Methylation Sensitivity of the Enhancer from the Human Papillomavirus Type 16*

(Received for publication, September 3, 1993, and in revised form, December 23, 1993)

Heinz-Joachim List, Volker Patzel, Ulrich Zeidler, Arndt Schopen, Gabriele Rühl, Jürgen Stollwerck, and Gerd Klock**

From the Institut für Biochemie, Technische Hochschule Darmstadt, Petersenstrasse 22, D-64287 Darmstadt, Germany

The human papillomavirus type 16 is associated with anogenital cancer. Transcription of the viral transforming genes E6 and E7 is under the control of an epithelial cell type-specific enhancer. In the enhancer core, we have identified a regulatory element that is recognized by a novel nuclear factor named MSPF (methylation-sensitive papillomavirus transcription factor). Mutating the MSPF binding site strongly affects the enhancer activity. The MSPF recognition sequence 5'-ATGGCNGNNCCGCTC-3' contains two CpG dinucleotides, potential targets for 5-cytidine methylation. DNA recognition by MSPF is strictly methylation-sensitive, since introduction of 5-methylcytidine into either CpG abolishes complex formation. Moreover, CpG methylation of the MSPF binding site suppresses the activity of the enhancer and of the MSPF enhancer subfragment in vivo. In the cervical carcinoma cell line CaSki, which has integrated multiple transcriptionally inactive human papilloma virus 16 genomes, a few of the viral genomes are methylated at the MSPF binding site. These findings suggest that viral transcription can be suppressed by methylation of the regulatory region, an event that prevents binding of the cellular transcription factor MSPF.

with the cellular tumor suppressor proteins p53 and Rb (9–11). HPV DNA can be detected in about 90% of cervical carcinomas, where the viral DNA is frequently found integrated in the cellular genome (1, 12, 13). During integration, the E6 and E7 genes are conserved (13, 14), but the viral E2 open reading frame is usually lost or disrupted, an event that may lead to increased E6 and E7 transcription (15).

Since the E6 and E7 proteins play a key role in tumorigenesis of HPVs, it is important to understand how their gene expression is regulated. Cellular factors are thought to repress viral gene expression in nontumorigenic cells but may be lost in tumorigenic cells (1, 16, 17). The E6 and E7 genes from HPV-16 are transcribed from the promoter P97 in the long control region (LCR; 18). P97 is under control of an upstream enhancer that shows a marked specificity for keratinocytes and cervical carcinoma cells (19, 20). This cell specificity is retained on enhancer subfragments that are about 90 bp long (21–23). The basal enhancer activity can be stimulated further by steroid hormones through a glucocorticoid response element (GRE)/progesterone response element in the 90-bp enhancer core (19, 24). The viral enhancer is bound by a number of additional transcription factors (25), some of which have been identified as AP-1 (21, 26), NF-1 (27), and TEF-1 (28).

Like cellular genes, several viral DNA genomes have been found to be methylated (29–31), a modification that introduces 5-methylcytidine into 5'-CpG-3' dinucleotides (for review, see Refs. 32 and 33). In general there exists a correlation between hypomethylation at CpG sites of vertebrate genes and gene expression. The correlation is best with cell type-specific genes, which are usually found unmethylated in expressing cells but methylated in nonexpressing cells (32). The concept that DNA methylation could be responsible for the cell type specificity is supported by the expression behavior of transfected genes. When a methylated copy of the muscle-specific α-actin gene is introduced into myoblast cells, demethylation at specific CpG sites and gene activation are observed (34). DNA methylation may also play a role in developmentally regulated gene expression, as both de novo methylation and demethylation of specific genes are observed in embryonic development (33, 35–37). Transcriptional regulation by DNA methylation may result from the failure of transcription factors to bind to their methylated recognition sites. Recently several factors have been identified which show CpG methylation-sensitive DNA binding in vitro (38–44). An alternative model proposes that nuclear proteins bind to methylated CpGs (5′-5-mCpG-3′ dinucleotides) and thereby repress transcription (45–47).

