Post-Transcriptional Regulation of Papillomavirus Gene Expression

Carl C. Baker

Basic Research Laboratory, DBS, National Cancer Institute,
National Institutes of Health, Bethesda, Maryland 20892-5055
E-mail: ccb@helix.nih.gov

Papillomavirus genomes are transcribed as one large complex transcription unit with multiple promoters, two poly(A) sites (early and late), and considerable alternative splicing. As a result, there are multiple opportunities for post-transcriptional regulation of papillomavirus gene expression. Although this area has not been as extensively studied as transcriptional regulation, there are a few well documented examples of post-transcriptional regulatory elements which have been identified in the papillomaviruses. This review will focus on elements which work at levels up through translation, but will not include post-translational regulation.

Alternative splicing of papillomavirus pre-mRNAs is extensive and has been described in detail in the maps presented elsewhere in this compendium. Splicing patterns and splice sites will not be detailed here except in cases where the regulation of splicing has been studied. In general, alternative splicing allows the generation of multiple mRNAs from one pre-mRNA and is essential for the differential expression of genes in a compact genome such as the papillomaviruses.

Alternative splicing may play an important role in the regulation of the E6 and E7 proteins of the high risk genital papillomaviruses. Splicing in the E6 ORF to give the E6*I, E6*II, and E6*III species certainly affects the amount of full length E6 protein that can be translated from an E6 mRNA and may also produce E6 proteins with altered functions (see Myers and Androphy E6 chapter and references therein). In addition, this splicing may also regulate translation of the E7 ORF by affecting translation initiation at the AUG in the E7 ORF, although there is some question about the validity of this concept (22,23). The cis elements and splicing factors which regulate splicing within the E6 ORF are not yet known, but will be important to identify. Expression of the E7 ORF in HPV-16 is also dependent on the 5′ splice site at nt 880 (5). Mutation of this splice site leads to skipping of the exon containing the E7 ORF, as would be predicted by the exon definition model of splicing (6,7). Thus changes in the recognition of either splice site flanking the E7 exon could regulate E7 expression in the high risk HPVs.

The early to late switch in BPV-1 gene expression is regulated at least partially at the level of splicing. At early to intermediate stages of the viral life cycle, a common 3′ splice site at nt 3225 is used for splicing almost all early pre-mRNAs as well as the late pre-mRNA. Most of these pre- mRNAs are also polyadenylated at the early poly(A) site. In contrast, at late stages of the life cycle, a novel 3′ splice site at nt 3605 is activated (2). Coupled with this is a switch in poly(A) site use and synthesis of the L1 mRNA. Thus there seems to be a link between 3′ splice site selection and poly(A) site choice. This switch in 3′ splice site utilization is dependent on differentiation and takes place as the keratinocyte progresses from the spinous layer to the granular layer (4). In addition, this change in 3′ splice site utilization may be a key step in the transition from the early to late stages of the viral life cycle since activation of a suboptimal 5′ splice site at nt 3764 is linked to use of the nt 3605 3′ splice site through limitations on exon size (2,7). In turn, activation of the nt 3764 5′ splice site may be required to suppress polyadenylation at the early poly(A) site and allow polyadenylation to occur at the downstream late poly(A) site (1).

Considerable information is now available about the mechanisms regulating selection of these alternative 3′ splice sites and is summarized in the model shown in Figure 1. Both the nt 3225 and nt 3605 3′ splice sites are suboptimal, allowing their recognition to be regulated by additional cis elements (30). Immediately downstream of the nt 3225 3′ splice site is a bipartite splicing regulatory element consisting of a purine-rich exonic splicing enhancer (SE1) followed immediately by a pyrimidine-rich exonic splicing suppressor (ESS) (Table 1 and Figures 1 and 2A) (30). This element is capable of

