the second is the E7 promoter, which does not exist in HPV-16, and the third is the E1/E4 promoter downstream of the HPV-16 sequences that we investigated. HeLa cells (44) and probably other cervical carcinoma-derived cell lines (our unpublished observations) have high CDP/Cut activity and in this respect resemble undifferentiated epithelial cells. This property leads to low enhancer-promoter activity of transiently transfected genital HPVs (55).

Roman and colleagues used large genomic segments of HPV-6 for the study of CDP/Cut binding, which did not allow exact localization and quantification of the binding sites. Despite this limitation, comparison of their research on HPV-6 with ours on HPV-16, and alignments of the LCRs and E6 genes of many other genital HPVs (43), suggest that these viruses contain a cluster of CDP/Cut binding sites in the 5' LCR, a directly repeated site between the enhancer and the promoter overlapping with the replication origin (47, 48), and another cluster of CDP/Cut binding sites in the E6 gene. These three elements appear to cooperate to repress HPV, probably through interaction of CDP/Cut with HDAC1 (38), which leads to deacetylation and structural changes of nucleosomes resulting in decreased transcription factor access. This mech-

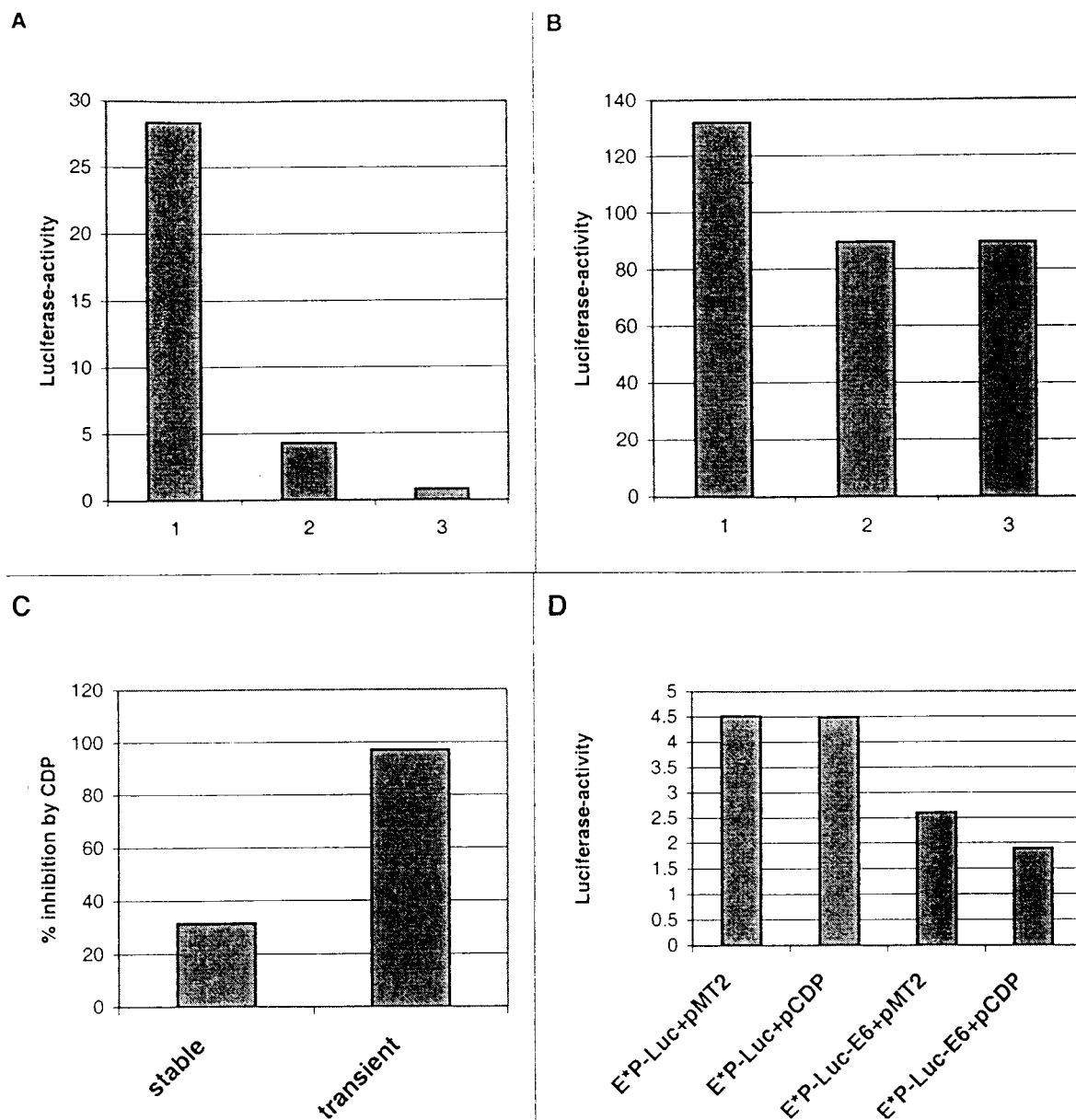


FIG. 9. Effect of CDP/Cut overexpression on transient and stable transfectants. (A) Transient transfection of HeLa cells with 1 μ g of E6-EP-Luc-E6 and 4 μ g of pMT2 (column 1) and 4 and 12 μ g of pMT2-CDP (columns 2 and 3). (B) Pools of HeLa cells stably transfected with E6-EP-Luc-E6 were transiently transfected with 4 μ g of pMT2 (column 1) and 4 and 12 μ g of pMT2-CDP (columns 2 and 3). (C) Repression by CDP/Cut in stable and transient transfection, as expressed by comparison of columns 1 and 3 in panels A and B. (D) CDP/Cut binding site mutants of PSM (E*P-Luc) without or with the E6-MAR tested in transient transfections of HeLa cells.