The enhancer from HPV-16 contains a potential methylation site with three CpG dinucleotides. We show here that a sequence containing two of these CpGs is recognized by a nuclear protein which we call MSPF. 5-Cytidine methylation of the CpGs both inhibits MSPF binding and suppresses transcription in vivo. Therefore, HPV-16 early transcription responds to

Human papillomaviruses (HPVs) infect epithelial cells and induce benign hyperproliferative lesions, known as papillomas or warts (1). Among more than 60 different HPV types that have been identified, the viruses of one HPV group, represented by HPV-16 and -18, are linked to anogenital cancer and cause cervical intraepithelial neoplasia that have a high risk of malignant progression (1–3). The viral early genes E6 and E7 induce benign hyperproliferative lesions, known as papillomas (4–8), which appear to result from their ability to form complexes with the cellular tumor suppressor proteins p53 and Rb (9–11).

This work was supported by Deutsche Forschungsgemeinschaft Grant KI 5292-1 and the Mildred Scheel Stiftung für Krebsforschung Grant M13/91/1. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

* This work was supported by Deutsche Forschungsgemeinschaft Grant KI 5292-1 and the Mildred Scheel Stiftung für Krebsforschung Grant M13/91/1. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

** To whom correspondence should be addressed.

† Present address: Angewandte Tumorvirologie, Im Neuenheimer Feld 280, D-6900 Heidelberg, Germany.
‡ Present address: Fakultät Biologie, Universität Konstanz, D-78434 Konstanz, Germany.
§ Present address: Heinrich-Pette-Institut, Universität Hamburg, D-2000 Hamburg 20, Germany.
¶ Present address: Applied Biosystems, D-6100 Weiterstadt, Germany.

1 The abbreviations used are: HPV(s), human papillomavirus(es); LCR, long control region; bp, base pair(s); GRE, glucocorticoid response element; MSPF, methylation-sensitive papillomavirus transcription factor; CAT, chloramphenicol acetyltransferase; EMSA, electrophoretic mobility shift assay; kb, kilobase(s); GRU, glucocorticoid response unit; NF, nuclear factor; AP, activator protein; TEF, transcriptional enhancer factor.
the methylation status of the viral regulatory region by recruiting MSPF, a methylation-sensitive transcription factor.

MATERIALS AND METHODS

Oligonucleotides and Plasmids—Oligonucleotides were synthesized on an Applied Biosystems 380A DNA synthesizer by the phosphoramidite method and were purified by high performance liquid chromatography before use. For the introduction of methylated cytidines, 5-methylcytidine derivatives of the phosphoramidites were used. For DNA binding studies (electrophoretic mobility shift assays, EMSA), complementary strands were designed to leave four nucleotides 5’ sticky ends. The sequence of one strand of each double-stranded oligonucleotide after the fill-in is given 5’ to 3’ (capital letters, HPV-16 [48] wild type or mutant sequence).

fp4e-P, (7646–7656); gatcATTGTTCATGgatc.

fp4e-Q (7610–7639); catgATACCTGCATATGGCACTACTGGAATCTCAACAG.

fp5e-L (7667–7691); tcgacATCACTGCGCAACGCTTACATAGgccag.

fp5e-L-mt1 (7667–7691); tcgacATCCTACGTAGTCGCTGTTATGAGgccag.

fp5e-L-mt2 (7667–7691); tcgacATCTGGCACCACCGTCTATGAggccag.

fp5e-L-mt3 (7667–7691); tcgacATCTGGCACCACCGTCTATGAggccag.

fp5e-L-mt4 (7667–7691); tcgacATCTGGCACCACCGTCTATGAggccag.

fp5e-L-mt5 (7667–7691); tcgacATCTGGCACCACCGTCTATGAggccag.

fp7e-L (7732–7756); tcgacATCAGCACTTGGCAATAGGTTAggccag.

fp8e-9e-L (7785–7812); tcgtcattcACCTGCATGACCAGCGTAAAGGTTAggccag.

tat-GRU-D (−3248 to −2403); tcgaAAGCTTCCATGCTAGGTCCG.

tccagaTACACTGCACTATGTGCAACTACTGAA-

tcgacTCACTGCGCAACGCTTACATAGgccag.

tcgacGATTCCAGCGTAGTCGCAAGTTGA-

tcgacGATTCATGCGTAGTCGCCTGTTGA-

tcgacaTCACTATGATCCAAGTCCTTACATAgg-

tcgacTTACAGCATATlTGGCATAAGGTTAgg-

tcgacGTTAGGCACAT-3'. With the first method, we followed the pro-
ger2 DNA polymerase (Sequenase Version 2.0). EMSAs were done essentially as described (26). In short, 5–10 µg of nuclear protein was preincubated for 10 min on ice with 1.5 µg of poly (dA-dT). After the addition of a 50-fold molar excess (or 100-fold, where indicated) of unlabeled competitor oligonucleotide followed by a 10-min incubation on ice, 18 fmol of 32P-labeled oligonucleotide was added. After incubation at 25 °C for 10 min, the complexes were separated on a 5% polyacrylamide gel (60:1 bisacrylamide) containing 2.5% glycerol in 50 mM Tris base, 380 mM glycine, and 2 mM EDTA, pH 7.8. The gel was dried and analyzed by autoradiography.