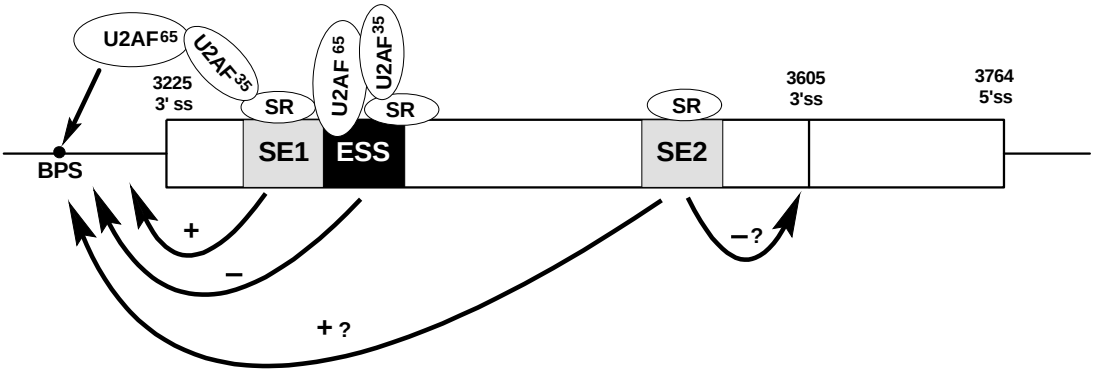

Figure 1. Model for the regulation of BPV-1 alternative splicing by exonic sequences. The region in the vicinity of the nt 3225 and 3605 3′ splice sites (ss) and nt 3764 5′ splice site are shown with the locations of the exonic splicing enhancers and suppressor indicated by shaded boxes. Protein-RNA and protein-protein interactions are indicated above the map. Arrows below indicate whether the interactions are stimulatory (+), inhibitory (-), or hypothetical (?).

Table 1 Cis-elements and trans-factors involved in papillomavirus post-transcriptional regulation

Virus	Name	Location	Nucleotide Position	Mechanism of Action	Trans-acting Factors	Figure	Reference
BPV-1	5′ splice site	late 3′UTR	7136–7144	inhibitor of polyadenylation	U1 snRNP	2B	(12, 13)
BPV-1	exonic splicing enhancer 1 (SE1)	early region	3256–3305	enhancer of splicing	SR proteins	2A	(30, 31)
BPV-1	exonic splicing enhancer 2 (SE2)	early region	3477–3526	enhancer and/or repressor of splicing	SR proteins	2A	(30, 31)
BPV-1	exonic splicing suppressor (ESS)	early region	3306–3354	repressor of splicing	U2AF65 SR proteins	2A	(30, 32)
BPV-1	short upstream ORFs	late leader	7250–7385	inhibition of translation			(8)
HPV-16	NRE	late 3′UTR	7130–7208	mRNA instability? polyadenylation? translation?	U2AF65	2E	(11, 13, 17, 18, 26)
HPV-16	ARE (A-rich element)	early 3′UTR	4005–4213	mRNA instability		2D	(14)
HPV-16		L1 ORF	5813–6150	decreases both mRNA and protein levels			(26)
HPV-1a		late 3′UTR	6958–7035	decreases both mRNA and protein levels	38, 52 kDa nuclear and 50, 74 kDa cytoplasmic proteins	2C	(27, 29)

both positively and negatively regulating use of the nt 3225 3′ splice site. A short distance upstream of the nt 3605 3′ splice site is a second purine-rich exonic splicing enhancer (SE2) (Figures 1 and 2A and Table 1). Both SE1 and SE2 are required for the preferential use of the nt 3225 3′ splice site in

undifferentiated cells. SE2 could work as either an exonic splicing enhancer on the nt 3225 3′ splice site or as an intronic splicing suppressor on the nt 3605 3′ splice site as seen for the adenovirus 3RE (16). SE1 and SE2 bind multiple members of a family of constitutive and alternative splicing factors known as SR proteins (31). SR proteins enhance use of suboptimal 3′ splice sites by recruiting the essential splicing factor to the 3′ splice site. U2AF in turn recruits the U2 snRNP to the branch point (reviewed in (20)). The ESS binds both SR proteins and U2AF65 and may compete with the suboptimal 3′ splice site for these factors (32). This particular arrangement of positive and negative elements allows the coordinated regulation of splicing by the same set of factors. However, it is not yet known how these factors are affected by keratinocyte differentiation and/or viral proteins.

