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

Felix Hoppe-Seyler* and Karin Butz

Forschungsschwerpunkt Angewandte Tumorvirologie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 242, D-6900 Heidelberg, Germany

Gene expression of the cancer-associated human papillomavirus (HPV) type 18 is modulated by cis-regulatory elements located within the viral upstream regulatory region (URR). All cellular factors identified so far involved in viral gene control bind to the 3' portion of the HPV-18 URR. In contrast, very little is known about regulatory elements within the 5' portion of the URR. We therefore analysed this region of unknown function to delineate potential cis-regulatory elements contained therein. By utilizing transient expression assays, an 84 bp fragment could be identified in the 5' portion of the URR that exhibits orientation-independent cis-stimulatory activity in HeLa cervical carcinoma cells and primary human fibroblasts. Gel retardation assays and competition experiments indicated specific binding of cellular proteins to the 84 bp fragment. By functional dissection, a regulatory element with intrinsic cis-stimulatory activity could be mapped within the 84 bp fragment. Binding studies indicate that this cis-stimulatory element contains a sequence-aberrant Sp1 recognition site. Transient luciferase assays performed with mutated templates demonstrate that this Sp1 binding site behaves as a functional Sp1 element in vivo and is a main determinant of the cis-stimulatory activity exerted by the 84 bp fragment. These data show that the 5' portion of the HPV-18 URR contains cis-activating elements and indicate an important functional role for the cellular transcription factor Sp1 in their regulation.

Human papillomaviruses (HPVs) are small DNA viruses and causative agents of various benign epithelial proliferative disorders. During the natural course of viral infection, gene expression and replication of HPVs are tightly controlled by the differentiation status of the epithelial host cell (Dürst et al., 1991; Stoler et al., 1990). Certain types of these viruses are also closely associated with the development of human anogenital malignancies. HPV-16 and -18 DNA has been found in over 70% of cervical carcinomas, usually integrated into host cell chromosomes (zur Hausen, 1989). Although integration is often accompanied by deletion of large portions of the viral DNA, the transforming (Bedell et al., 1989; Münger et al., 1990; Storey et al., 1988) and growth-promoting (von Knebel Doeberitz et al., 1988) genes E6 and E7 are normally preserved and transcribed in carcinomatous tissue as well as in tumour-derived cell lines (Schwarz et al., 1985).

HPV-18 E6/E7 gene expression is controlled by a transcriptional control region, the upstream regulatory region (URR), which is located within the viral genome between the end of the late open reading frame (ORF) L1 and the start of the E6 early gene coding sequence (Fig. 1). The 825 bp long HPV-18 URR has been shown to interact with a variety of DNA-binding proteins, including the cellular transcription factors AP1, NF1, KRF, Oct-1, Sp1 and the glucocorticoid receptor (Garcia-Carranca et al., 1988; Gloss et al., 1989; Mack & Laimins, 1991; Nakshatri et al., 1990; Hoppe-Seyler & Butz, 1993). All these factors so far identified bind to the 436 bp 3' portion of the viral URR, which encompasses the viral constitutive enhancer and the E6/E7 promoter region (Fig. 1).

Until now, however, there have been very little data about regulatory elements within the 389 bp long 5' portion of the URR, which constitutes almost half of the complete HPV-18 URR. Four sites of protein–DNA interactions have been detected by footprint analysis in this part of the URR, caused by cellular factors not yet known (Garcia-Carranca et al., 1988). The functional significance of these contacts is not yet clear.

To investigate whether the 389 bp 5' terminal RsaI–RsaI fragment of the HPV-18 URR possesses direct enhancer activity, we cloned it in sense and antisense orientations upstream of a truncated herpes simplex virus (HSV) thymidine kinase (TK) promoter (nucleotides −79 to 0; numbering according to McKnight & Kingsbury, 1982) into the luciferase reporter plasmid.
Neither the 5'-terminal 133 bp into plasmid pBtk*L and were tested for their individual negative regulatory influences from flanking sequences in HeLa cells and primary human fibroblasts (Fig. 1).

