Short communication

Functional promoters in the genome of human papillomavirus type 6b

S. Karlen, E. A. Offord and P. Beard*

Swiss Institute for Experimental Cancer Research, CH-1066 Epalinges, Switzerland

Viral mRNAs from lesions containing human papillomavirus type 6 (HPV-6) have previously been mapped on the viral DNA but relatively little is known about the control of mRNA production, or whether the mapped RNA termini correspond to promoters. By analysis of run-off transcripts synthesized \textit{in vitro}, primer extension and measurements of promoter activity in fragments of the viral DNA introduced into cells, we have identified three promoters in the early region of the HPV-6b genome. These are: (i) at the end of the long control region upstream of the E6 open reading frame; (ii) upstream of E7 and (iii) upstream of E1. The promoter upstream of E1 was the most active. These results contrast with results of similar assays with HPV-18, in which the strongest promoter was that controlling expression of the transforming genes E6 and E7. In addition, a novel promoter was detected close to E5a, upstream of the late genes.

Human papillomaviruses (HPVs) are associated with both benign and malignant lesions of the skin or mucosal epithelia. Among the viruses affecting genital tissues, HPVs 6 and 11 are most frequently found in benign condylomata acuminata, which only occasionally progress to cancer, whereas other types (notably including HPV-16 and HPV-18) are often detected in cervical carcinomas (Howley, 1990). Therefore, infections with HPV-6 and HPV-11 are considered low risk and infections with HPV-16 and HPV-18 are considered high risk. A potentially important difference between low risk and high risk HPV types is in the manner in which transcription of the viral genes, in particular the transforming genes E6 and E7, is regulated (Smotkin \textit{et al.}, 1989). The HPV-6-specific transcripts in genital condylomata acuminata have been mapped (Smotkin \textit{et al.}, 1989; Chow \textit{et al.}, 1987). The great majority of the RNAs are derived from the early region of the genome, most abundant species being a spliced transcript encoding part of the E1 and the E4 open reading frames (ORFs). The 5'-ends of these abundant transcripts were found in the E7 ORF near the boundary with E1 (Chow \textit{et al.}, 1987). Other RNA species had their 5'-ends just upstream of the first ATG of the E6 ORF and within the E6 coding sequences (Smotkin \textit{et al.}, 1989). RNA mapping studies have also indicated that rare mRNA species could be initiated at the end of the early region, just upstream of the L2 ORF (Chow \textit{et al.}, 1987). This RNA could, however, have been a result of degradation or processing of transcripts initiated in the early region.

Although these HPV-6 mRNAs have been mapped, little is known about the promoters that control their synthesis. It was not clear whether the RNA ends which were mapped correspond to promoters or to sites of processing or cleavage. Most studies on the regulation of gene expression in HPVs have focused on the promoter for the early genes which is located near the junction of the long control region (LCR) and the E6 ORF. In the work reported here we mapped promoters active in the genome of HPV-6b, \textit{in vitro} and in transfected cells. We compared their locations and relative activities with the promoters detected using similar methods in the DNA of the high risk HPV-18 (Karlen & Beard, 1993).

In order to screen for transcriptional promoters, the HPV-6b genome was divided into two parts. The viral DNA (a kind gift of Prof H. zur Hausen, German Cancer Research Center, Heidelberg) cloned into the vector pBR322 at the \textit{Bam}HI site was digested with \textit{EcoR}I which cuts once in the vector and once in the HPV-6b sequence at nucleotide (nt) 2188 (Schwarz \textit{et al.}, 1983), thus yielding two fragments (Fig. 1a). The larger fragment containing most of L2, all of L1, the LCR, E6, E7 and most of E1, together with vector sequences, was isolated and used as direct template for transcription reactions (template p6bL). The smaller fragment, containing E2, E4, E5a, E5b, part of L2 and a small piece of pBR322, was isolated to give the template p6bS.

The p6bS template was tested for transcriptional activity \textit{in vitro} using a whole-cell extract of HeLa cells.

* Author for correspondence. Fax +41 21 652 69 33.

