Papillomavirus gene expression is strictly linked to the differentiation state of the infected cell and is highly regulated at the level of transcription and RNA processing. All papillomaviruses make extensive use of alternative mRNA polyadenylation and splicing to control gene expression. This chapter contains a compilation of all known alternatively spliced papillomavirus mRNAs and it summarizes our current knowledge of viral RNA elements, and viral and cellular factors that control papillomavirus mRNA processing.

© 2013 Elsevier Inc. All rights reserved.

**Introduction**

Papillomaviruses are small DNA viruses with a circular double stranded DNA genome (zur Hausen, 2002). Their genomes typically encode early and late genes that are expressed in a temporal and highly regulated manner (Howley and Lowy, 2006). Gene expression is regulated at the level of transcription (Bernard, 2002; Thierry, 2009) and RNA processing (Graham, 2008; Schwartz, 2008; Zheng and Baker, 2006) and the papillomavirus proteins are produced from a myriad of alternatively spliced and polyadenylated mRNAs. Maps of all known papillomavirus mRNAs are shown in Appendix I. Advantages of the extensive use of alternative splicing and polyadenylation include the ability to express many genes from a compact genome, as well as the ability to individually regulate expression of each gene during the viral life cycle. Papillomavirus gene expression is tightly linked to the differentiation program of infected epithelial cells. The most obvious example is perhaps the well-conserved delay in late L1 and L2 gene expression to the uppermost layers with terminally differentiated cells in the squamous epithelium (Chow et al., 2010; Doorbar, 2005; Moody and Laimins, 2010). Activation of L1 and L2 expression requires a viral promoter-switch, a change of polyA signal and derepression of two alternative splice sites. Merely a promoter switch does not suffice, as experiments in which the late HPV-16 promoter p670 was replaced by the constitutively active human cytomegalovirus promoter did not activate late gene expression (Orru et al., 2012; Zhao et al., 2004). Therefore, regulation at the level of RNA processing plays an important role in papillomavirus gene expression. Furthermore, HPV-1 Rev and RRE that are required for efficient nuclear export of partially spliced HIV-1 mRNAs (Felber and Pavlakis, 1993), enhance BPV-1, HPV-1 and HPV-16 late gene expression (Barksdale and Baker, 1995; Tan et al., 1995; Tan and Schwartz, 1995), and adenovirus E4orf4 that regulates the switch from early to late gene expression in adenoviruses by dephosphorylating splicing factors (Akusjarvi and Stevenin, 2003), can induce HPV-16 L1 mRNA production by enhancing viral mRNA splicing (Somberg et al., 2009). Significant effects on papillomavirus gene expression by relatively subtle mutations at RNA processing signals in complete papillomavirus genomes of different types have been reported (Andrew and DiMaio, 1993; Deng et al., 2003; Hubert and Laimins, 2002; Klumpp et al., 1997; Poppelreuther et al., 2007; Terhune et al., 2001, 1999). These results underscore the importance of RNA processing in the papillomavirus gene expression program.

This chapter discusses cis-acting papillomavirus RNA elements and viral and cellular trans-acting factors that regulate papillomavirus gene expression.