Enhancer activity employing transient luciferase assays. Alternatively, potential cis-stimulatory elements, although present, might be functionally silent either due to an over long distance to the test promoter or due to the absence of enhancer elements within the 389 bp region.

The complete 389 bp 5'-URR and subfragments derived therefrom (nucleotide positions as indicated) were cloned in sense (s) or antisense (as) orientations upstream of the truncated HSV TK promoter into luciferase reporter plasmid pBtk*L. Regions of protein–DNA interactions within the 389 bp 5' URR (Fp1 to Fp4) as assessed by footprint analysis (Garcia-Carranca et al., 1988) are indicated. Relative luciferase activities (fold activation above pBtk*L) after correction for variations in transfection efficiency by cotransfected pAc-Gal (Hoppe-Seyler et al., 1991a) in HeLa cells and primary human fibroblasts (Fib) are shown in the columns on the right.

Footprint analysis of the complete HPV-18 URR (Garcia-Carranca et al., 1988) suggested two sites of protein–DNA interactions, designated Fp2 and Fp3, within the 84 bp MaeIII–MaeIII subregion (Fig. 1). Thus Fp2 and Fp3 might represent the regulatory modules contributing to the transcriptional stimulatory activity exerted by the 84 bp fragment. To assess the functional role of these sequences individually, synthetic oligonucleotides corresponding to Fp2 (HPV-18, positions 7252 to 7286) and Fp3 (7306 to 7333) (see Fig. 3 for nucleotide sequences) were subcloned as monomeric or trimeric units in enhancer configuration upstream of the HSV TK promoter into reporter plasmid pBLCAT2 (Luckow & Schütz, 1987) and were tested in HeLa cells and primary human fibroblasts. Fp2 did not exhibit significant cis-stimulatory activity, neither as a monomer nor as a trimer (Fig. 2a). In contrast, Fp3 activated the TK promoter weakly, but reproducibly, as a monomer, whereas trimerization of the sequence led to strong enhancer activity in both cell types (Fig. 2b). These findings localize an intrinsic cis-stimulatory potential to the Fp3 module.

In the DNA sequence of Fp2 we did not notice any obvious consensus sequence which could serve as a recognition site for known transcription factors. To visualize the factors binding to Fp2, gel retardation assays were performed using a HeLa cell nuclear extract. A single shifted complex A (Fig. 2a, lane 2) could be detected, which was efficiently competed for by a 100-fold molar excess of homologous unlabelled oligonucleotide Fp2 (lane 3), but not by a 100-fold molar excess of a heterologous oligonucleotide (lane 4) containing an AP1 recognition site (AP1E, see Fig. 3), thus demonstrating the binding specificity for the detected protein complex to sequences within Fp2.

Close inspection of the DNA sequences contained within the Fp3 module revealed two motifs reminiscent of consensus sequences for the binding of known transcription factors (Fig. 3): a central GCGCCC motif, carrying a single mismatch when compared to the

Fig. 1. Functional dissection of the 5' URR in HPV-18. The HPV-18 genome is shown schematically at the top. Below, the 389 bp 5' portion (dashed area), the constitutive enhancer (open area) and the E2-responsive subregion (stippled area) within the HPV-18 URR are indicated. The complete 389 bp 5'-URR and subfragments derived therefrom (nucleotide positions as indicated) were cloned in sense (s) or antisense (as) orientations upstream of the truncated HSV TK promoter into luciferase reporter plasmid pBtk*L. Regions of protein–DNA interactions within the 389 bp 5' URR (Fp1 to Fp4) as assessed by footprint analysis (Garcia-Carranca et al., 1988) are indicated. Relative luciferase activities (fold activation above pBtk*L) after correction for variations in transfection efficiency by cotransfected pAc-Gal (Hoppe-Seyler et al., 1991a) in HeLa cells and primary human fibroblasts (Fib) are shown in the columns on the right.