anism appears to apply to HPV-16, as we recently observed release of repression by the CDP/Cut-dependent silencer PSM in response to the HDAC1 inhibitor trichostatin A (47), probably due to alteration of two specifically positioned nucleosomes that limit transcription factor access to the HPV-16 enhancer and promoter (61). This finding can probably be extended to the transcription of HPV-11 and the replication of all genital HPVs, as there are reports on the stimulation of HPV-11 transcription under the influence of trichostatin A (68) and inhibition of papillomavirus replication by nucleosomes (37a) and its release by association of the E1 protein with components of the SWI-SNF complex (37). The upper part of Fig. 10 summarizes the repression of HPV-16 by coop-

eration of CDP/Cut, HDAC1, and nucleosomes in the form of a model. The lower part of Fig. 10 shows the binding of transcription factors in the absence of nucleosomes and proposes an additional somewhat more speculative activation mechanism by displacement of CDP/Cut through unknown components of the nuclear matrix.

The MARs of HPV-16 are context-dependent transcriptional enhancers. The clusters of CDP/Cut binding sites in the 5' part of the LCR and in E6 (but probably not the CDP/Cut binding sites of PSM) function in many genital HPVs as MARs, as judged by attachment assays, clusters of topoisomerase II targets, and computerized sequence analyses (62). The data in this paper do not answer the questions of what kind of DNA-