**RNA elements in papillomavirus late and early 3′-UTR sequences**

**Late 3′-UTR sequences**

Inhibitory sequences in the late untranslated region of papillomavirus mRNA were originally discovered in BPV-1 and in HPV-16 (Furth and Baker, 1991; Kennedy et al., 1990, 1991) and are relatively well characterized (Graham, 2008). One may speculate that the role of these sequences in the viral life cycle is either to prevent premature late gene expression or to serve as landing pads for cellular RNA binding proteins that activate late gene expression in response to cellular differentiation. Inhibitory sequences are present in the late 3′-UTR of all HPVs that have been analyzed, including HPV-1, HPV-2, HPV-6, HPV-16, HPV-18, HPV-31, HPV-41 and HPV-61 (Cumming et al., 2002; Kennedy et al., 1990, 1991; Tan and Schwartz,
In general, these sequences are not well conserved but many are AU- or GU-rich and often contain multiple copies of the G-U3-G sequence which resembles the AUUUA-motif often found in AU-rich RNA instability elements in the 3'-UTR on cellular mRNAs (Zhao et al., 2007b). These late UTR elements have an inhibitory function in mitotic cells including cancer cells that must be overcome in terminally differentiated cells that are permissive for late gene expression. Best characterized are the late UTR elements in BPV-1 (Fig. 1), HPV-1 (Fig. 2) and HPV-16 (Figs. 3 and 4).

It is well established that the negative element in the BPV-1 late UTR is a relatively short, 5'-splice site-like sequence that binds specifically to the U1snRNA part of the cellular U1snRNP complex (Fig. 1) (Furth et al., 1994). This binding inhibits polyadenylation of the late BPV-1 mRNAs (Furth et al., 1994; Gunderson et al., 1998). In contrast, the inhibitory element in the HPV-1 late UTR is a classical AU-rich RNA instability element (ARE) (Fig. 2) (Sokolowski et al., 1997; Tan and Schwartz, 1995), similar to those originally discovered in a subset of short-lived cellular mRNAs,

**Fig. 1.** (A) Schematic drawing of the BPV-1 genome. Boxes represent protein coding sequences, arrows promoters and black and white triangles, 5'- and 3'-splice sites, respectively. Early and late polyA signals named pAE and pAL are indicated. A subset of viral mRNAs is shown below the genome and the most likely translation product for each mRNA is indicated to the right. (B) Left: The positions of splicing silencer (red) (ESS1 and ESS2) and splicing enhancer (green) (SE1, SE2 and SE4) elements and the cellular proteins they interact with are indicated (Zheng and Baker, 2006). See text for details and Table 1 for SR-protein nomenclature (Manley and Krainer, 2010). Colored arrows show the effect of the regulatory RNA elements on the various BPV-1 splice sites. Right: The BPV-1 late UTR encodes a negative regulatory RNA element that binds specifically to cellular U1snRNP. This interaction inhibits late mRNA processing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
including interleukin-, interferon- and c-fos-mRNAs (Barreau et al., 2006). Similarly to the c-fos ARE, the HPV-1 ARE reduces mRNA stability and inhibits translation (Sokolowski et al., 1997; Wiklund et al., 2002), and binds specifically to hnRNP C1/C2 and HuR (Fig. 2) (Sokolowski et al., 1999, 1997; Sokolowski and Schwartz, 2001; Zhao et al., 1996). HuR is a shuttling protein whereas hnRNP C1/C1 is strictly nuclear. Both proteins regulate the stability, polyadenylation and nuclear export of cellular mRNAs. HuR binds specifically to the AUUUA- and AUUUUUA-motifs in the HPV-1 AU-rich RNA element (Sokolowski et al., 1999), while hnRNP C1/C2 binds the penta-U nucleotides (Sokolowski and Schwartz, 2001; Sokolowski et al., 1997). The inhibitory activity of the HPV-1 AU-rich element is lower in cell lines in which a large fraction of the HuR protein resides in the cytoplasm (Carlsson and Schwartz, 2000), suggesting that HuR interacts with the HPV-1 AU-rich element in cells. The HPV-1 AU-rich RNA element also interacts with polyA-binding protein (PAPB) (Fig. 2) (Wiklund et al., 2002), indicating that the inhibitory effect on translation exerted by the HPV-1 ARE is caused by binding to PAPB.