Lack of enhancer activity could be explained by the absence of enhancer elements within the 389 bp region. Alternatively, potential cis-stimulatory elements, although present, might be functionally silent either due to an over long distance to the test promoter or due to negative regulatory influences from flanking sequences in the context of the complete 389 bp fragment. Therefore we dissected the 389 bp Rsal–Rsal fragment by MaeIII digestion into three distinct subregions of 133 bp, 84 bp and 182 bp, respectively (Fig. 1). The corresponding subregions were cloned in sense or antisense orientations into plasmid pBtk*L and were tested for their individual enhancer activity employing transient luciferase assays (for conditions see Hoppe-Seyler et al., 1991b) in both HeLa cells and primary human fibroblasts (Fig. 1). Neither the 5'-terminal 133 bp Rsal–MaeIII fragment (p133s/tk*L and p133as/tk*L) nor the 3'-terminal 182 bp MaeIII–Rsai (p182s/tk*L and p182as/tk*L) fragment significantly stimulated luciferase expression in either orientation. However, the centrally located 84 bp MaeIII–MaeIII fragment (p84s/tk*L and p84as/tk*L) led to a five- to sevenfold stimulation of luciferase activities when tested both in sense and antisense configurations. Dimerization of the 84 bp fragment led to a more than additive increase in the cis-stimulatory potential, resulting in an approximately 30-fold enhancement of luciferase activities in both cell types (Fig. 1). These results indicate the presence of a hitherto unidentified cis-activating element within the 5' portion of the HPV-18 URR, which is active both in epithelial and non-epithelial cells.

Enhancer activity employing transient luciferase assays. Alternatively, potential cis-stimulatory elements, although present, might be functionally silent either due to an over long distance to the test promoter or due to negative regulatory influences from flanking sequences in the context of the complete 389 bp fragment. Therefore we dissected the 389 bp Rsal–Rsal fragment by MaeIII digestion into three distinct subregions of 133 bp, 84 bp and 182 bp, respectively (Fig. 1). The corresponding subregions were cloned in sense or antisense orientations into plasmid pBtk*L and were tested for their individual enhancer activity employing transient luciferase assays (for conditions see Hoppe-Seyler et al., 1991b) in both HeLa cells and primary human fibroblasts (Fig. 1). Neither the 5'-terminal 133 bp Rsal–MaeIII fragment (p133s/tk*L and p133as/tk*L) nor the 3'-terminal 182 bp MaeIII–Rsai (p182s/tk*L and p182as/tk*L) fragment significantly stimulated luciferase expression in either orientation. However, the centrally located 84 bp MaeIII–MaeIII fragment (p84s/tk*L and p84as/tk*L) led to a five- to sevenfold stimulation of luciferase activities when tested both in sense and antisense configurations. Dimerization of the 84 bp fragment led to a more than additive increase in the cis-stimulatory potential, resulting in an approximately 30-fold enhancement of luciferase activities in both cell types (Fig. 1). These results indicate the presence of a hitherto unidentified cis-activating element within the 5' portion of the HPV-18 URR, which is active both in epithelial and non-epithelial cells.

Footprint analysis of the complete HPV-18 URR (Garcia-Carranca et al., 1988) suggested two sites of protein–DNA interactions, designated Fp2 and Fp3, within the 84 bp MaeIII–MaeIII subregion (Fig. 1). Thus Fp2 and Fp3 might represent the regulatory modules contributing to the transcriptional stimulatory activity exerted by the 84 bp fragment. To assess the functional role of these sequences individually, synthetic oligonucleotides corresponding to Fp2 (HPV-18, positions 7252 to 7286) and Fp3 (7306 to 7333) (see Fig. 3 for nucleotide sequences) were subcloned as monomeric or trimeric units in enhancer configuration upstream of the HSV TK promoter into reporter plasmid pBLCAT2 (Luckow & Schütz, 1987) and were tested in HeLa cells and primary human fibroblasts. Fp2 did not exhibit significant cis-stimulatory activity, neither as a monomer nor as a trimer (Fig. 2a). In contrast, Fp3 activated the TK promoter weakly, but reproducibly, as a monomer, whereas trimerization of the sequence led to strong enhancer activity in both cell types (Fig. 2b). These findings localize an intrinsic cis-stimulatory potential to the Fp3 module.