0001-3509 © 1996 SGM
RNA transcripts were synthesized in the presence of $[^{32}P]UTP$ and the RNAs produced were denatured with glyoxal and analysed by gel electrophoresis (Tack & Beard, 1985). Two DNA templates were used in the assay: the entire p6bS fragment and p6bS shortened by cleaving with NcoI (p6bS/NcoI) (see Fig. 1a). Transcription from both templates led to comparable conclusions and Fig. 1(b) shows the results of a typical transcription reaction with the p6bS/NcoI template. Lane 1 shows a main transcript of around 620 nt and an additional band about 10-fold fainter of 1150 nt. This suggests that an RNA initiation site is present at the end of the early region, located around position 3800 just in front of E5a and possibly a second weaker one around position 3270. The mapping of these RNA start sites was confirmed by primer extension. This is shown for the stronger promoter at nt 3800 in Fig. 1(c). RNA produced from the p6bS/NcoI template was hybridized to the primer 6b-3977 and extended by reverse transcriptase. The results show that the 5' termini of the extended...
products mapped to two sites close together at positions 3770 and 3772, thus confirming the presence and orientation of a promoter at this position.

To identify promoters in the early region of HPV-6b, the p6bL template was transcribed and the RNA products analysed as above. A strong run-off transcript of about 1510 nt was observed (Fig. 2a), which would correspond to an RNA start site around nt 680. Two weaker transcripts of approximately 1900 and 2100 nt were also detected (the transcripts are shown by arrows). The starts of these two transcripts were mapped to around nt 290 and nt 90, respectively. Above these bands there was a heterogeneous smear of RNA. Although this RNA could have initiated in the upstream part of the LCR or in L1, discrete 5'-ends were not detected by subsequent primer extension mapping so these RNAs have not yet been characterized further. None of these transcripts was seen when α-amanitin (2.5 μg/ml) was present in the reaction (lane 2), confirming that they were synthesized by RNA polymerase II. Only bands near the bottom of the gel, which are believed to be end-labelled small RNAs, were unaffected by α-amanitin.

To confirm these initiation sites and map them more precisely, the RNA synthesized from p6bL was hybridized to three different oligonucleotides complementary to the HPV-6b sequence and analysed by primer extension (Fig. 2b). Reactions carried out with the oligonucleotide 6b-834 (5' end at nt 834) gave two main extended products (162 and 146 nt in length) with termini mapping to nt 672 and 688, respectively. These results indicated the presence of a strong promoter within the E7 coding sequences. When the oligonucleotides 6b-226 and 6b-627 were used in the primer extension reactions, two additional RNA initiation sites were mapped to positions 80 (just upstream of E6) and about 300 (within E6), respectively.

The analyses by run-off transcription and by primer extension therefore agreed fairly well as to the locations
of promoters active in vitro in the E6–E7 region of plasmid p6bL. Of these promoters, the one around nt 680 in the E7 sequence was the strongest.

To further establish the activities of the promoters identified within E7 and near E5a and to test whether they function in vivo, DNA fragments were isolated from HPV-6b and cloned into the promoterless chloramphenicol acetyltransferase (CAT) reporter vector pBLCAT3. Two constructs were found to be active after transfection into HeLa cells, as indicated by the appearance of acetylated forms of chloramphenicol. The first construct contained the NsiI (nt 534)–RsaI (nt 742) restriction fragment of HPV-6b and was referred to as HPV-6-742CAT. The second construct, HPV-6-3880CAT, contained the viral sequence from the NsiI site at position 2955 to the RsaI site at nt 3880. Several other constructs, including nt 242–534 (upstream of the E7 promoter), 6652–7270 (L1 LCR junction) and 2955–3916 (extending further downstream from the E5 promoter, past an ATG codon) gave no detectable CAT activity over background. These assays were repeated several times and representative results are shown in Fig. 3. By examination of the autoradiograms and quantification of the acetylation, the two HPV-6 promoters were reproducibly less active than the SV40 promoter in the control pSV2CAT.

The positions of the three promoters in the early region of the HPV-6b genome that we detected in these experiments correspond well with 5′-ends of RNAs previously mapped in vivo. We conclude that the start sites we identified are important in vivo. The promoter we detected around nt 80, just upstream of the E6 ORF, fits with 5′-ends of mRNA from condylomata acuminata mapped by the R-loop technique (Chow et al., 1987) or S1 analysis (Smotkin et al., 1989). It is therefore probably responsible for synthesis of mRNA for E6 and possibly,
following splicing, for other viral proteins downstream. Analysis of the DNA sequence shows that this promoter contains a TATA homology at nt 64 and a CCAAT homology at nt 9. A weak promoter detected in HPV-6g by Wu & Mounts (1988) may correspond to this promoter.