The HPV-16 negative regulatory RNA element can be divided into a 5′-half with multiple, weak 5′-splice site-like motifs that all contribute to the inhibitory activity of this region, and a GU-rich 3′-half (Figs. 3 and 4). While the former interacts with U1snRNP (Cumming et al., 2003; Furth et al., 1994), like the BPV-1 late UTR element (Furth et al., 1994), the 3′-half has been reported to bind a number of proteins including CUG-BP1 (Goraczniak and Gunderson, 2008), U2AF65 (Cumming et al., 2009; Koffa et al., 2000; McPhillips et al., 2004), HuR (Cumming et al., 2009; Koffa et al., 2000), hnRNP A1 (Cheunim et al., 2008), weakly to CstF-64 (Cumming et al., 2002; Koffa et al., 2000) and indirectly to ASF/SF2 (McPhillips et al., 2004), now named SRSF1 (Fig. 4B) (see Table 1 for new nomenclature of SR-proteins (Manley and Krainer, 2010)). The HPV-31 late 3′-UTR can also bind U2AF65, HuR and CstF-64 (Cumming et al., 2002). With the exception of

Fig. 2. (A) Schematic drawing of the HPV-1 genome. Boxes represent protein coding sequences, the arrow a promoter and black and white triangles, 5′- and 3′-splice sites, respectively. Early polyA signal pAE and the two late polyA signals pAL1 and pAL2 are indicated. A subset of viral mRNAs is shown below the genome and the most likely translation product of each mRNA is indicated to the right. (B) Blow up of the HPV-1 late UTR shows the position of the AU-rich inhibitory RNA element (red) (Graham, 2008; Schwartz, 2008; Zheng and Baker, 2006). The exact sequence and protein-binding partners of the AU-rich element are shown below the schematic drawing of the UTR. See text for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
CUG-BP1, specific binding sites for these proteins have not been identified. Under slightly different experimental conditions, UGUUUGU- or UGUUU-motifs in the HPV-16 late UTR bind primarily a 55 kDa nuclear protein that is distinct from U2AF65 (Zhao et al., 2007b) and was identified as CUG-BP1 (Goraczniak and Gunderson, 2008). The CUG-BP1 acts in synergy with the upstream 5′-splice site like elements to inhibit polyA site activity, nuclear mRNA export and translation (Goraczniak and Gunderson, 2008). One may speculate that the HuR protein binds to the GUUUG-motifs that resemble AUUUA-motifs to which it normally binds, and therefore competes with CUG-BP1 for the overlapping UGUUUGU-sites, perhaps explaining why HuR interactions with the HPV-16 late 3′-UTR are not always detected (Goraczniak and Gunderson, 2008) (Fig. 4B). HuR could potentially affect HPV-16 late mRNA half-life or nuclear export. Overexpression of HuR in chronically HPV-16 infected keratinocytes induced HPV-16 late gene expression, and knock-down of HuR reduced late gene expression (Cumming et al., 2009), but the exact role of HuR in the HPV-16 infectious cycle remains to be determined.

