[Frontiers in Bioscience 13, 5646-5663, May 1, 2008]

Papillomavirus 3' UTR regulatory elements

Sheila V Graham

Division of Infection and Immunity, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow, G12 8TA, Scotland, UK

TABLE OF CONTENTS

1. ABSTRACT

Papillomaviruses infect epithelial cells causing mainly benign lesions or warts. In some rare instances, these may progress to malignancy. For example, human papillomavirus type 16 (HPV16) is the causative agent of 60% of cases of cervical cancer. The replication cycle of papillomaviruses is intimately linked to epithelial differentiation. In particular, late gene expression is completed exclusively in the upper epithelial layers. Regulation of late gene expression is largely by post-transcriptional means. RNAs encoding the late proteins, the virus capsid proteins L1 and L2, can be detected in the lower layers of infected epithelia but protein is detected only in the upper layers. It is clear that cellular factors mediate this gene regulation. RNA regulatory elements that bind RNA processing factors that mediate post-transcriptional control have been identified in the 3' untranslated regions of a number of papillomaviruses. These elements, the proteins they bind and the mechanisms by which they are proposed to act are discussed. Further understanding of such host-virus molecular interactions may lead to development of novel strategies for abrogating virus infection.

2. INTRODUCTION

Papillomaviruses are nonenveloped double stranded DNA viruses with a virion size of around 55 nm and a genome of approximately 8 kbp. This group of viruses infects cutaneous and mucosal epithelial cells and usually causes benign lesions or warts (1). There are probably upwards of 300 papillomavirus types and possibly up to 200 different human types identified. Of these, over 120 are fully sequenced and can be divided into seventeen different genera in a phylogenetic tree (1). The two major groups are the alpha group that contain species that infect anogenital mucosal epithelia and the beta group that infect the skin. Papillomaviruses can also be designated "low risk" or "high risk" depending on the risk of progression to cancer of cells in infected tissue.

Among the first papillomaviruses studied were bovine papillomavirus type 1 (BPV1), a delta group papillomavirus, and human papillomavirus type 1 (HPV1) in the mu group. BPV1 causes benign fibropapillomas in cattle but can cause penile cancer in bulls and equine sarcoids that can transform to malignancy. In cell culture the virus can transform rodent fibroblasts. HPV1 causes

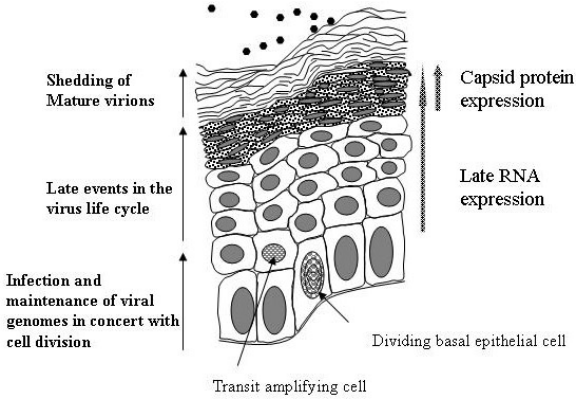

Figure 1. A schematic diagram of a differentiating epithelium. Nuclei are shown in grey. The granular layer cells are stippled. Hexagons show virus particles released from disintegrating cornified layer cells. A generalized pattern of major events in the virus life cycle is indicated on the left of the diagram. A generalized view of late gene expression is indicated on the right of the diagram. Arrows indicate a basal layer stem cell were papillomaviruses may establish latent infection and a transit amplifying cell where the virus begins its replicative life cycle.

verrucae in humans and is never associated with malignant transformation (2). More recently, research has focused on human papillomavirus type 16 (HPV16) an alpha group species 9 papillomavirus, and related species. This is because HPV16 infection causes anogenital lesions, in particular cervical dysplasias that in some rare cases can progress to cervical cancer. HPV16 and other ano-genital infective HPVs comprise the most common sexually transmitted disease in the world today. Twelve years ago the World Health Organization designated HPV16 a human carcinogen. It is now clear that the virus is the primary cause of around 60% of all cases of cervical cancer worldwide. The remaining 40% of cases are caused by other HPVs of the alpha group, for example, HPV18, 31, 33 and 45 (3) Cervical cancer is a major health problem in the world today, especially in the third world where cervical screening programs to detect HPV infection are not available to large sections of the population. It is the second most important cause of cancer deaths in women with nearly 300,000 deaths per annum and half a million people affected on an annual basis. Thus an in depth understanding of the replication cycle of these papillomaviruses is crucial in attempting to identify new targets for development of antiviral therapies. An HPV vaccine has been licensed and is being rolled out this year in some countries in the developed world but it will take up to 50 years to discover its full efficacy. This is a prophylactic vaccine and the plan is to vaccinate pre-pubescent girls so that they are protected from infection when they become sexually active. Meanwhile women who have not been vaccinated will already be infected and/or become infected and so be at risk of developing cervical lesions that may progress to cancer (4). Development of novel therapies is still very important to aim to reduce the level of infection in the population and so reduce the overall incidence of cervical lesions and cancer.