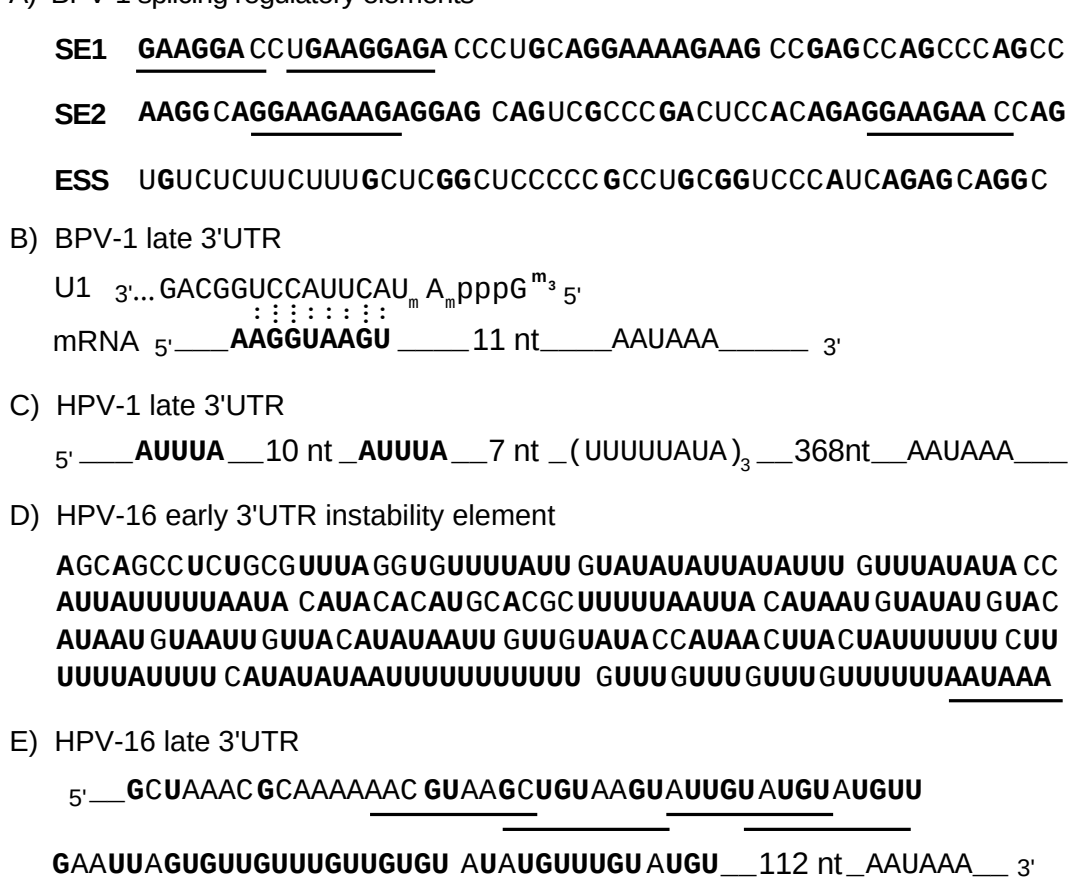

Figure 2. Sequences of papillomavirus post-transcriptional regulatory elements. References and nucleotide positions are given in Table 1. A) The sequences of BPV-1 splicing regulatory elements SE1, SE2, and the ESS are shown. Purine nucleotides are shown in bold. The underlined sequences in SE1 and SE2 are potential binding sites for the SR protein ASF/SF2 and either match most closely with the A13 sequence (AGAAGGAC) or the consensus sequence (RGAAGAAC) obtained through a SELEX procedure in reference (25). B) The BPV-1 late 3′UTR element (bold) is a regulatory 5′ splice site which base pairs with the 5′ end of the U1 snRNA and inhibits polyadenylation. C) Two AUUUA sequences (bold) in the HPV-1 late 3′UTR are essential for inhibitory activity, although three UUUUUAUA sequences also contribute to full activity of the element. D) The sequence of the HPV-16 early 3′ UTR is shown. Although the exact mRNA instability element has not been mapped, it most likely consists of an AU-rich region followed by a U-rich region (21). A and U residues are shown in bold. E) The HPV-16 late 3′UTR element has 4 overlapping nonconsensus 5′ splice site-like sequences (underlined) followed by a GU-rich region (Gs and Us shown in bold). In B-E, AAUAAA is the polyadenylation signal.