**Early 3′-UTR elements**

Less is known about RNA elements in the papillomavirus early 3′-UTR, but it could potentially regulate early mRNA stability, translation, or polyadenylation. Fip1 is a member of the cellular polyadenylation complex CPSF and enhances polyadenylation by binding to U-rich sequences immediately upstream of polyadenylation signals (Kaufmann et al., 2004). Fip1 binds to U-rich sequences in the HPV-16 early 3′-UTR that has a modest stimulatory effect on the early polyadenylation signal pAE (Zhao et al., 2005b) (Figs. 3 and 4). Polypyrimidine tract binding protein (PTB), also known as hnRNP I (Han et al., 2010), is a cellular splicing factor that binds U-rich sequences in the HPV-16 early UTR that are located between late 5′-splice site SD3632 and the early
polyadenylation signal pAE (Zhao et al., 2005b) (Fig. 4A). Over-expression of PTB also induces HPV-16 late gene expression from subgenomic HPV-16 expression plasmids (Somberg et al., 2008), indicating that PTB either inhibits pAE, and/or activates SD3632. The cellular HuR protein can induce HPV-16 late gene expression from subgenomic HPV-16 plasmids that lack the HPV-16 late 3′-UTR (Johansson et al., 2012), demonstrating that HuR can act on other HPV-16 sequences than the late 3′-UTR as described above. As HuR induced primarily L2 mRNAs as asposed to spliced L1 mRNAs, it is likely to inhibit polyadenylation at pAE, perhaps by binding to the U-rich early 3′-UTR (Johansson et al., 2012), demonstrating that HuR can act on other HPV-16 sequences than the late 3′-UTR as described above. As HuR induced primarily L2 mRNAs as asposed to spliced L1 mRNAs, it is likely to inhibit polyadenylation at pAE, perhaps by binding to the U-rich early 3′-UTR (Johansson et al., 2012). hnRNP C1 and C2 are two related, strictly nuclear proteins that affect cellular mRNA polyadenylation, stability and transport (Han et al., 2010). They bind directly to U-rich sequences in the early UTR of HPV-16 and may also regulate pAE (Zhao et al., 2005b) (Fig. 4A). The HPV-16 early 3′-UTR contains cytoplasmic polyadenylation elements (CPEs) that interact with cytoplasmic polyadenylation element binding protein (CPEB) (Glahder et al., 2010; Vinther et al., 2005), suggesting that cytoplasmic polyadenylation machinery may be involved in HPV-16 gene expression. In addition, the HPV-16 eUTR can reduce the half-life of early HPV-16 mRNAs under certain conditions (Jeon et al., 1995; Jeon and Lambert, 1995), while it does not under other conditions (Häfner et al., 2008; Zhao et al., 1996). The role of each viral 3′-UTR element and its cognate trans-acting factor in the viral life cycle remains to be determined.

Splicing regulatory elements

BPV-1 mRNA splicing

BPV-1 encodes numerous splice sites, like most papillomaviruses. The regulation of three splice sites (SA3225, SA3605 and SD3764) located in the central portion of the genome has been studied in detail (Jia and Zheng, 2009). SA3225 is used for production of many early BPV-1 mRNAs and one of the two potential L2 mRNAs termed LL2 (Fig. 1A). It is a suboptimal 3′-splice site (Zheng et al., 2000a) and splicing is stimulated by two purine rich splicing enhancers (SE1 and SE2) (Zheng et al., 1996) that interact with multiple splicing factors of the SR-protein family (SRSF1, SRSF4, SRSF5 and SRSF6) (Zheng et al., 1997) (Fig. 1) (see Table 1 for SR protein nomenclature (Manley and Krainer, 2010)). These factors are known to enhance splicing of cellular
mRNAs, but may also have a splicing inhibitory role. SRSF3 appears to stimulate splicing to SA3225 (Liu et al., 2003). In addition, SA3225 is under control of a pyrimidine rich splicing silencer (ESS1) that counteracts the enhancers (Zheng et al., 1996, 1999). ESS1 interacts with PTB, U2AF65 and SR proteins, but only SR proteins appear to contribute to its splicing-inhibitory function (Zheng et al., 1998) (Fig. 1). Production of BPV-1 L1 mRNAs is dependent on an alternative 3′-splice site named SA3605 that is located downstream of SA3225 (Fig. 1). Splicing to SA3605 results in further splicing between SD3764 and SA5609, two splice sites that are used exclusively to produce L1 mRNAs. This creates a small exon in between SA3605 and SD3764 that harbors splicing enhancer and silencer elements (SE4 and ESS2) that regulate SA3605, and perhaps SD3764 (Zheng et al., 2000b) (Fig. 1). Cellular factors that bind SE4 to stimulate splicing to SA3605 have not been identified, but binding of SRSF3 (Table 1) (Manley and Krainer, 2010) to SE4 enhances splicing to SA3225 while simultaneously inhibiting SA3605 (Jia et al., 2010, 2009) (Fig. 1). High levels of SRSF3 therefore promote production of BPV-1 early mRNAs and one of the potential L2 mRNAs (L2L), while inhibiting production of L1 and L25 mRNAs. As the L25 mRNA is the best candidate L2 mRNA since it contains fewer ORFs upstream of L2 than the L2L mRNA, it appears that SRSF3 acts to promote BPV-1 early gene expression.