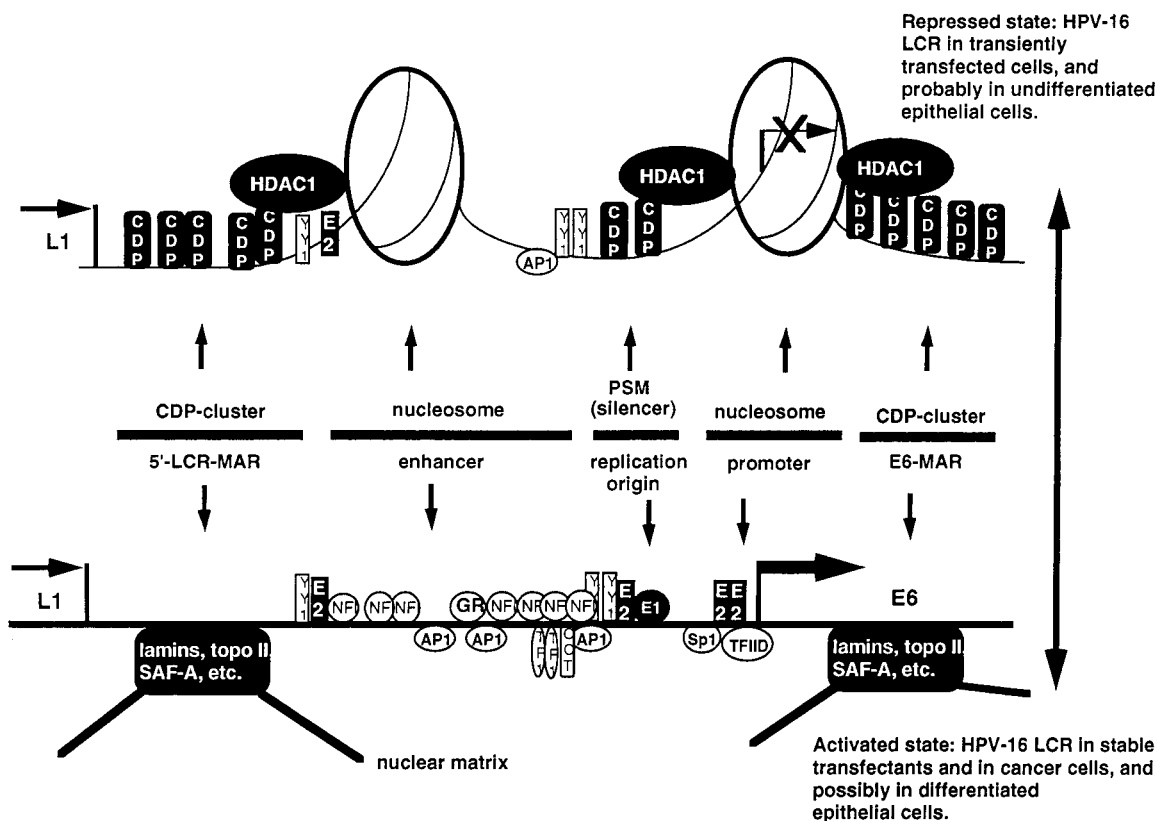


FIG. 10. Model of a switch of the HPV E6 promoter between repression by CDP/Cut, HDAC1, and two specifically positioned nucleosomes (top) and activation by interaction between unknown components of the nuclear matrix and cellular transcription factors (bottom). The figure shows 12 CDP/Cut binding sites in two MARs flanking the EP segment of HPV-16 and in a silencer motif that overlaps with the binding site of the viral replication origin, as well as the binding site of the replication protein E1 and of the most extensively researched transcriptional regulators. Elements identified previously are the MAR (62), the enhancer and its elements (46), the silencer elements (47 and 48), the replication origin (46), and the promoter elements (17 and 63). Abbreviations for genes: L1, L1 gene; E6, E6 gene; E7, E7 gene. Abbreviations for transcription factor binding sites: AP1, activator protein 1; CDP, CCAAT displacement factor/Cut; GR, glucocorticoid receptor; NFI, nuclear factor 1; TF1, transcription enhancer factor 1; oct, octamer binding factor 1.

protein interactions (i) give rise to the affinity of the HPV-16 MARs to the nuclear matrix and (ii) lead to transcriptional activation after integration of HPV-16 into the cellular DNA. It is unlikely that CDP/Cut itself is responsible for these activities, as it is a soluble factor, not an intrinsic part of the nuclear matrix, and as repression is its only known function.

CDP/Cut is sometimes referred to as a component of the nuclear matrix (5) based on the observation that it binds certain MARs (5, 66). The assumption that MAR behavior and CDP/Cut binding are expressions of the same underlying mechanism may require further investigation, as two laboratories presented evidence for an alternative scenario, in which CDP/Cut is a soluble repressor that prevents MARs from interacting with the nuclear matrix, which would induce transcription. Wang et al. (66) demonstrated that the activity of the immunoglobulin heavy-chain enhancer depends on the interaction between a MAR and the matrix-bound activator Bright and that this interaction is inhibited by CDP/Cut. In another study, it was found that SATB1 and CDP/Cut can both bind a MAR at the promoter of a mouse mammary tumor virus, thereby repressing it (39). SATB1 and CDP/Cut are able to bind to one another in the absence of DNA binding sites. The authors proposed that equimolar concentrations of SATB1 and CDP/Cut annihilate their DNA binding ability, and transcriptional induction may result from the interaction of the vacant SATB1 and CDP/Cut sites with the nuclear matrix. Similar mechanisms could apply to HPV-16, and CDP/Cut

might not only repress through the HDAC1 pathway but also limit matrix access.