3. THE PAPILLOMAVIRUS LIFE CYCLE

In order to infect epithelial tissue papillomaviruses must gain access to cells of the basal layer, either the basal stem cells, or their daughter transit amplifying cells. This is because these are the only dividing cells in the epidermis (Figure 1) and the virus requires to establish infection by commandeering the host cell replication machinery to make copies of its own genome. The virus seems to infect tissue by falling down micro abrasions in the epithelium. It may attach to the receptor molecule heparin sulphate proteoglycan (5), perhaps with some involvement of alpha 6 integrin (6).

Immediately upon infection, the virus genome that has been deposited as an episome in the nucleus of an infected cell undergoes an initial amplification resulting in production of up to 100 genome copies. During a normal infection the virus persists in this manner in the basal layer of the epithelium, probably in the transit amplifying cells which can divide within the basal layer. This leads to maintenance of the infection as a reservoir of virus genomes within cell nuclei. Persistence of infection, over a long period of time, is thought to be a predisposing factor for tumor progression associated with HPV16. The virus may undergo a productive life cycle when a basal layer cell begins to differentiate and move to the upper epithelial layers. The productive life cycle results in expression of the virus early proteins, E1, E2, E5, E6 and E7, in both the basal and upper layers. The infected cells differentiate in the upper epithelial layers but due to low level expression of the virus oncoproteins E6 and E7 these differentiating cells retain the ability to enter S-phase. This capacity for cell cycle progression is essential for a productive life cycle so that the 100 copies of the virus genome present in transient amplifying cells that have begun to differentiate are amplified at least 10-fold (7). At this stage virus late gene expression has begun. The first and most abundantly expressed virus late protein is E4. It is expressed as a chimera with the first 5 amino acids of E1 attached through splicing onto the N-terminus of the E4 protein (E1^E4) (8). Expression of this protein commences in the spinous layer. Although encoded in the early region of the genome it is a late protein. E1^E4 is co-expressed as part of polycistronic mRNAs that also encode the capsid proteins L1 and/or L2 (9). The virus capsid proteins, once expressed, are imported into the nucleus of infected cells, where the virus genome has undergone vegetative viral DNA replication. Formation of mature virus involves encapsidation of episomal virus genomes. Virus particles are then release from the top of the epithelium from disintegrating cornified layer cells. One of the most remarkable features of the virus life cycle is restriction of expression of the virus capsid proteins to cells of the granular layer (Figure 1). When in contact with host immune surveillance the capsid proteins can raise a strong immune response. Restriction of their expression to the outermost layers of the epithelium reduces the possibility that upon expression the protein will trigger an immune response. Indeed, the current HPV vaccines are virus-like particles composed of these proteins.

4. HOW PAPILLOMAVIRUS GENE EXPRESSION IS REGULATED

The aim of a large body of work over the last 15 years has been to elucidate regulation of late gene expression of papillomaviruses in response to epithelial differentiation. Of particular interest is the nature of the cellular signals, and the *cis*-acting viral sequences that interact with these, that regulate the tight restriction of capsid protein expression to the cells of the granular layer of the epithelium. This section describes in very general terms how papillomavirus gene expression is regulated.