RESULTS

CpG Methylation Inhibits Specific Protein Binding to fp5e in the HPV-16 Enhancer—DNAse I footprinting experiments have revealed that numerous nuclear proteins recognize the HPV-16 enhancer (25). The 90-bp core enhancer contains in the fp5e region an NF-1 semipalindromic sequence (half-site) TTGGC and three CpG dinucleotides (Fig. 1). Since CpGes are potential targets for 5-cytidine methylation, we decided to investigate the binding specificity of the fp5e sequence covering two Cpgs (Fig. 1) by EMSA. An EMSA with the fp5e oligonucleotide and HeLa nuclear extracts yields one specific complex (MSPP, see below), which can be competed by the homologous oligonucleotide (Fig. 2A, lanes 1 and 2). In contrast, an excess of a second NF-1 half-site, TTGGCsC, a somewhat closer to the consensus NF-1 site, does not compete the complex (lane 3). This is in agreement with previous observations (60), but purified NF-1 could also bind to fp5e (27). Two unrelated DNA fragments from the HPV-16 enhancer, fp4e-Q and fp8e-9e-L, do not compete the fp5 complex (Fig. 2A, lanes 4 and 5). Conversely, a specific complex formed on the fp7e fragment was competed well by an excess of the homologous sequence, but only weakly by fp5e-L (data not shown), suggesting that the
two sequences bind different proteins. A second band with a
more blurred appearance marked with X shows a similar
competition as the specific band (Fig. 2, A and B). We cannot rule
out that the band X complex may also be specific; however, we
consider this unlikely since the complex appeared with largely
varying intensities in a number of experiments (data not shown).

To narrow down the sequence in fp5e responsible for forma-
tion of the specific complex, we synthesized several fp5e-L mu-
tants. fp5e-L-mt2, which retains only an 18-bp wild type se-
quency in the center of fp5e, still competes the fp5e complex
(Fig. 2B, lane 3). The mutant fp5e-L-mt1 in which the two CpG
dinucleotides were exchanged does not compete (lane 5). A
third mutant, fp5e-L-mt3, in which the NF-1 pentamer is mu-
tated from TTGGC to ABC, competes the fp5e complex (Fig.
2B, lane 4). When this mutant was radioactively labeled and
used in the EMSA, it gave a complex with the same mobility
and the same competition behavior as the wild type 25-mer
fp5e-L (data not shown). We noticed a close similarity between
fp5e-L and a sequence containing two CpGs in the rat TAT gene
with DNA that is completely unmethylated on both CpGs (for nomencla-
ture of partially methylated double strands, see Fig. 2D). When these DNAs were
used as competitors of

A

B

C

FIG. 1. Binding sites in the HPV-16 enhancer core. Panel A, the viral LCR is flanked by L1 and E6 open reading frames. It contains a cell
type-specific enhancer that controls the promoter P97. The enhancer is located on a 399-bp fragment (position 7455–7853) (19, 48), and full
transcriptional activity is found on a 232-bp fragment (shaded), which includes the 90-bp enhancer core (black bar) (21, 22). Panel B, sequence
of the protein binding site fp5e identified in DNase I footprinting experiments (25, 60). The fp5e-L oligonucleotide used in EMSAs is marked (black
bar; see "Materials and Methods"). fp5e contains three CpG dinucleotides and overlaps with a TTGGC pentamer (arrow; 27). Panel C, sequence
comparison with a related motif from the HPV-16 enhancer. fp7e-L, which binds to the NF-1 protein (27), was used in EMSAs. Bases identical with
the sequence in panel B are underlined. Base exchanges were introduced in fp7e-L (marked by dots) to destroy an Oct-1/NF-A binding site at the
right end of fp7e-L (23).

pentamer TTGGC is not required, whereas mutating both CpG
dinucleotides completely abolishes binding. A sequence contain-
ing only two pentamers flanking the NF-1 motif is suffi-
cient for formation of the complex (Fig. 2F). The MSPF consen-
sus sequence that we have derived, 5'-ATGCGNNNGCGCTT-
3', is an imperfect palindrome with the repeated half-site
CGCCY'T.