There is presently no general explanation for the transcriptional activation by MARs. MARs have been reported to alter nucleosomal organization, to bring a gene into regions of the nucleus with high concentrations of transcription factors, RNA polymerases and topoisomerases, and to cause topological stress of flanking regions due to local unwinding (9, 10, 24, 41, 66). To achieve this, the nuclear matrix and MARs interact through DNA binding proteins other than typical transcription factors, such as SAF-A (54), SATB1 (19), nucleolin (20), or lamins (41). These proteins have neither nucleotide recognition specificity nor transcription activation domains, and targeting and unwinding of MARs are possibly a cause of transcriptional induction. It is unclear why this mechanism is efficient when a MAR is part of the chromosomal DNA, and why transient transfection may not allow the MAR to exert this function. Against this general lack of knowledge, we do not know whether transcription from transiently transfected HPV DNA resembles the gene expression from a stable episome. Alternatively, it is quite possible that episomally replicating HPV DNA is regulated by the nuclear matrix in the same manner as we observed only in stable transfectants. These alternatives require further research, as do the molecular properties that make the E6-MAR an strong enhancer and the 5'-LCR-MAR a weak one, although the latter one binds the matrix more strongly.

While the nuclear matrix consists of structural proteins with little sequence specificity, sequence-specific transcription factors such as YY1, nuclear factor I, and steroid receptors (for a review, see reference 11) and the RNA polymerase II holoenzyme (33) can associate with the nuclear matrix. Should CDP/Cut also be able to associate with the nuclear matrix (39), one could envisage a scenario where HPV MARs are alternatively bound by structural matrix proteins or by matrix attached CDP/Cut, resulting in a surface-bound switch between an active and an inactive state.

MARs, cis-responsive elements, and transcription factor binding sites are located within HPV genes. In cellular genes, MARs, cis-responsive elements, and transcription factor binding sites are normally in nontranscribed regions or in introns. Papillomaviruses provide an increasingly impressive example of how the lack of space on condensed viral genomes has led to deviations from such paradigms. Papillomaviruses diverge from most cellular genes by having no noncoding introns, by using some cistrons for more than one protein, and by having polycistronic mRNAs with very short nontranslated leaders. Published studies (1, 52, 62) and our data reported here show that they have (i) MARs in the E6, E5 gene, and other genes and (ii) transcription factor binding sites and cis-responsive functions in the E6, E7, and E5 genes. The combination of these observations raises the possibility that 50% or even more of the viral genome may make contact with DNA binding nuclear factors that modulate the HPV life cycle. These factors may influence other aspects of HPV biology such as replication, partition, and particle maturation, as pointed out for the nuclear compartmentalization of simian virus 40 (18).

Is HPV-16 transcription induced during carcinogenesis by integration into cellular chromosomes? As the normal life cycle of HPVs involves only episomal genomes, transcriptional induction during integration into cellular chromosomes cannot be a natural function of viral MARs. However, even in HeLa cells with their high concentration of CDP/Cut (44), chromosomal integration of HPV-16 MARs, and of the endogenous HPV-18 copies, overcomes the repression by CDP/Cut, possibly by bringing the viral genome into nuclear territories favorable to transcription. This suggests that transcriptional induction may also occur when HPV genomes integrate into cellular DNA in situ, with the result of higher expression of the E6 and E7 oncoproteins and increased oncogenic properties of the affected cell.