In the DNA sequence of Fp2 we did not notice any obvious consensus sequence which could serve as a recognition site for known transcription factors. To visualize the factors binding to Fp2, gel retardation assays were performed using a HeLa cell nuclear extract. A single shifted complex A (Fig. 2a, lane 2) could be detected, which was efficiently competed for by a 100-fold molar excess of homologous unlabelled oligonucleotide Fp2 (lane 3), but not by a 100-fold molar excess of a heterologous oligonucleotide (lane 4) containing an AP1 recognition site (AP1E, see Fig. 3), thus demonstrating the binding specificity for the detected protein complex to sequences within Fp2.

Close inspection of the DNA sequences contained within the Fp3 module revealed two motifs reminiscent of consensus sequences for the binding of known transcription factors (Fig. 3): a central GCGCCC motif, carrying a single mismatch when compared to the
consensus binding site for the cellular transcription factor Sp1 (CCGCCC) (Briggs et al., 1986) and a 3'-terminal located sequence TGAGTAA, exhibiting a single difference from the AP1 recognition consensus (TGAGTCA) (Angel et al., 1987). To investigate whether one of these factors binds to Fp3, gel shift competition assays were performed. Incubation of HeLa cell crude nuclear extract with a radiolabelled Fp3 probe led to the formation of a specific DNA binding complex B (Fig. 2b, lane 2), which was efficiently competed for by a 100-fold molar excess of homologous unlabelled oligonucleotide Fp3 (lane 3), but not by the heterologous oligonucleotide AP1E (lane 4). These competition data indicate that complex B is formed by a protein specifically binding to a motif within the DNA sequence of Fp3 and suggest that this protein is not AP1. Importantly, the formation of complex B could be efficiently prevented by a 100-fold molar excess of oligonucleotide Sp1K (lane 5) containing a well defined consensus Sp1-binding site (Westin & Schaffner, 1988, for sequence see Fig. 3). This indicates that Fp3 represents a recognition site for the cellular transcription factor Sp1. The second, faster migrating complex observed with Fp3 (lane 2) shows the same binding specificity and might represent a degradation product of Sp1 or a different modification of the protein (Jackson & Tjian, 1988). The conclusion that Fp3 contains a bona fide Sp1 binding site is supported experimentally by the observation that Fp3 is specifically
contains a genuine Spl binding site is derived from gel crude nuclear extract. In contrast, Fp3 is not bound by complex migrating similarly to complex B obtained with the HPV-18 constitutive enhancer.

consensus Spl binding site (Westin & Schaffner, 1988) contains the Sp 1 recognition sequence contained therein and binding assays. Oligonucleotides Fp2 and Fp3 are derived from the HSV IE-3 gene promoter and contains a well defined API binding element (Garcia-Carranca et al., 1988) and is derived from the HPV-18 constitutive enhancer.

bound by purified human Sp1 protein (Stratagene) (Fig. 2b, lane 7), leading to the formation of a protein DNA complex migrating similarly to complex B obtained with crude nuclear extract. In contrast, Fp3 is not bound by the purified AP2 protein (not shown), which also exhibits binding affinity to GC-rich sequences.