A number of additional negative regulatory elements have been identified in both coding regions and 3′ UTRs of several of the papillomaviruses. These elements are summarized in Table 1 and the sequences are shown in Figure 2. Perhaps the best studied of these elements resides in the BPV-1 late 3′UTR and inhibits the expression of the L1 and L2 mRNAs (12,13). This element has been mapped to a 5′ splice site which plays a regulatory role but is not used for splicing (Figure 2B). At least one splicing factor, the U1 snRNP, binds to the 5′ splice site and is required for its function. There are at least two mechanisms by which a 5′ splice site can inhibit gene expression: nuclear retention and inhibition of polyadenylation. It should be noted that the second mechanism may also lead to nuclear retention since in most cases polyadenylation is required for nucleocytoplasmic transport. Both *in vivo* and *in vitro* studies have demonstrated that a 5′ splice site can inhibit polyadenylation (1,10). In favor of a nuclear retention mechanism is the observation that the HIV-1 Rev protein in trans and the Rev-binding site in cis can block the inhibitory effect of a 5′ splice site (3). The Rev protein promotes the nucleocytoplasmic transport of mRNAs to which it is bound. Additional studies will be required to distinguish between these mechanisms. The HPV-16 late 3′ UTR contains an element which also inhibits expression of late mRNAs and has some similarities with the BPV-1 late 3′UTR (Table 1 and Figure 2E). Treatment of keratinocytes with PMA, an agent which induces differentiation, relieves inhibition by this element, suggesting that the late 3′UTR plays an important role in regulating differentiation- dependent late gene expression (11). This element appears to have two parts (see Figure 1E): the 5′ sequence has four overlapping motifs that look like weak (nonconsensus) 5′ splice sites (13). No single motif is sufficient for inhibition expression. However, multiple 5′ splice sites inhibit expression much more that a single 5′ splice site, suggesting that the four weak sites may act together to inhibit expression (3). A second downstream region consists of a GU-rich sequence that may bind the splicing factor U2AF65 (11). The mechanism by which this element functions is not clear. Tan et al. (26) used a p17gag reporter system *in vivo* and found a 64-fold decrease in protein levels and only a four-fold decrease in mRNA levels, suggesting that the 3′UTR element may affect multiple processes. *In vitro* studies demonstrated an effect on mRNA stability, but not polyadenylation (17,18). However, a role in mRNA destabilization has not been shown *in vivo*.

Although several papillomavirus late 3′UTRs have negative regulatory elements, the sequences of these elements are not conserved. This may relate to different tissue tropisms for the viruses. The HPV-1 late 3′UTR has an element which looks unlike either of the elements described above (27). This element has two essential AUUUA motifs followed by three U5AUA motifs which are less important (Table 1 and Figure 2C). Like the HPV-16 late 3′UTR element, the HPV-1 element affects protein levels much more that mRNA levels, suggesting that it acts predominantly as an inhibitor of translation. AUUUA sequences in the 3′UTRs of many cellular genes have been shown to destabilize mRNAs in the cytoplasm and inhibit translation (reviewed in (21)).

Inhibitory elements are not restricted to the late 3′UTRs. The HPV-16 early 3′UTR contains an element which destabilizes early mRNAs in vivo (14), although a destabilization phenotype for this element has not been universally seen (27). Although this element has not been finely mapped, the early 3′UTR consists of a long AU-rich region followed by a U-rich region (Figure 2D). This structure is typical of mRNA instability elements found in the 3′UTRs of other unstable mRNAs (21). Changes in early mRNA stability may play an important role in the increased expression of E6 and E7 proteins seen at late stages in malignant progression of HPV-16 lesions. In most cancers, the HPV genomes are integrated into the host chromosomes. Integration generally disrupts the E1 and E2 regions so that E6 and E7 mRNAs are expressed as viral/host chimeras. If integration occurs upstream of the poly(A) site of a cellular gene with stable mRNAs, the chimeric viral mRNAs would be more stable than the viral mRNAs expressed from extrachromosomal viral genomes, leading to increased viral mRNA levels in the cancers.