In malignant lesions, HPV-16 genomes are more frequently chromosomally integrated than in episomal form (16, 64). Integration occurs most often between the E7 and E2 genes, leaving oncoprotein expression intact but annihilating expression of E2 (4, 58). As the HPV-16 and HPV-18 E2 proteins are repressors of E6 and E7 transcription (17, 63), such a recombination derepresses the oncogenes and may contribute to a more aggressive cancer cell phenotype. Although this concept is most likely valid, we propose that MAR-dependent transcriptional stimulation might precede the release of E2-dependent repression. In a lesion containing only episomal copies, some viral genomes, but not simultaneously all of them, may integrate into a chromosome of any individual cell. While these integrated genomes would cease to transcribe E2, the E2 repressor is still made from the remaining episomes, leaving the cellular phenotype unaltered. Only MAR-dependent transcriptional stimulation of intrachromosomal copies constitutes a dominant genotype that may overcome E2 repression, establish enhanced E6 and E7 expression, and confer a phenotype favorable to carcinogenesis.

ACKNOWLEDGMENTS

We thank S. H. Orkin and E. Neufeld for providing plasmids pMT2 and pMT2-CDP and anti-CDP antiserum, and we thank J. Bode for stimulating discussions.

REFERENCES

- Ai, W., E. Toussaint, and A. Roman. 1999. CCAAT displacement protein binds to and negatively regulates human papillomavirus type 6 E6, E7, and E1 promoters. *J. Virol.* **73**:4220–4229.
- Auborn, K. J., and B. M. Steinberg. 1991. A key DNA-protein interaction determines the function of the 5' URR enhancer in human papillomavirus type 11. *Virology* **181**:132–138.
- Auftero, B., E. J. Neufeld, and S. H. Orkin. 1994. Sequence-specific DNA binding of individual cut repeat of the human CCAAT displacement/cut homeodomain protein. *Proc. Natl. Acad. Sci. USA* **91**:7757–7761.
- 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.
- Banan, M., I. C. Rojas, W. H. Lee, H. L. King, J. V. Harriss, R. Kobayashi, C. F. Webb, and P. D. Gottlieb. 1997. Interaction of the nuclear matrix associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene. *J. Biol. Chem.* **272**:18440–18452.
- Barberis, A., G. Superti-Furga, and M. Busslinger. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. *Cell* **50**:347–359.
- Berezney, R., and D. S. Coffey. 1974. Identification of a nuclear protein matrix. *Biochem. Biophys. Res. Commun.* **60**:1410–1470.
- Berezney, R., M. J. Mortillaro, H. Ma, X. Wei, and J. Samarabandu. 1995. The nuclear matrix: a structural milieu for nuclear genomic function. *Int. Rev. Cytol.* **162A**:1–65.
- Bode, J., Y. Kohwi, L. Dickinson, T. Joh, D. Klehr, C. Mielke, and T. Kohwi-Shigematsu. 1992. Biological significance of unwinding capability of nuclear matrix-associating DNAs. *Science* **255**:195–197.
- Bode, J., M. Stengert-Iber, V. May, T. Schlake, and A. Dietz-Pfeilstetter. 1996. Scaffold/matrix-attached regions: topological switches with multiple regulatory functions. *Crit. Rev. Eukaryot. Gene Expr.* **6**:115–138.
- Boulikas, T. 1995. Chromatin domains and prediction of MAR sequences. *Int. Rev. Cytol.* **162A**:279–388.
- Chan, W. K., G. Klock, and H. U. Bernard. 1989. Progesterone and glucocorticoid response elements occur in the long control regions of several human papillomaviruses involved in anogenital neoplasia. *J. Virol.* **63**:3261–3269.
- Chong, T., D. Apt, B. Gloss, M. Isa, and H. U. Bernard. 1991. The enhancer of human papillomavirus-16: binding sites for the ubiquitous transcription factors Oct-1, NFA, TEF-2, NF1, and AP1 participate in the epithelial specific transcription. *J. Virol.* **65**:5933–5943.
- Cockerill, P. N., M. H. Yuen, and W. T. Garrard. 1987. The enhancer of the immunoglobulin heavy chain locus is flanked by presumptive chromosomal loop anchorage elements. *J. Biol. Chem.* **262**:5394–5397.
- 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.
- Daniel, B., G. Mukherjee, L. Seshadri, E. Vallikad, and S. Krishna. 1995. Changes in the physical state and expression of human papillomavirus type 16 in the progression of cervical intraepithelial neoplasia lesions analysed by PCR. *J. Gen. Virol.* **76**:2589–2593.
- Demeret, C., C. Desaintes, M. Yaniv, and F. Thierry. 1997. Different mechanisms contribute to the E2-mediated transcriptional repression of human papillomavirus type 18 viral oncogenes. *J. Virol.* **71**:9343–9349.
- Deppert, W., and R. Schirmbeck. 1995. The nuclear matrix and virus function. *Int. Rev. Cytol.* **162A**:485–537.
- Dickinson, L. A., T. Joh, Y. Kohwi, and T. Kohwi-Shigematsu. 1992. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* **70**:631–645.
- Dickinson, L. A., and T. Kohwi-Shigematsu. 1995. Nucleolin is a matrix attachment region DNA binding protein that specifically recognizes a region with high base-unpairing potential. *Mol. Cell. Biol.* **15**:456–465.
- Dignam, J. D., R. M. Lebowitz, 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.
- Dollard, S. C., T. R. Broker, and L. T. Chow. 1993. Regulation of the human papillomavirus type 11 E6 promoter by viral and host transcription factors in primary human keratinocytes. *J. Virol.* **67**:1721–1726.
- Dong, G., T. R. Broker, and L. T. Chow. 1994. Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements. *J. Virol.* **68**:1115–1127.