4.1. Transcriptional regulation

Bovine papillomavirus transcribes its early genes from a set of six promoters located in the early region (10, 11). For the alpha human papillomaviruses, early gene expression appears to be regulated from one or more constitutively active early promoters located in the E6 and E7 coding regions. For example, HPV16 uses the P_{97} promoter and HPV31, a similar promoter located at P_{99} (12-16). In a normal infection transcriptional activity of these promoters is weak and it is thought that they are negatively regulated by the virus replication/transcription factor E2 (17). Clearly, late gene expression is regulated transcriptionally by activation of one or more late-acting promoters in the virus genome. The key late promoter that is activated only in cells of the spinous layer and above is P_L for BPV1, P_{670} for HPV16 and P_{745} for HPV31. These promoters direct transcription through the early region before transcribing the late region downstream. For some papillomaviruses there is evidence of other downstream promoters that might direct synthesis of late transcripts but whether these are activated in a differentiation-dependent manner is not clear (9-11, 18). In differentiated epithelial cells, for BPV1, HPV1 and HPV31 there is good evidence that transcription read through of the early polyadenylation site is regulated in synchrony with terminal differentiation so that in undifferentiated epithelial cells mature late mRNAs are not fully synthesized (19-22). Presumably any incomplete RNAs in undifferentiated epithelial cells would undergo rapid nuclear degradation. So this is one way in which late region expression may be abrogated in inappropriate cell types.

4.2. Post-transcriptional regulation

One key observation made many years ago was that despite detection of expression of L1 proteins only in granular layer cells, RNAs encoding these proteins could be detected in the lower epithelial layers (23, 24). These experiments involved RNA fluorescence *in situ* hybridization of tissue samples infected with HPV 6, 11 and 33, three alpha group papillomaviruses. The probes used would detect unprocessed and processed virus RNAs and so could not differentiate between transcripts that may be destined for degradation and mRNAs that could be translated. For HPV16, there is evidence that in undifferentiated epithelial cells late gene transcripts are confined to the nucleus (25). To add to this story we demonstrated recently that in undifferentiated HPV16-infected epithelial cells the late RNAs that could be detected were nonpolyadenylated, pre-mRNAs that are very unstable (9). These data indicate that papillomavirus capsid gene expression is regulated, at least in part, at one or more post-transcriptional levels.

5. POST-TRANSCRIPTIONAL REGULATION OF GENE EXPRESSION

Post-transcriptional regulation of gene expression can involve regulation of any of the steps in the mRNA biogenesis and utilization pathways, for example, splicing (26), polyadenylation (27), nuclear export (28), RNA stability (29), and translation (30). The processes of RNA capping, splicing and polyadenylation occur co-transcriptionally and are physically linked via the carboxyl terminal domain (CTD) of RNA polymerase II (27,31,32). The CTD acts as a landing pad for RNA processing factors allowing coordination of RNA processing events. Addition of the 7 methyl guanosine cap stabilizes the RNA, is essential for export to the cytoplasm and for translation on the polysomes. Splicing removes introns from the pre-mRNA and splices together exonic coding regions. Splicing of alternative exons (alternative splicing) in mRNAs is emerging as a major means of production of multiple proteins from one RNA. Splicing and alternative splicing is regulated through binding of the splicing machinery and through binding of splicing regulatory proteins such as the SR and hnRNP factors (33, 34). Splicing is important for nuclear export of mRNA to the cytoplasm and for appropriate translation on the polysomes. In polyadenylation, poly (A) polymerase adds up to 200 adenosine residues to the cleaved 3' end of a pre-mRNA. Two major protein complexes, cleavage polyadenylation specificity factor (CPSF) that binds RNA upstream of the cleavage site and cleavage stimulatory factor (CstF) that binds downstream of the cleavage site are essential in priming the RNA for the cleavage event but the identity of the endonuclease itself is not yet known. Polyadenylation efficiency can be controlled through regulation of the affinity of these protein complexes for their binding sites by various *cis*-acting factors (35).

Upon transcription and RNA processing RNA molecules are formed into ribonucleoprotein particles (RNP) whose many protein components can interact in a dynamic fashion (36). One function of RNP formation is to maintain the mRNA in stable form. However, sometimes components of the RNPs can render an RNA stable or target it for decay, in both the nucleus and the cytoplasm. The mRNA decay pathway begins with removal of the poly(A) tail at the 3' end of the transcript. Then either the 5' cap is removed by Dcp1 and Dcp2 proteins and the body of the message degraded by the Xrn1 exonuclease or the message is degraded in a 3' to 5' direction by the exosome (29). In addition to RNP formation, fully processed mRNA molecules that leave sites of transcription are molecularly marked with proteins that reveal the various processing events that the RNA has undergone. One of the major protein depositions is the exon junction complexes (EJC) that define exon boundaries in mRNAs (37). EJC formation on RNAs is extremely important in determining, during translation on the polysomes, whether an mRNA contains premature termination codons and should be degraded or