Several DNA-binding proteins that recognize sequences con-
taining CpG dinucleotides do not bind to their methylated bind-
ing sites (see Introduction), whereas at least one factor does not
discriminate between methylated and nonmethylated CpGs
(61). Since mutating the two CpG dinucleotides destroyed the
fp5e binding site, we investigated the influence of 5-cytidine
methylation on specific protein-DNA complex formation. The
oligonucleotide fp5e-L-me1234, which contains 5-methylcyti-
dine at all four positions of the two CpGs in fp5e, does not
compete the fp5e complex (Fig. 2D, lane 4), indicating that
5-cytidine methylation of fp5e inhibits protein-DNA complex
formation. In addition, neither the methylated oligonucleotide
fp5e-L-1234 nor the mutant fp5e-L-mt1 formed complexes in an
EMSA (data not shown).

To determine which methylated cytidines affect protein bind-
ing, oligonucleotides were prepared which were either un-,
fully, or fully methylated at each of the CpGs (for nomenclature
of partially methylated double strands, see Fig. 2D and
text of Fig. 2). When these DNAs were used as competitors of
the fp5e-L complex, none of the methylated oligonucleotides
was found to be an effective competitor (Fig. 2F). We conclude
that the protein-DNA complex on fp5e-L can only be formed
with DNA that is completely unmethylated on both CpG
dinucleotides. Since, to our knowledge, a protein with this bind-
ing specificity has not been described, we have named the pro-
tein that binds to unmethylated fp5e DNA MSPF (methylation-
sensitive papillomavirus transcription factor).

MSPF Binding Activity in Cell Lines Does Not Correlate with
the Epithelial Cell Specificity of the HPV-16 Enhancer—The
transcriptional enhancer from HPV-16 shows a marked prefer-
ence for epithelial cells. To determine the possible contribu-

---

*Methylation-sensitive HPV-16 Enhancer*
of MSPF to this cell type specificity, MSPF binding activities were examined in nuclear extracts from cell lines of epithelial and nonepithelial origin. The binding specificity was tested by homologous and heterologous competition of fp5e complexes in an EMSA. The epithelial cell types comprised SiHa (cervical carcinoma), HPK IA (HPV-16-immortalized keratinocyte), and MCF-7 (breast adenocarcinoma) cell lines and the nonepithelial cells Balb-3T3 (fibroblast) and SK-MEL-28 (melanoma). Specific binding activity to fp5e-L characteristic of MSPF was observed in nuclear extracts from all five cell lines, with the greatest binding activity in SiHa cells (Fig. 3A). The complexes formed with the different cell lines all have identical mobility and could only be competed by an excess of homologous oligonucleotide, but not by the mutant fp5e-L-mt1 or the fp7e-L fragment (Fig. 3A). We conclude that a factor with binding properties similar to MSPF is present in all cells tested, includ-
Methylation-sensitive HPV-16 Enhancer

**A**

<table>
<thead>
<tr>
<th>Competitor</th>
<th>MSPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>Relative CAT-activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>1.5</td>
</tr>
<tr>
<td>3.2</td>
</tr>
<tr>
<td>3.7</td>
</tr>
<tr>
<td>5.5</td>
</tr>
</tbody>
</table>

**C**

<table>
<thead>
<tr>
<th>Relative CAT-activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>1.3</td>
</tr>
<tr>
<td>1.2</td>
</tr>
<tr>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
</tr>
</tbody>
</table>