24. Forrester, W. C., C. van Genderen, T. Jenuwein, and R. Grosschedl. 1994. Dependence of enhancer-mediated transcription of the immunoglobulin μ gene on nuclear matrix attachment regions. *Science* **265**:1221–1225.
25. Gasser, S. M., and U. K. Laemmli. 1986. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell* **46**:521–530.
26. Gloss, B., H. U. Bernard, K. Seedorf, and G. Klock. 1987. The upstream regulatory region of the human papillomavirus-16 contains an E2 protein independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones. *EMBO J.* **6**:3735–3743.
27. Gloss, B., T. Chong, and H. U. Bernard. 1989. Numerous nuclear proteins bind the long control region of human papillomavirus type 16: a subset of 6 of 23 DNase I-protected segments coincides with the location of the cell-type-specific enhancer. *J. Virol.* **63**:1142–1152.
28. Harada, R., G. Berube, O. J. Tamplin, C. Denis-Larose, and A. Nepveu. 1995. DNA-binding specificity of the cut repeats from the human cut-like protein. *Mol. Cell. Biol.* **15**:129–140.
29. Hoppe-Seyler, F., and K. Butz. 1993. A novel cis-stimulatory element maps to the 5' portion of the human papillomavirus type 18 upstream regulatory region and is functionally dependent on a sequence-aberrant Sp1 binding site. *J. Gen. Virol.* **74**:281–286.
30. Howley, P. M. 1996. Papillomavirinae: the viruses and their replication, p. 2045–2076. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), *Fields virology*. Lippincott-Raven, Philadelphia, Pa.
31. International Agency for Research on Cancer. 1995. IARC monographs on the evaluation of carcinogenic risks to humans, vol. 64. Human papillomaviruses. International Agency for Research on Cancer, Lyon, France.
32. Kanaya, T., S. Kyo, and L. A. Laimins. 1997. The 5' region of the human papillomavirus type 31 upstream regulatory region acts as an enhancer which augments viral early expression through the action of YY1. *Virology* **237**:159–169.
33. Kimura, H., Y. Tao, R. G. Roeder, and P. R. Cook. 1999. Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. *Mol. Cell. Biol.* **19**:5383–5392.
34. Klehr, D., K. Maass, and J. Bode. 1991. Scaffold-attached regions from the human interferon β domain can be used to enhance the stable expression of genes under the control of various promoters. *Biochemistry* **30**:1264–1270.
35. Kohwi-Shigematsu, T., K. Maass, and J. Bode. 1997. A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-lined reporter genes. *Biochemistry* **36**:12005–12010.
36. Lamond, A. L., and W. C. Earnshaw. 1998. Structure and function in the nucleus. *Science* **280**:547–553.
37. Lee, D., H. Sohn, G. V. Kalpana, and J. Choe. 1999. Interaction of E1 and hSNF5 proteins stimulates replication of human papillomavirus DNA. *Nature* **399**:487–491.
- 37a. Li, R., and M. R. Botchan. 1994. Acidic transcription factors alleviate nucleosome-mediated repression of DNA replication of bovine papillomavirus type 1. *Proc. Natl. Acad. Sci. USA* **91**:7051–7055.
38. Li, S., L. Moy, N. Pittman, G. Shue, B. Aufiero, E. J. Neufeld, N. S. LeLeiko, and M. J. Walsh. 1999. Transcriptional repression of the cystic fibrosis transmembrane conductance regulator gene, mediated by CCAAT displacement protein/cut homolog, is associated with histone deacetylation. *J. Biol. Chem.* **274**:7803–7815.