The second promoter activity we detected in the early region (near nt 300) is of interest since it corresponds to the 5'-ends of RNAs encoding the E7 protein (Smotkin et al., 1989). E7 is the main transforming gene in human papillomaviruses. In benign genital lesions containing DNA of low risk HPVs, expression of E7 is low. In contrast, the amount of E7 is greater in malignant lesions containing DNA of high risk HPVs (Smotkin & Wettstein, 1986). These observations have led to the hypothesis that an important difference between low risk and high risk HPV types is in the manner in which E7 is expressed. Our results support this hypothesis. The activity of the promoter we found upstream of the E7 ORF in HPV-6 was relatively weak and this may account for the low level of expression of E7 in benign lesions.

In high risk HPVs, E6 and E7 are transcribed from a single promoter; this primary transcript gives rise to mRNA species in which part of E6 is spliced out. The splice may generate an mRNA from which E7 can be efficiently translated (Smotkin et al., 1989). It has been observed that the spliced version is much more abundant than the full-length transcript, which is present in very small amounts in cancer cells (Schneider-Gädicke & Schwarz, 1986). In contrast, the E6/E7 region of low risk papillomaviruses lacks such splicing and E7 appears to be translated from an mRNA initiated from a weak promoter located within the E6 ORF (Smotkin et al., 1989 and this study). Therefore, high risk and low risk papillomaviruses appear to have different strategies for generating E7 mRNA. Of course this is not the only difference between high and low risk HPV types: the high risk viruses encode an E7 protein which transforms rodent cells, whereas E7 from low risk types is inefficient in similar assays. Therefore, a high level of E7 expression, in combination with the higher potential of E7 (and E6) to transform cells, may account for the greater oncogenic activity of high risk HPV types.

The strongest promoter activity in the early region, as judged from the intensities of the run-off transcript bands, gave rise to RNA initiated around position 680, in the E7 ORF. It is an interesting correlation therefore that the most abundant viral mRNA from genital condylomata starts in E7 and encodes a fusion protein of parts of E1 and E4 (Chow et al., 1987). The E1/E4 fusion protein of HPV-16 interacts with the cytokeratin matrix of the infected cell, promoting its degradation (Doorbar et al., 1991), and its function may be to help the release of the virus at the end of the replication cycle.

Interestingly, we did not find an equivalent RNA initiation site during an analysis of the transcription of the HPV-18 E6-E7-E1 region (Karlen & Beard, 1993). However, the 5'-ends of transcripts have been mapped within E7 of HPV-16 (Rohlfs et al., 1991) and HPV-31b (Hummel et al., 1992). Transcription of these RNAs appears to be induced at later stages of keratinocyte differentiation. In contrast, the mRNA coding for the E1/E4 protein of HPV-6 and HPV-11 is synthesized very early during the differentiation of epithelial cells (Stoler et al., 1989) and we observed that the promoter for this mRNA is active in the in vitro transcription assay. These observations suggest that a major difference between high risk (HPV-16, HPV-18 and HPV-31b) and low risk (HPV-6 and HPV-11) papillomaviruses is in the way in which the production of the E1/E4 mRNA is regulated during host cell differentiation.

Finally, a promoter was found in the region of the E5 ORF, upstream of the late genes. The function of this promoter is not known. It could give rise to an RNA encoding E5. However, it has several characteristics of a late promoter. It does not have an apparent TATA box and in this resembles the late promoter of cottontail rabbit papillomavirus (Wettstein et al., 1987), bovine papillomavirus-1 (Baker & Howley, 1987) and SV40 (Brady et al., 1982), none of which contains the TATA sequence. It has, at a position approximately 30 bp upstream of the initiation site the 11 nt sequence GCAACAAAGGC. This sequence has some homology to the BPV-1 late promoter (Baker & Howley, 1987) and to the SV40 late promoter sequence GTACCTAACC, which has been shown to be important for efficient utilization of the SV40 late start site. In human condylomata acuminata (Chow et al., 1987) viral mRNA apparently initiated near the beginning of the late genes has been identified. Similar RNAs have also been detected in cells containing DNA of HPV-16 and BPV (Rohlfs et al., 1991; Baker & Howley, 1987). These RNAs could represent truncated forms of transcripts initiated in the early region of the genome. Alternatively, they could be transcribed from a promoter just upstream of the late genes. This latter hypothesis is supported by our finding of a promoter activity in this region of the viral DNA.

We are very grateful to B. Hirt for helpful discussions, to H. zur Hausen for the gift of cloned HPV-6b DNA, to B. Bentele for skilful technical assistance and to A.-M. Rodel and K. Wingate for help in preparing the manuscript. The work was supported by the Fonds National suisse de la recherche scientifique and the Ligue suisse contre le cancer.

References


(Received 14 July 1995; Accepted 13 September 1995)