whether it can be translated into full length protein (38). However, a myriad of other proteins have the capacity to regulate mRNA stability and translation in the cytoplasm. The general scenario is that *cis*-acting RNA signal sequences interact with cellular proteins or protein complexes to inhibit or to decrease or increase the efficiency of RNA processing, RNA stability and translation. *cis*-acting RNA sequences that interact with regulatory proteins can exist all along the mRNA molecule, not just at exon junctions. In general terms, there are many examples where the 5' UTR can contain signals that control mRNA export, stability and efficiency of translation; the body of the message can contain elements that regulate splicing, bind export proteins and proteins involved in mRNA stability/decay and translation while the 3' UTR can contain sequence motifs that regulate polyadenylation and splicing efficiency, mRNA stability/decay, and translation. Overall, the fate of an mRNA is dependent upon its nucleotide sequence and upon the changing patterns of proteins that it binds as it progresses from sites of transcription in the nucleus to translation on the polysomes in the cytoplasm.

5.1. Viruses and post-transcriptional regulation

In virus infection it is well known that post-transcriptional processes are targeted by the virus to optimize viral gene expression and to repress or abolish host gene expression. A good example of viral regulation of host mRNA processing involves the RNA-binding ICP27 protein of herpes simplex virus type 1 (HSV-1). ICP27 is an immediate early virus protein whose levels start to accumulate early in infection. At this stage the protein enters the nucleus to inhibit host cell splicing resulting in degradation of the host transcripts (39). This phenomenon is an essential portion of host cell shut-off engineered by HSV-1 that means that virus mRNAs do not have to compete with cellular mRNAs for export to the cytoplasm and for translation. Most HSV-1 RNAs have no introns making them immune to inhibition of splicing. Normally, intronless transcripts are not exported efficiently from the nucleus so HSV-1 has also to regulate nuclear export using ICP27 as a recruitment factor for key RNA export pathways such as TAP/REF and CRM1 (40). Most viruses, especially RNA viruses, express proteins that interfere with host cell translation. A good example of this type of post-transcriptional regulation comes from study of the *Picornaviridae* (41)**.** Picornavirus infection results in shut off of host cell translation due to cleavage of the translation initiation factor eIF4G that forms a scaffold for the main players in translation initiation, eIF4E and eIF4A. Cleavage means that the scaffold falls apart, the full translation initiation complex cannot be formed and cellular capped mRNAs can no longer be translated. Picornavirus mRNAs are not capped but instead contain an internal ribosome entry site in their 5' UTR that, in the absence of binding of the eIF4 factors, attracts in the 40S ribosomal subunit allowing initiation of transcription in a cap-independent manner.

Similarly, viruses can utilize the host cell RNA processing machinery to maximize efficient expression of their genomes. For example, HIV RNAs require either to remain unspliced, for production of new genomic RNAs or for translation into the major virus Gag-Pol protein precursor, or to undergo alternative splicing to yield the Env precursor and the virus accessory and regulatory proteins. The switch between no splicing and splicing depends on differential use of various splice acceptor and donor sites. This is regulated through *cis*-acting sequences in the RNAs but also through binding of cellular SR proteins, key proteins that regulate efficiency of splicing and use of alternative splice sites (42). The HIV genome possesses two long terminal repeats (LTR). The 5' 97 nt repeat region of both contains polyadenylation sites but to produce mRNAs only the 3' LTR polyadenylation site can be active. There are several mechanisms that allow the upstream site to remain silent and the downstream site to be recognized efficiently. The upstream site is repressed by its positioning close to the transcription initiation site in the LTR. Binding of the massive transcription initiation complex may inhibit recognition of the 5' polyadenylation signal or disallow binding of the polyadenylation machinery. Inhibition is also through close proximity of the 5' polyadenylation signal to a major splice donor site located only in the 5' LTR. U1snRNP binding is necessary for this inhibition so again, there is likely competition over space for binding of the early splicing complex and the polyadenylation complex (43). The molecular mechanism of inhibition has been worked out and is described in the next section. Finally, the 5' polyadenylation signal resides in a stem loop structure and stabilization of this structure has been shown to inhibit binding of polyadenylation factors (44). The mechanisms for enhancing use of the 3' polyadenylation site relies on an upstream sequence element (USE: see section 5.2) which binds the 160 kDa subunit of CPSF increasing efficiency of poly (A) addition (45).