39. Liu, J., A. Barnett, E. J. Neufeld, and J. P. Dudley. 1999. Homeoproteins CDP and SATB1 interact: potential for tissue-specific regulation. *Mol. Cell. Biol.* **19**:4918–4926.
40. Liu, Y., J. Chen, Q. Gao, S. Dalal, Y. Hong, C. P. Mansur, V. Band, and E. Androphy. 1999. Multiple functions of human papillomavirus type 16 E6 contribute to the immortalization of mammary epithelial cells. *J. Virol.* **73**:7297–7307.
41. Luderus, M. E. E., J. L. den Blaauwen, O. J. B. de Smit, D. A. Compton, and R. van Driel. 1994. Binding of matrix attachment regions to lamin polymers involves single-stranded regions and the minor groove. *Mol. Cell. Biol.* **14**:6297–6305.
42. Mailly, F., G. Berube, R. Harada, P. L. Mao, S. Phillips, and A. Nepveu. 1996. The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy. *Mol. Cell. Biol.* **16**:5346–5357.
43. Myers, G., H. U. Bernard, H. Delius, M. Favre, J. Icenogle, M. van Ranst, and C. Wheeler. 1994. Human papillomaviruses. A compilation and analysis of nucleic acid and amino acid sequences. Los Alamos National Laboratory, Los Alamos, N.Mex.
44. Neufeld, E. J., D. G. Skalnick, P. M. J. Lievens, and S. H. Orkin. 1992. Human CCAAT displacement protein is homologous to the *Drosophila* homeoprotein, cut. *Nat. Genet.* **1**:50–55.
45. Nickerson, J. A., G. Krockmalnic, K. M. Wan, and S. Penman. 1997. The nuclear matrix revealed by eluting chromatin from a cross-linked nucleus. *Proc. Natl. Acad. Sci. USA* **94**:4446–4450.
46. O'Connor, M., S. Y. Chan, and H. U. Bernard. 1995. Transcription factor binding sites in the long control regions of genital HPVs, p. 21–40. *In* G. Myers, H. U. Bernard, H. Delius, C. Baker, J. Icenogle, A. Halpern, and C. Wheeler (ed.), *Human papillomaviruses 1995, part III-A. Compendium*. Los Alamos National Laboratory, Los Alamos, N.Mex.
47. O'Connor, M. J., W. Stünkel, C. H. Koh, H. Zimmermann, and H. U. Bernard. 2000. The differentiation-specific factor CDP/Cut represses transcription and replication of human papillomaviruses through a conserved silencing element. *J. Virol.* **74**:401–410.
48. O'Connor, M. J., W. Stünkel, H. Zimmermann, C. H. Koh, and H. U. Bernard. 1998. A novel YY1-independent silencer represses the activity of the human papillomavirus type 16 enhancer. *J. Virol.* **72**:10083–10092.
49. O'Connor, M. J., S. H. Tan, C. H. Tan, and H. U. Bernard. 1996. YY1 represses human papillomavirus type 16 transcription by quenching AP-1 activity. *J. Virol.* **70**:6529–6539.
50. Parker, J. N., W. Zhao, K. J. Askins, T. R. Broker, and L. T. Chow. 1997. Mutational analysis of differentiation-dependent human papillomavirus type-18 enhancer elements in epithelial raft cultures of neonatal foreskin keratinocytes. *Cell Growth Differ.* **8**:751–762.
51. Pater, M. G., G. A. Hughes, D. E. Hyslop, H. Nakshatri, and A. Pater. 1988. Glucocorticoid-dependent oncogenic transformation by type 16 but not type 11 human papilloma virus DNA. *Nature* **335**:832–835.
52. Pattison, S., D. G. Skalnick, and A. Roman. 1997. CCAAT displacement protein, a regulator of differentiation-specific gene expression, binds a negative regulatory element within the 5' end of the human papillomavirus type 6 long control region. *J. Virol.* **71**:2013–2022.