**Fig. 3.** MSPF activity in several cell lines does not show specificity for epithelial cell types. Panel A, EMSAs were performed with fp5e-L and nuclear extracts from several cell lines of epithelial (SiHa, HPK IA, MCF-7) and non epithelial origin (Balb-3T3, SK-MEL-28). For each extract tested, competitors were added as follows: lane A, none; lane B, fp5e-L (homologous); lane C, fp5e-L-mt1; lane D, fp7e-L. The MSPF complex is observed in all lanes but is strongest in the cervical carcinoma cell line SiHa. Panels B and C, transcriptional activation of CAT expression by the MSPF binding site. Wild type and mutant oligonucleotides were cloned into pBLCAT5 upstream of the thymidine kinase (TK) promoter. The orientation of the inserts is given by arrows pointing to the right for sense and to the left for antisense orientation. The plasmids were transiently transfected into the cell lines listed, and CAT enzymatic activity was determined. The given CAT values were normalized to the activities obtained with pBLCAT5 for each cell line. Panel B, transcriptional enhancer effect of the MSPF binding site. The vector pBLCAT5 and the plasmids pH16ML-1 to -4, which contain the fp5e-L sequence, were transfected into three epithelial (HeLa, SiHa, MCF-7) and one fibroblast cell line (Balb-3T3). Increasing numbers of MSPF binding sites lead to a transcriptional stimulation which is maximal with the tetrameric insertion in pH16ML-4. Panel C, the plasmids pH16ML-mt1 to -mt4 contain the mutant sequence fp5e-L-mt1, which does not bind MSPF. Single or multiple inserts do not lead to an increased CAT expression compared to the vector pBLCAT5. The HPV-16 enhancer is complex, comprising regulatory elements.
methylation-sensitive HPV-16 Enhancer

Fig. 4. Site-directed mutagenesis of the whole enhancer reveals that protein binding to CpG dinucleotides correlates with enhancer activity. Mutations were introduced into two sites of the HPV-16 enhancer in pH16EPH-wt, which carries the wild type enhancer fragment (position 7455–7853) upstream of the thymidine kinase (TK) promoter from herpes simplex virus and the CAT gene. The mutation in the MSPF binding site destroys MSPF binding (pH16EPH-5eL); the MSPF- and TEF-1 site (pH16EPH-5eK) mutants carry base exchanges in the CpG dinucleotides. The plasmids were transiently transfected into the two cervical carcinoma cell lines HeLa and SiHa, respectively, and CAT activity was determined. For each mutation, two independent isolates were used: one mutant was constructed by a standard mutagenesis protocol and one by a variant protocol using gapped duplex DNA. The CAT activities for each plasmid represent the average of several independent transfections with the two isolates; the resulting CAT activities of each mutant were normalized to the

CpG Methylation of the MSPF Binding Site Abolishes Its Enhancing Activity in Vivo—The CpG-containing region fp5e/ fp6e in the enhancer core plays an important role in the function of the whole enhancer (see Fig. 4). The finding that introduction of 5-methylcytidine into CpG dinucleotides inhibits the in vitro binding of MSPF (Fig. 2) raises the possibility that DNA methylation may affect the HPV-16 enhancer function in vivo. One of the two CpG dinucleotides in the MSPF binding site is part of the sequence 5′-GGCG-3′, a recognition site for the HhaI DNA methylase. We used HhaI methylase to introduce 5-methylcytidine as the modified nucleotide into the CpGs of all HhaI sites of the following plasmids: pH16EN-1, which carries the 400-bp enhancer fragment; pH16ML-4, containing the MSPF binding site tetramer (see Fig. 3B); and the vector pBLCAT5 (the HhaI sites in the regulatory regions of these plasmids are shown in Fig. 5). The three methylated plasmids and their unmethylated counterparts were transiently transfected into HeLa cells, and the CAT enzyme activity was determined. The complete experiment, including DNA methylation and transfection, was repeated twice, and the average CAT activities from the three transfection experiments are shown in Fig. 5.

CAT expression of the vector pBLCAT5 alone is not affected by the methylation. In contrast, HhaI methylation suppresses the activity of the enhancer plasmid pH16EN-1 (Fig. 5). This inhibiting effect can be attributed to methylation of the CpG dinucleotide in fp5e, the binding site for MSPF, since only one HhaI site is present in the 400-bp enhancer. Moreover, the enhancer activity of the MSPF binding site is completely suppressed by methylation of one CpG dinucleotide in each copy of the tetrameric fp5e insert (Fig. 5). In conclusion, these results clearly demonstrate that the fp5e-L fragment functions as a 5-C-methylation-sensitive enhancer. The CpG in the HhaI site in fp5e is one of two methylation signals in the binding site for MSPF, whose DNA binding activity is inhibited following methylation of either CpG dinucleotide (Fig. 2).