A specific mutation introduced into the Sp1-binding motif, resulting in oligonucleotide Fp3M (Fig. 3), abolished both the interaction of purified Sp1 with the fragment and the formation of HeLa complex B (Fig. 2b, lanes 8 and 6, respectively). This indicates that both the purified Sp1 protein and complex B from HeLa crude nuclear extract contact the oligonucleotide probe at the Sp1-like recognition motif. Further evidence that Fp3 contains a genuine Sp1 binding site is derived from gel shift competition experiments in which unlabelled Fp3 oligonucleotide efficiently competed for the binding of Sp1 to its well defined consensus sequence within the HSV (immediate early 3) IE-3 gene promoter (data not shown).

To assess the functional significance of the Sp1–DNA interaction for the regulatory activity of the 84 bp fragment, the Sp1 recognition sequence contained therein was specifically mutated by a polymerase chain reaction-based protocol (Higuchi, 1990) using Fp3M (Fig. 3) as a primer. Introduction of this mutation (Fig. 4a) led to a strong decrease of the cis-stimulatory activity exerted by the 84 bp fragment when tested in vivo employing transient luciferase assays in HeLa cells and primary human fibroblasts (Fig. 4b; compare p84sM/tk*L with p84sM/tk*L). The correlation between the loss of binding in vitro (Fig. 2b) and the loss of cis-activating function in vivo (Fig. 4) strongly suggests that the sequence-aberrant Sp1 recognition motif within Fp3 represents a functional Sp1 element and that the cellular transcription factor Sp1 significantly contributes to the activation potential of the 84 bp region.

This work identifies an 84 bp cis-regulatory element within the 5′-terminal portion of the HPV-18 URR, exhibiting both specific protein–DNA interactions and significant cis-activation potential. The fragment shows features of a true enhancer region, since its activity is orientation-independent and dimerization leads to a synergistic generation of cis-stimulatory activity, both findings being characteristic of genuine enhancer elements (Fromenthal et al., 1988).

Analysis of the 84 bp fragment employing gel retardation assays confirmed the existence of two sites of specific DNA–protein interaction, designated Fp2 and Fp3, as suggested previously by footprint analysis (Garcia-Carranca et al., 1988). When tested individually for their enhancer activity in transient assays, Fp3 exhibited intrinsic stimulatory potential, indicating that it can function as a cis-activatory element. In contrast, no stimulatory activity could be detected for Fp2, suggesting that it either does not possess cis-activation potential in the cells tested or that it might require an heterologous cis-regulatory element for stimulatory function.

The observed lack of cis-stimulatory activity for the 84 bp fragment in the context of the complete 389 bp 5′ URR could be explained in two ways. First, it is possible that the regulatory activity of the 84 bp fragment is

![Fig. 4. (a) Luciferase test constructs carrying the wild-type 84 bp sequence (p84sM/tk*L) and the mutated fragment bearing a 4 bp change (as indicated) within the Sp1 recognition site of Fp3 which abolishes Sp1 binding (see Fig. 2b). (b) Relative luciferase activity (fold activation above pBtk*L of the constructs after transient transfection into HeLa cells (stippled bars) and primary fibroblasts (solid bars).](image-url)
dependent on the distance to the test promoter. This possibility could provide an explanation for the lack of cis-stimulatory activity measured for constructs bearing the whole 389 bp fragment, in which the 84 bp element is approximately 250 or 200 bp removed from the TK promoter (Fig. 1, p389s/tk*L and p389as/tk*L, respectively) as compared to plasmids containing the 84 bp element at a distance of approximately 70 bp to the test promoter (Fig. 1, p84s/tk*L and p84as/tk*L, respectively). Second, we cannot rule out the possibility that the 84 bp enhancer might underlie influences from negative regulatory elements within flanking sequences, which have also been speculated to be present in the 5' URR of HPV-11 (Auborn & Steinberg, 1991). Further work is required to distinguish these possibilities.