53. Rapp, L., and J. J. Chen. 1998. The papillomavirus E6 proteins. *Biochim. Biophys. Acta* **1378**:F1–F19.
54. Romig, H., F. O. Fackelmayr, A. Renz, U. Ramsperger, and A. Richter. 1992. Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment elements. *EMBO J.* **11**:3431–3440.
55. Sailaja, G., R. M. Watts, and H. U. Bernard. 1999. Many different papillomaviruses have low transcriptional activity in spite of strong epithelial specific enhancers. *J. Gen. Virol.* **80**:1715–1724.
56. Schiffman, M. H., and L. A. Brinton. 1995. The epidemiology of cervical carcinogenesis. *Cancer* **76**:1888–1901.
57. Schübeler, D., C. Mielke, K. Maass, and J. Bode. 1996. Scaffold/matrix-attached regions act upon transcription in a context-dependent manner. *Biochemistry* **35**:11160–11169.
58. 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* **314**:111–114.
59. Skalnick, D. G., E. C. Strauss, and S. H. Orkin. 1991. CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter. *J. Biol. Chem.* **266**:16736–16744.
60. Strouboulis, J., and A. P. Wolffe. 1996. Functional compartmentalization of the nucleus. *J. Cell Sci.* **109**:1991–2000.
61. Stünkel, W., and H. U. Bernard. 1999. The chromatin structure of the long control region of human papillomavirus type 16 represses viral oncoprotein expression. *J. Virol.* **73**:1918–1930.
62. Tan, S. H., D. Bartsch, E. Schwarz, and H. U. Bernard. 1998. Nuclear matrix attachment regions of human papillomavirus type 16 point toward conservation of these genomic elements in all genital papillomaviruses. *J. Virol.* **72**:3610–3622.
63. Tan, S. H., L. E. C. Leong, P. Walker, and H. U. Bernard. 1994. The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. *J. Virol.* **68**:6411–6420.
64. Vernon, S. D., E. R. Unger, D. L. Miller, D. R. Lee, and W. C. Reeves. 1997. Association of human papillomavirus type 16 integration in the E2 gene with poor disease-free survival from cervical cancer. *Int. J. Cancer* **74**:50–56.
65. Wang, H., K. Liu, F. Yuan, L. Berdichevsky, L. B. Taichman, and K. Auburn. 1996. C/EBP β is a negative regulator of human papillomavirus type 11 in keratinocytes. *J. Virol.* **70**:4839–4844.
66. Wang, Z., A. Goldstein, R. T. Zong, D. Lin, E. J. Neufeld, R. H. Scheuermann, and P. W. Tucker. 1999. Cux/CDP homeoprotein is a component of NF- μ NR and represses the immunoglobulin heavy chain intronic enhancer by antagonizing the Bright transcription enhancer. *Mol. Cell. Biol.* **19**:284–298.
67. Wei, X., J. Samarabandu, R. S. Devdhar, A. J. Siegel, R. Acharya, and R. Berezney. 1998. Segregation of transcription and replication sites into higher order domains. *Science* **281**:1502–1506.
68. Zhao, W., F. Noya, W. Y. Chen, T. M. Townes, L. T. Chow, and T. R. Broker. 1999. Trichostatin A up-regulates human papillomavirus type 11 upstream regulatory region-E6 promoter activity in undifferentiated primary human keratinocytes. *J. Virol.* **73**:5026–5033.
69. Zong, R. T., and R. H. Scheuermann. 1995. Mutually exclusive interaction of a novel matrix attachment region binding protein and the NF- μ NR enhancer repressor. *J. Biol. Chem.* **270**:24010–24018.
70. zur Hausen, H. 1996. Papillomavirus infections—a major cause of human cancers. *Biochim. Biophys. Acta* **1288**:F55–F78.