Most cervical carcinomas and carcinoma cell lines contain HPV DNA integrated in the cellular genome; for example, HeLa cells have integrated HPV-18 DNA (13), and SiHa cells harbor a single, incomplete copy of the HPV-16 genome (63). The cervical carcinoma cell line CaSki contains more than 500 copies of the HPV-16 genome but only one or very few of them
are transcriptionally active (63). One mechanism contributing to transcriptional silence of the majority of the viral genomes in CaSki cells could be CpG methylation of the HPV-16 regulatory region. To analyze the methylation pattern of the HPV-16 control region (LCR) in the two cell lines SiHa and CaSki and in the HPV-16-immortalized keratinocyte cell line HPK IA (4), we used the two methylation-sensitive restriction enzymes EcoRI and HhaI. The EcoRI site at genome position 7453 is part of a 16-bp palindrome with the sequence 5'-CAACGAACTTCG-GTTG-3', which contains a binding site for the virus-encoded E2 protein (15); 5-cytidine methylation at either CpG will inhibit EcoRI digestion (64). Similarly, methylation of the HhaI site at position 7672 will prevent enzyme cleavage. The HhaI site contains one CpG, which is part of the MSPF enhancer region (see Fig. 2 and 5). Genomic DNA was digested with PstI, EcoRI, and HhaI and analyzed by Southern blot using probe A (position 7463–7618; Fig. 6A). An EcoRI/PstI double digestion of SiHa or HPK IA DNA yields a 1.3-kb band (theoretical 1326 bp; Fig. 6B, lanes 1 and 2). In contrast, in CaSki cells the EcoRI digestion is incomplete because in addition the 1776-bp PstI fragment appears. We can exclude unspecific inhibition of the EcoRI enzyme since this experiment was repeated several times with different DNA preparations (data not shown; one example is given in Fig. 6C). The incomplete EcoRI digestion indicates 5-cytidine methylation in CaSki cells at one or both CpG dinucleotides which overlap with the EcoRI site.

When genomic DNA from SiHa and HPK IA cells was digested with PstI, EcoRI, and HhaI, a 219-bp band resulting from EcoRI/HhaI digestion is observed (Fig. 6B, lanes 4 and 5), indicating that the CpG dinucleotide in the HhaI site is not methylated in either cell line. In contrast, when PstI/HhaI-digested CaSki DNA is analyzed with probe B, an incomplete cleavage of the 1776-bp PstI fragment by the CpG methylation-sensitive HhaI enzyme is observed (Fig. 6C, lane 3) leading to two PstI/HhaI fragments (1354 and 1107 bp). This demonstrates that CpG dinucleotides adjacent the EcoRI and within HhaI sites in the viral LCR are methylated. Taken together, these results provide evidence that the control regions of a fraction of the HPV-16 genome copies are methylated in CaSki cells. Together with the observed suppression of the HPV-16 enhancer function by 5-cytidine methylation, this result correlates with the finding that only one of the 500 HPV-16 genome copies was found to be transcriptionally active (63), and it suggests that methylation may be one important factor in suppressing the activity of integrated HPV genomes.

**DISCUSSION**

The transcriptional enhancer from HPV-16 regulates the expression of the viral transforming genes E6 and E7 in an epithelial cell type-specific fashion. We have described here the identification of a regulatory element in a region of the core enhancer with three CpG dinucleotides, which is potential sites for transcriptional repression by DNA methylation. A nuclear protein, which we have called MSPF, binds to this element and recognizes a sequence containing two CpGs in a strictly methylation-sensitive manner. Mutating the two CpGs has a strong negative effect on transcription, suggesting that the MSPF binding site is an essential element of the complete enhancer. A consensus sequence derived from the binding experiments contains both CpG dinucleotides but allows sequence exchanges between the CpGs. Several nuclear proteins that recognize sequences containing a single CpG dinucleotide do not bind the DNA when the CpG is methylated, for example, the adenovirus major late transcription factor (MLTF or USF; 38) or the transcription factor CREB (40). It will be interesting to see whether there are other transcription factors that recognize DNA sequences with two CpGs in a methylation-sensitive fashion. Mutating the CpG dinucleotide in the TEF-1 binding site also decreased the enhancer activity. This effect is probably due to reduced affinity of TEF-1 to the mutant sequence, since a different mutation changing the same sequence to 5'-'ACGACATATTTT', thereby making it identical with the TEF-1 site in fp66 (24), did not change the enhancer activity. The HPV-16 LCR contains additional protein binding sites with CpGs, such as the palindromic consensus 5'-ACCCGANTCCTGATG-3', with two CpG dinucleotides that bind the viral E2 protein. Whether binding of E2 or TEF-1 to its recognition sequence is sensitive to CpG methylation is not known.