**HPV-16 mRNA splicing**

In contrast to BPV-1, HPV-16 has only one major 3′-splice site in the central portion of the genome (SA3358) (Fig. 3) that is used for production of both early and late mRNAs, with the exception of E1 and E2 mRNAs (Fig. 3). Splice site SA3358, is the most commonly used splice site on the HPV-16 genome (Schmitt et al., 2010). Splicing to the corresponding splice site in HPV-31 (SA3295) is detected as early as 8 h post-infection (Ozbun, 2002). Mutational inactivation of SA3295 in HPV-31 activated a cryptic splice site three nucleotides further down (Klump et al., 1997), strongly suggesting that this splice site is under control of splicing enhancer elements. Indeed, the corresponding splice site in HPV-16, SA3358, is suboptimal and totally dependent on a downstream splicing enhancer (Rush et al., 2005) (Fig. 3). This enhancer coincides with a number of predicted binding sites for the SR protein SRSF1 (Somberg and Schwartz, 2010) (Fig. 4) (Table 1) (Manley and Krainer, 2010). Mutational inactivation of all predicted SRSF1 binding sites destroyed the splicing enhancer (Somberg and Schwartz, 2010). However, the exact binding site for SRSF1 needs to be determined. Interestingly, inactivation of the SA3358 splicing enhancer also shuts down the polyA signal pAE, demonstrating that the splicing enhancer at SA3358 is very efficient and that splicing to SA3358 enhances polyadenylation at pAE (Rush et al., 2005) (Fig. 3). However, overexpression of SRSF1 can induce HPV-16 late gene expression, suggesting that it may act on other splice sites on the HPV-16 genome, or that SA3358 activity is dependent on the exact concentration of SRSF1 (Somberg and Schwartz, 2010). Similarly, overexpression of SRSF9 (Table 1) (Manley and Krainer, 2010), which is the closest relative of SRSF1 in the SR protein family, causes skipping of the exon between SA3358 and SD3632 in HPV-16 and redirects splicing to SA5639 (Somberg et al., 2011). Binding of SRSF9 to the enhancer at SA3358 may negatively interfere with its function. Alternatively, SRSF9 stimulates splicing to SA5639 by another mechanism. In addition, a 28-nucleotide sequence overlapping predicted SRSF1 binding sites number 5 and 6 of the SA3358 enhancer was found to bind SRSF3, SRSF4, SRSF6, “Srp30s” (possibly including SRSF1, SRSF2 and/or SRSF9) and YB-1 (Jia et al., 2009) (see Table 1 for new SR protein nomenclature (Manley and Krainer, 2010)). Mutational inactivation of the SRSF3 binding sites or knock down of SRSF3 also enhanced HPV-16 late gene expression, suggesting that SRSF3 affects SA3358 (Jia et al., 2009). However, the exact binding site for SRSF3 needs to be determined. The role of the remaining factors remain unclear. In conclusion, at least three SR proteins bind to the enhancer region downstream of SA3358 and could regulate HPV-16 late gene expression: SRSF1 (Somberg and Schwartz, 2010), SRSF3 (Jia et al., 2009) and SRSF9 (Jia et al., 2009; Somberg et al., 2011) (Table 1) (Manley and Krainer, 2010).

mRNAs that are spliced to HPV-16 SA3358 can either be polyadenylated at pAE to generate mRNAs encoding E6, E7, E4 or E5, or polyadenylated at pAL to produce L2 mRNAs (Fig. 3). Alternatively, the 5′-splice site SD3632 that is located between SA3358 and pAE is active and generates L1 mRNAs by splicing to SA5639. SD3632 and pAE are utilized in a mutually exclusive manner and therefore compete with each other (Fig. 3). Consequently, SD3632 is suppressed during the early stage of the HPV-16 life cycle in which the majority of the HPV-16 mRNAs are polyadenylated at pAE. The small exon between SA3358 and SD3632 encodes a splicing silencer that efficiently inhibits SD3632 (Rush et al., 2005) (Figs. 3 and 4). Proteins binding to this silencer are likely to regulate L1 expression in a differentiation-dependent manner, but remains to be identified.