Binding studies, competition experiments and mutational analyses reveal that Fp3 contains a sequence-aberrant Sp1 binding element. Functionally, the Fp3 module behaves like a typical Sp1 element, showing modest transcriptional stimulation as a monomer but exhibiting synergistic activation after trimerization (Pascal & Tjian, 1991). Although we cannot rule out the formal possibility that the factor binding to the Sp1 motif in vitro is not equivalent to the factor that functions in vivo, our experimental results strongly suggest that the cellular transcription factor Sp1 is a main determinant of the cis-stimulatory activity exerted by the 84 bp fragment.

The HPV-18 E6/E7 promoter p105, which is located at the 3' end of the viral URR, represents the only promoter of HPV-18 described thus far (Thierry et al., 1987). In transient transfection experiments using deletion mutants, the 5'-terminal portion of the HPV-18 URR contributes only about 20% to the transcriptional enhancement of p105 by the complete viral URR (Hoppe-Seyler et al., 1991; Thierry et al., 1987a). Thus, the presence of the cis-stimulatory 84 bp element within the 5' portion of the viral URR is likely to contribute only marginally, if at all, to the extent of E6/E7 gene expression, suggesting that the primary function of these sequences may not result in the transcriptional activation of the E6/E7 promoter.

Interestingly, a constitutive enhancer element has also been recently identified within the 5'-terminal portion of the bovine papillomavirus type 1 (BPV-1) URR (Vande Pol & Howley, 1990). Analogous to the HPV-18 84 bp region described in the present study, an Sp1 element is essential for the transcriptional activity of this enhancer. Subsequent work showed that the 5' enhancer within the BPV-1 URR is important for both transformation and replication of the viral genome (Vande Pol & Howley, 1992). From these findings, one could speculate that the 84 bp enhancer element within the corresponding 5' portion of the HPV-18 URR might play a role in viral replication. Alternatively, these sequences could function as a transcriptional activatory element for an uncharacterized upstream viral promoter. As the major late promoter of BPV-1 (Baker & Howley, 1987) is located within the 5' portion of the URR, it is tempting to speculate that promoter elements for HPV late gene expression could be present in the 5' part of their URR. Thus, the 84 bp enhancer might be involved alternatively in the transcriptional regulation of HPV structural genes.

The cellular transcription factor Sp1 appears to be involved in several aspects of the biology of papillomaviruses. In addition to its important role for the activity of the BPV-1 constitutive enhancer, an Sp1-binding site has been shown to be essential for the basic activity of the BPV-1 p2443 promoter, which controls the expression of the BPV-1 E2 and the E5 ORFs (Spalholz et al., 1991). Recent reports indicated a functional cooperativity between Sp1 and the BPV-1 E2 trans-activator protein (Ham et al., 1991; Li et al., 1991), suggesting that the interaction of viral factors with Sp1 might play an important role in the transcriptional control of papillomaviruses. In addition, it has been shown that an Sp1 element activates the HPV-16 E6/E7 promoter (Gloss & Bernard, 1990) and is negatively regulated through the displacement of Sp1 by the binding of the viral E2 protein to an overlapping recognition sequence (Tan et al., 1992). It should be noted that a cis-active Sp1 binding site indispensable for efficient stimulation of the HPV-18 E6/E7 promoter is also located within the corresponding 5' portion of the HPV-18 URR (Hoppe-Seyler & Butz, 1993). The definition in the present study of an Sp1 element essential for the cis-activity of the 84 bp region within the 5' portion of the HPV-18 URR is consistent with the notion that Sp1 plays an important role in the life cycle of papillomaviruses and indicates a regulatory role for this factor in the control of HPV-18 gene expression beyond the modulation of E6/E7 transcription.

We thank Dr. H. zur Hausen for continuous support, A. Ullmann for excellent technical assistance, Drs. E. Schwarz, M. von Knebel Doeberitz and H. zur Hausen for critical reading of the manuscript, Dr. H. Delius for oligonucleotide synthesis, Dr. E. M. de Villiers for primary human fibroblasts and Dr. A. Klotzibucher and M. Rentrop for luciferase vectors pBL and pBk*I.

References


(Received 5 August 1992; Accepted 5 October 1992)