Site-specific DNA methylation impairs the activating function of the HPV-16 enhancer; the introduction of 5-methylcytidine into one CpG of the MSPF binding site in the complete

---

enhancer suppresses enhancer activity. In addition, the same modification inhibits enhancer activity of the MSPF binding site, suggesting that this fragment serves as a methylation-sensitive enhancer module. We did not find a repression of the thymidine kinase promoter alone by HhaI methylation, as was reported by others (70) when transfecting C33a cells. This difference may be due to the different vector that we have used, pBLCAT5, which was constructed by deleting a fragment upstream of the thymidine kinase promoter and replacing it with polyadenylation signals (50). Repression of gene transcription by CpG methylation has been demonstrated for other systems, such as the human globin genes (34, 65); however, the mechanism responsible for this repression is unclear. The inhibition of binding of positively acting transcription factors is likely to contribute to repression, but alternative mechanisms involving proteins binding to Cpg-methylated DNA have been proposed (45, 46).

Based on our results and previously published data, we present a model for the regulation of the HPV-16 enhancer by CpG methylation (Fig. 7A). The region in the enhancer containing two CpG dinucleotides in fp5e in its unmethylated state is complexed to MSPF. 5-C methylation at either of the two CpG dinucleotides prevents binding of MSPF to fp5e, which causes a repression of viral early transcription (Fig. 7A). We cannot rule out the possibility that MSPF is replaced by another protein when the MSPF binding site is methylated, but our data do not support this idea. One may speculate that both mechanisms, displacement of positively acting factors and binding of repressors, contribute to regulation of gene expression by DNA methylation. In addition, it is possible that other effects, such as alterations in chromatin structure, are also involved and are brought about by changes in the DNA methylation status.

At present, it is an open question as to whether DNA methylation plays a role in the normal life cycle of HPVs when the viral genome is replicated episomally. However, it is intriguing to assume that MSPF plays an important role not only in viral early transcription but also in the regulation of cellular genes, depending on their methylation status. MSPF, like other transcription factors, has been recruited from the transcription machinery of the keratinocyte for HPV-16 gene expression. In Fig.
not observed an increase in the E6/E7 mRNA level in CaSki cells. It has been shown previously that integrated HPV-18 genome copies were methylated in nontumorigenic HeLa X fibroblast hybrid cells but unmethylated in the parental HeLa cells or in tumorigenic hybrids (31). The methylation status of the HPV-16 genome copies in CaSki cells has also been investigated recently by Rösel and co-workers (70), confirming our results of HhaI site methylation and response to 5-azacytidine. The same authors also show that the HPV-18 early promoter can be suppressed by CpG methylation (70). The finding that HPV DNA can be methylated in cell lines is consistent with similar observations that were obtained with other DNA viruses; for example, herpesvirus saimiri DNA was found to be methylated in lymphoid cell lines (41), and the integration of adenovirus type 12 DNA into the cellular genome is followed by de novo methylation of the viral DNA (30).

The propagation of human papillomaviruses is restricted to epithelial cells, where gene expression and genome replication are linked to cellular differentiation. The mechanism underlying this cell type specificity is not known, but the enhancer regulating HPV-16 early gene expression is active exclusively in cell types of epithelial origin; this specificity is also observed with a 90-bp enhancer core (23, 28). However, at least three of the enhancer modules that constitute the core bind nuclear proteins that show only a limited or no cell type specificity: the glucocorticoid receptor and the factors NF-1 and AP-1. In addition, the nuclear factor MSPF that we have described here shows no epithelial cell specificity, but the highest MSPF activity is observed in the cervical carcinoma cell line SiHa. The strong down-effect in transcription which is observed with single site mutants of the complete enhancer suggests a synergistic action of the factors binding to these sites. As one possibility, cell type-specific factors may exist which are not detectable in binding studies and which coordinate the action of the factors bound to the enhancer. Methylation-sensitive transcription factors are likely to be involved in cellular gene expression depending on cell type or differentiation (33), and one can speculate that they are an integral part of a regulatory hierarchy; the methylation status of a cell type-specific gene may be controlled from a higher level in the hierarchy and will determine whether the regulatory region is bound by methylation-sensitive factors such as MSPF. Thereby methylation could determine whether a gene is expressed or not. According to this model, the methylation-sensitive factors need not be cell type-specific but are involved in cell type-specific regulation of transcription due to their position in the hierarchy. It will be interesting to see how factors such as MSPF and other methylation-sensitive DNA-binding proteins are involved in cell type and methylation-regulated expression of cellular genes.