The HPV-16 late 3′-splice site SA5639 is used exclusively to produce late L1 mRNAs and is under control of a splicing enhancer within the first 17 nucleotides immediately downstream of SA5639 (Zhao et al., 2007a) (Figs. 3 and 4). However, multiple splicing silencers located downstream of the enhancer override the enhancer and suppress SA5639 (Zhao et al., 2004, 2007a). These splicing silencers interact with hnRNP A1 in a sequence-specific manner (Zhao et al., 2004, 2007a; Zhao and Schwartz, 2008). hnRNP A1 is a pleiotropic cellular protein that shuttles between nucleus and cytoplasm and has among all, been shown to bind cellular splicing silencers to inhibit splicing of cellular mRNAs (Han et al., 2010). Mutational inactivation of the hnRNP A1 binding sites alleviates inhibition of SA5639 (Zhao et al., 2004, 2007a; Zhao and Schwartz, 2008). hnRNP A1 is highly expressed in the lower to mid layers of the epithelium, but is undetectable in terminally differentiated keratinocytes that are permissive for papillomavirus late gene expression, lending support to the idea that hnRNP A1 inhibits HPV-16 late gene expression in a cell-differentiation dependent manner (Fay et al., 2009; Zhao et al., 2007a).

Utilization of splice sites SD226, SA409, SA526 and SA742, upstream of SD880 in HPV-16 gives rise to alternatively spliced mRNAs expressing various forms of the HPV-16 E6 oncoprotein and E7, but little is known about cis-acting RNA elements that regulate these splicing events. However, the branch site used in some of these splicing events was mapped to AACAACC with the actual branch point at A385 in the HPV-16R genome (Ajíro et al., 2012). Sequence polymorphism in this region of the genome has been shown to give rise to variations in splicing efficiency between various HPV-16 genomes, thereby affecting expression levels of the viral oncoproteins (Lopez-Urrutia et al., 2012). Mapping of splicing regulatory elements in this region of the HPV-16 genome is therefore of interest.

Splicing between SD226 and SA409 generates mRNAs that are translated to E6*I and E7*I proteins by leaky scanning (Stacey et al., 1995), by a ribosomal shunting mechanism (Remm et al., 1999) or by a translation-reinitiation mechanism (Tang et al., 2006), whereas mRNAs that remain unspliced between SD226 and SA409 produce full-length E6. However, the latter mRNAs are rare since splicing between SD226 and SA409 is very efficient. Splicing between SD226 and SA409 was stimulated by hnRNP A1 and hnRNP A2, but binding sites for these two proteins on the HPV-16 E6/E7 mRNAs have not been identified (Rosenberger et al., 2010). hnRNP A1 and hnRNP A2 are two closely related splicing factors that often inhibit mRNA splicing (Han et al., 2010). For example, hnRNP A1 binds to splicing silencers that inhibit HPV-16 late L1
mRNA splicing as described above (Zhao et al., 2004, 2007a; Zhao and Schwartz, 2008) (Fig. 4B). High levels of hnRNP A1 would therefore inhibit L1 production and promote E6* and E7 production. Indeed, HPV-16 containing cervical cancer cells are characterised by high hnRNP A1 levels (Fay et al., 2009) and efficient splicing between SD226 and SA409 (Cornelissen et al., 1990; Smotkin et al., 1989), but little L1 mRNA production.