Other negative regulatory elements have also been found in coding regions. An element in the HPV-16 L1 ORF inhibits expression of a CAT reporter gene (Table 1) (26). Like many of the other elements described above, this one also reduces protein levels much more than mRNA levels. Neither the sequences which are required for the function of this element nor the mechanism of action are known. Several of the elements described above decrease protein levels more than mRNA levels, suggesting that papillomavirus late gene expression is subjected to translational regulation. This is supported by

in situ hybridization and immunoperoxidase studies of an HPV-6-infected vulvar condyloma which demonstrated that only a subset of the cells containing L1 mRNA also expressed high levels of L1 protein (24). Translation of papillomavirus proteins from multicistronic mRNAs can be blocked by the presence of the upstream ORFs. The E1∧E4 ORF in an HPV-6 or 11 E1∧E4∧L1 mRNA blocks the translation of the downstream L1 ORF (9,28). Likewise, four short ORFs in the first exon of the BPV-1 L1 mRNA inhibit translation of the L1 ORF (8). This effect is presumably due to either a block to 40S ribosome scanning or to a specific inhibitory peptide translated from one of the uORFs (reviewed in (19)). In contrast, the HPV-16 E5 ORF appears to be translated only if the upstream E2 ORF is also translated (15). The mechanism for this coupling is not known.

References

[1] Baker, C. C. 1997. (UnPub)

[2] Barksdale, S. K. and C. C. Baker. 1993. Differentiation-specific expression from the bovine papillomavirus type 1 P2443 and late promoters. *J. Virol.* **67**:5605–5616.

[3] Barksdale, S. K. and C. C. Baker. 1995. Differentiation-specific alternative splicing of bovine papillomavirus late mRNAs. *J. Virol*. **69**:6553–6556.

[4] Barksdale, S. K. and C. C. Baker. 1995. The human immunodeficiency virus type 1 Rev protein and the Rev-responsive element counteract the effect of an inhibitory 5′ splice site in a 3′ untranslated region. *Mol. Cell. Biol*. **15**:2962–2971.

[5] Belaguli, N. S., M. M. Pater, and A. Pater. 1992. Nucleotide 880 splice donor site required for efficient transformation and RNA accumulation by human papillomavirus type 16 E7 gene. *J. Virol*. **66**:2724–2730.

[6] Belaguli, N. S., M. M. Pater, and A. Pater. 1995. Splice sites of human papillomavirus type 16 E6 gene or heterologous gene required for transformation by E7 and accumulation of E7 RNA. *J. Med. Virol*. **47**:445–453.

[7] Berget, S. M. 1995. Exon recognition in vertebrate splicing. *J. Biol. Chem*. **270**:2411–2414.

[8] Binder, R., S. K. Barksdale, A. M. Del Vecchio, J. Tsai, and C. C. Baker. 1997. (UnPub)

[9] Brown, D. R., L. Pratt, J. T. Bryan, K. H. Fife, and K. Jansen. 1996. Virus-like particles and E1 E4 protein expressed from the human papillomavirus type 11 bicistronic E1 E4 L1 transcript. *Virology* **222**:43–50.

[10] Cooke, C. and J. C. Alwine. 1996. The cap and the 3′ splice site similarly affect polyadenylation efficiency. *Mol. Cell. Biol*. **16**:2579–2584.

[11] Dietrich-Goetz, W., I. M. Kennedy, B. Levins, M. A. Stanley, and J. B. Clements. 1997. A cellular 65-kDa protein recognizes the negative regulatory element of human papillomavirus late mRNA. *Proc. Natl. Acad. Sci. USA* **94**:163–168.

[12] Furth, P. A. and C. C. Baker. 1991. An element in the bovine papillomavirus late 3′ untranslated region reduces polyadenylated cytoplasmic RNA levels. *J. Virol*. **65**:5806–5812.

[13] Furth, P. A., W.-T. Choe, J. H. Rex, J. C. Byrne, and C. C. Baker. 1994. Sequences homologous to 5′ splice sites are required for the inhibitory activity of papillomavirus late 3′ untranslated regions. *Mol. Cell. Biol*. **14**:5278–5289.

[14] Jeon, S. and P. F. Lambert. 1995. Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: Implications for cervical carcinogenesis. *Proc. Natl. Acad. Sci. USA* **92**:1654–1658.