Acknowledgments-We thank Drs. H. zur Hausen, M. Durst, and M. von Knobel Doberitz for cell lines; K. Bleuel, M. Glos, C. Klebe, G. Langst, A. Löh, T. Rothmann, T. Sanchez, C. Schirmer, V. Senner, and E. Steczer for assistance; G. Kelsey for critically reading the manuscript; F. Weih for discussion; and H. G. Gassen for generous support.

REFERENCES
Acad. Sci. U. S. A. 80, 3612–3615
3. Bossart, M., Gissmann, L., Ikenberg, H., Kleinheinz, A., Schurzen, W., and
4. Durst, M., Dziarska-Petreszuk, R. T., Boukamp, P., Fussing, N. E., and
Acad. Sci. U. S. A. 85, 7109–7113
EMBO J. 7, 1815–1820
Cancer Res. 48, 3780–3786

B ATCGGNNNNCCTT

CTATGCGCCCAACCGCTTA
CTCTGGCGATGCGCGGT
TATGCGCGCAACCGCTCA
GTCGCGCGCAACCGCTT
TGAGCGCGCAACCGCTTA
GGTTGGCGCCACGGCTGA

MSPF consensus
HPV-16 enhancer
rat TAT – 2.5 kb GRU
mouse Thy-1.2 promoter
human EGF-R promoter
human Raf-1 promoter
HTLV-I – 1 bp repeat

Fig. 7. Regulation of the HPV-16 enhancer by DNA methylation. Panel A, protein-DNA complexes in the HPV-16 enhancer core are shown which have been described here or published previously (19, 21, 26–28, 60). The f5pE-L region in its unmethylated state is bound by the nuclear factor MSPF, which is an essential factor for the enhancer function. 5-C methylation of the two CpG dinucleotides inhibits binding of MSPF and thereby represses enhancer activity. Panel B, sequences related to the HPV-16 MSPF binding site. Potential MSPF binding sites are listed which show homology to the MSPF site in the HPV-16 enhancer (underlined). The rat TAT MSPF site is located next to a GRE at position -2403 through -2428 in a DNase I-hypersensitive region which binds several nuclear proteins (46, 57). The other motifs are from the mouse Thy-1 gene promoter, position +24 through +7 (66); the human EGF receptor promoter, position: -93 through -76 (67); the human Raf-1 promoter, position +58 through +75 (68); and the human T-cell leukemia virus type 1 (HTLV-I) long terminal repeat promoter, imperfect 21-bp repeat, position -104 through -87 (69). EGF-R, epidermal growth factor receptor.

7B the MSPF binding sites from HPV-16 and from the rat TAT gene are compared with four potential MSPF sites from the following regulatory regions: the mouse Thy-1 gene promoter (66); the promoter of the human epidermal growth factor receptor gene (67); the human Raf-1 promoter (68); and the 21-bp repeat in the human T-cell leukemia virus type 1 long terminal repeat promoter (69). It should be noted that both the HPV-16 and the TAT MSPF sites are located in close vicinity of a GRE (19, 57); however, the functional significance of this region for TAT gene expression is unclear (49). The homologous sequence in the Thy-1 gene is located directly adjacent to the CAP site of the promoter (position +24 through +7) where it is part of a region important for Thy-1 promoter activity (66). Interestingly, both the Thy-1 and the Raf-1 promoter are GC-rich and are located in regions with a high frequency of CpG dinucleotides (HpaII tiny fragment islands) (66, 68). Therefore it should be rewarding to test whether these sequences bind MSPF and whether transcription of both genes responds to the methylation status of the promoter.

In the cervical carcinoma cell line CaSki, the majority of the integrated HPV-16 genomes is methylated at CpG dinucleotides in the viral control region. One of these CpGs overlaps with a binding site for the viral E2 protein, and a second CpG is in the MSPF binding site. Interestingly, all but one of the several hundred HPV-16 genome copies are thought to be transcriptionally inactive (63), and it is conceivable that CpG methylation of the viral LCR is one mechanism responsible for suppression of transcription. However, in experiments involving 5-azacytidine designed to investigate whether demethylation of the HPV-16 genomes can activate transcription, we have...
Methylation-sensitive HPV-16 Enhancer