The effect of individual splicing factors on papillomavirus gene expression is complex and difficult to predict as illustrated by the following example: Multiple SRSF1 sites downstream of HPV-16 3′-splice site SA3358 enhance splicing to SA3358 (Somberg and Schwartz, 2010). In an intracellular environment in which preferably the early p97 promoter is used, SRSF1 would primarily stimulate expression E6* and E7, as the most abundant HPV-16 mRNAs encoding these proteins are spliced to SA3358 (Fig. 3). However, when the late promoter p670 is activated and dominates, SRSF1 would enhance E4 and presumably E5 production (Fig. 3). At later stages of the infection when read-through at pAE into the late region commences, enhancement of SA3358 by SRSF1 would instead promote production of L2 mRNAs, and in case of simultaneous de-repression of late splice sites SD3632 and SA5639 and activation of L1 mRNA splicing, SRSF1 would stimulate L1 expression (Fig. 3). The same line of reasoning applies to many of the cellular and viral factors that regulate HPV gene expression.

In general, the majority of all RNA binding proteins that have been studied in relation to papillomavirus infection in cervical epithelium are highly expressed in the lower layers of the epithelium and are undetectable in the superficial layers of terminally differentiated keratocytes (Fay et al., 2009; Mole et al., 2009a). SR proteins and hnRNPs are often highly expressed in cervical cancer, as well as in high-grade cervical lesions (Fay et al., 2009; Mole et al., 2009a). This is probably attributable to the increased anabolic requirements of dividing cells, which predicts a greater need for RNA processing factors in cancer cells than in normal cells. Relative concentrations of splicing factors in papillomavirus infected cells may be as important as absolute concentrations.

**Viral factors that regulate papillomavirus RNA processing**

Expression levels of HPV E2 increases with cell differentiation (Xue et al., 2010) and recent experiments have shown that E2 from both mucosal and cutaneous HPV types can induce HPV late gene expression (Johansson et al., 2012). E2 causes a read-through at pAE into the late region commences, enhancement of SA3358 by SRSF1 would instead promote production of L2 mRNAs, and in case of simultaneous de-repression of late splice sites SD3632 and SA5639 and activation of L1 mRNA splicing, SRSF1 would stimulate L1 expression (Fig. 3). The same line of reasoning applies to many of the cellular and viral factors that regulate HPV gene expression.

In general, the majority of all RNA binding proteins that have been studied in relation to papillomavirus infection in cervical epithelium are highly expressed in the lower layers of the epithelium and are undetectable in the superficial layers of terminally differentiated keratocytes (Fay et al., 2009; Mole et al., 2009a). SR proteins and hnRNPs are often highly expressed in cervical cancer, as well as in high-grade cervical lesions (Fay et al., 2009; Mole et al., 2009a). This is probably attributable to the increased anabolic requirements of dividing cells, which predicts a greater need for RNA processing factors in cancer cells than in normal cells. Relative concentrations of splicing factors in papillomavirus infected cells may be as important as absolute concentrations.