[15] Johnsen, C. K., M. Stanley, and B. Norrild. 1995. Analysis of human papillomavirus type 16 E5 oncogene expression *in vitro* and from bicistronic messenger RNAs. Inter*Virology* **38**:339–345.

[16] Kanopka, A., O. Mühlemann, and G. Akusjärvi. 1996. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* **381**:535–538.

[17] Kennedy, I. M., J. K. Haddow, and J. B. Clements. 1990. Analysis of human papillomavirus type 16 late mRNA 3′ processing signals *in vitro* and *in vivo*. *J. Virol*. **64**:1825–1829.

[18] Kennedy, I. M., J. K. Haddow, and J. B. Clements. 1991. A negative regulatory element in the human papillomavirus type 16 genome acts at the level of late mRNA stability. *J. Virol*. **65**:2093–2097.

[19] Kozak, M. 1992. Regulation of translation in eukaryotic systems. *Annu. Rev. Cell Biol*. **8**:197–225.

[20] Manley, J. L. and R. Tacke. 1996. SR proteins and splicing control. *Genes Dev*. **10**:1569–1579.

[21] Ross, J. 1996. Control of messenger RNA stability in higher eukaryotes. Trends. Genet. **12**:171–175.

[22] Sedman, S. A., M. S. Barbosa, W. C. Vass, N. L. Hubbert, J. A. Haas, D. R. Lowy, and J. T. Schiller. 1991. The full-length E6 protein of human papillomavirus type 16 has transforming and trans-activating activities and cooperates with E7 to immortalize keratinocytes in culture. *J. Virol*. **65**:4860–4866.

[23] Stacey, S. N., D. Jordan, P. J. F. Snijders, M. Mackett, J. M. M. Walboomers, and J. R. Arrand. 1995. Translation of the human papillomavirus type 16 E7 oncoprotein from bicistronic mRNA is independent of splicing events within the E6 open reading frame. *J. Virol*. **69**:7023–7031.

[24] Stoler, M. H., S. M. Wolinsky, A. Whitbeck, T. R. Broker, and L. T. Chow. 1989. Differentiation-linked human papillomavirus types 6 and 11 transcription in genital condylomata revealed by in situ hybridization with message-specific RNA probes. *Virology* **172**:331–340.

[25] Tacke, R. and J. L. Manley. 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J*. **14**:3540–3551.

[26] Tan, W., B. K. Felber, A. S. Zolotukhin, G. N. Pavlakis, and S. Schwartz. 1995. Efficient expression of the human papillomavirus type 16 L1 protein in epithelial cells by using Rev and the Rev-responsive element of human immunodeficiency virus or the cis-acting transactivation element of simian retrovirus type 1. *J. Virol*. **69**:5607–5620.

[27] Tan, W. and S. Schwartz. 1995. The Rev protein of human immunodeficiency virus type 1 counteracts the effect of an AU-rich negative element in the human papillomavirus type 1 late 3′ untranslated region. *J. Virol*. **69**:2932–2945.

[28] Tomita, Y. and B. Simizu. 1993. Translational properties of the human papillomavirus type-6 L1-coding mRNA. *Gene* **133**:223–225.

[29] Zhao, C. P., W. Tan, M. Sokolowski, and S. Schwartz. 1996. Identification of nuclear and cytoplasmic proteins that interact specifically with an AU-rich, cis-acting inhibitory sequence in the 3′ untranslated region of human papillomavirus type 1 late mRNAs. *J. Virol*. **70**:3659–3667.

[30] Zheng, Z. M., P. J. He, and C. C. Baker. 1996. Selection of the bovine papillomavirus type 1 nucleotide 3225 3′ splice site is regulated through an exonic splicing enhancer and its juxtaposed exonic splicing suppressor. *J. Virol*. **70**:4691–4699.

[31] Zheng, Z. M., P. J. He, and C. C. Baker. 1997. Structural, functional, and protein binding analyses of bovine papillomavirus type 1 exonic splicing enhancers. *J. Virol*. (In Press)

[32] Zheng, Z. M., P. J. He, and C. C. Baker. 1997. (UnPub)