**Intragenic elements that regulate polyadenylation and translation**

The papillomavirus early polyadenylation signal occupies a particularly important position on the papillomavirus genome as it divides the viral genome into early and late genes. As such, it must be subject to regulation. Mutational inactivation of pAE in BPV-1, HPV-31 or HPV-16, results in increased read-through into the late L1 and L2 coding regions, but also causes efficient activation of multiple cryptic polyadenylation signals located upstream of pAE (Andrew and DiMaio, 1993; Terhune et al., 1999, 2001; Zhao et al., 2005b). These results suggest that papillomavirus pAE is controlled by positive regulatory RNA elements. Indeed, such elements were identified in the L2 coding region of HPV-16 (Oberg et al., 2005, 2003) and HPV-31 (Terhune et al., 1999, 2001). In HPV-16, the L2 RNA elements encode multiple GGG-motifs that interact with hnRNP H (Oberg et al., 2005) (Figs. 3 and 4). Mutual antagonism of the hnRNP H binding sites correlated with reduced polyadenylation efficiency at pAE, indicating that hnRNP H stimulates polyadenylation at HPV-16 pAE (Oberg et al., 2005). Cellular polyadenylation factor CstF-64 also binds to the L2 RNA sequences in both HPV-16 and HPV-31 and presumably enhance polyadenylation at pAE (Oberg et al., 2005; Terhune et al., 2001). Intriguingly, HPV-16 L1 protein binds to hnRNP H, suggesting that L1 could regulate late gene expression (Zheng et al., 2012). However, it remains to be seen if hnRNP H and L1 are expressed in the same cells in the cervical epithelium. High levels of polyadenylation factors would inhibit late gene expression. Similarly to many other RNA processing factors, polyadenylation factors are down regulated in response to cellular differentiation.

The HPV-16 L1 and L2 coding region contains multiple RNA sequences that inhibit gene expression (Collier et al., 2002; Mori et al., 2006; Oberg et al., 2003; Sokolowski et al., 1998; Tan et al., 1995). Similar sequences have been described in canine oral papillomavirus (COPV) (Berg et al., 2005). These sequences are particularly efficient in HPV-16 L1 and L2 compared to HPV-1 L1 and L2 (Sokolowski et al., 1998). They reduce mRNA levels and/or inhibit utilization of the mRNA. HPV genomes are relatively AT-rich compared to most mammalian genomes, and AT-rich genes are in general poorly expressed compared to GC-rich genes (Kudla et al., 2006). The presence of rare amino acid codons in the
BPV and HPV genomes may also contribute to inefficient translation of L1 and L2 (Cu et al., 2004; Zhao et al., 2005a). Binding of hnRNP A1 proteins to splicing silencers in the HPV-16 L1 coding region also inhibits L1 expression. Mutational inactivation of these hnRNP A1 binding sites in the splicing silencers resulted in a dramatic increase in L1 production from L1 CDNA expression plasmids (Collier et al., 2002; Zhao et al., 2004; Zhao and Schwartz, 2008). These mutant L1 plasmids produced enough L1 to evoke cellular and humoral immune responses when injected into mice, which was in stark contrast to wild type L1 plasmids (Rollman et al., 2002). The binding of cellular proteins to L2 RNA also correlates with poor L2 expression. RNA of the L2 coding sequences in wild type and other HPV genomes may also contribute to inefficient translation. It is therefore predicted to inhibit translation of all downstream ORFs when it is present on all mRNAs initiated at the HPV-16 late, differentiating cervical cell epithelial cells. J Virol. 80, 987–997.

Felder, B.K., Pavlakis, G.N., 1993. Molecular biology of HPV-1: positive and negative regulatory elements important for virus expression. AIDS 7, S51.

References


Gu, W., Li, M., Zhao, W.M., Fang, N.X., Bu, S., Frazer, I.H., Zhao, K.-N., 2004. mRNA’s (CGA) differentially regulates expression of wild-type and codon-modified papilloma virus L1 genes. Nucleic Acids Res. 32, 4448–4461.


Sokolowski, M., Schwartz, S., 2001. Heterogeneous nuclear ribonucleoprotein C binds exclusively to the functionally important AUUUUU motifs in the human papillomavirus type-1 AU-rich inhibitory element. Virus Res. 73, 163–175.


Somberg, M., Schwartz, S., 2010. Multiple ASF/SF2 sites in the HPV-16 E4-coding region promote splicing to the most commonly used 3′-splice site on the HPV-16 genome. J. Virol. 84, 8219–8230.


Zhao, X., Schwartz, S., 2008. Inhibition of HPV-16 L1 expression from L1 cDNAs correlates with the presence of hnRNP A1 binding sites in the L1 coding region. Virus Genes 36, 45–53.