5.2. 3' untranslated region-mediated gene regulation

Upstream sequence elements (USEs) that regulate polyadenylation efficiency have been delineated in 3' untranslated regions upstream of several human and particularly viral poly (A) signals. Some examples include human complement factor C2 (46), and human lamin B2 (47), adenovirus L1 (48), adenovirus L3 (49), ground squirrel hepatitis virus (50), cauliflower mosaic virus (51), SV40 (52,53), and HIV1 (45). These upstream sequence elements have little sequence or structural similarity, except that they are generally U-rich, and act in a position and orientation-dependent manner to regulate efficiency of 3' end processing. In some cases USE have been shown to exert their effect through binding of RNA processing proteins. The HIV1 USE recruits the core polyadenylation factor CPSF (45), the SV40 USE binds U1 snRNP protein A (U1A) which stabilizes the interaction of CPSF on the polyadenylation site (53), the complement C2 element binds polypyrimidine tract binding protein (PTB) that enhances RNA cleavage (54), and the lamin B USE recruits CPSF, among other proteins (47). Thus USEs often have an up-regulatory effect on 3' end processing by recruiting proteins that increase the efficiency of binding of the basal polyadenylation machinery. The effect is either direct, as in the case of the HIV1 USE or indirect, as in the case of the C2 USE.

However, USEs can also have inhibitory effects on 3' end formation. U1A inhibits its own expression by binding of two molecules of U1A to an USE, conserved in mammalian evolution, in the 3' UTR of the U1A transcript. This inhibition has been demonstrated *in vitro* and *in vivo* (55). As will be discussed below in sections 6-8, certain papillomaviruses contain USE in their late 3' untranslated regions that bind U1 snRNP. This binding is inhibitory of gene expression. These situations are complicated by the fact that certain elements and certain proteins can have both positive and negative regulatory effects on gene expression. For example, in experiments where the auxiliary splicing factor, $U2AF^{65}$ was tethered by means of an MS2 binding domain to an MS2 binding site located just upstream of an adenovirus L3 cleavage and polyadenylation signal poly (A) polymerase and polyadenylation was inhibited (56). In contrast, tethering $U2AF^{65}$ to the 3' UTR of a human β-globin gene enhanced 3' end formation (57). Similar differences have been observed with U1A as described above (55). In the case of autoregulation of U1A regulation is inhibitory of polyadenylation. Similarly, U1A inhibits poly (A) addition in the immunoglobulin M heavy chain mRNA by binding upstream of the poly (A) site (58). Surprisingly, U1A may also binds downstream of the same poly (A) site and contribute to the inhibition of expression reducing the efficiency of cleavage at the poly (A) site by interfering with binding of the polyadenylation factor CstF64 (59). Thus the location of the USE, or a downstream sequence element, with respect to the polyadenylation site may have a major bearing on the final activity of the protein bound. The theme of significance of location of an element regulating polyadenylation is carried on in results of experiments where U1 snRNP was engineered to allow binding to the terminal exon of several reporter genes. Such binding was postulated to inhibit poly (A) polymerase activity, i.e. poly (A) addition (60), in a similar manner to that demonstrated previously for the U1 70K component of U1 snRNP binding to the USE of bovine papillomavirus type 1 (61) (see section 6). However, U1 snRNP, when targeted to a downstream proximal site, is known to have a repressive effect on polyadenylation at the poly (A) site in the 5' LTR of HIV-1 (62). Mutations in U1 snRNA protein binding loops indicated that U1 70K was involved in this control mechanism which was likely inhibition of the cleavage event. In contrast, in experiments *in vitro* where U1 snRNP was bound at a 5'splice site downstream of the adenovirus L3 poly(A) site the cleavage reaction was clearly inhibited independent of interactions between U1A, U1 70K and poly (A) polymerase (63).

Many mRNAs, such as those of growth factors, lymphokines and cytokines involved in early growth response, have very short half lives. This is thought to be very important for their ability to respond rapidly to cellular signals. In contrast, other mRNAs, for example for housekeeping genes such as α and β tubulin, are very stable allowing translation of the proteins over a prolonged period of time. Thus, the relative stabilities of different mRNAs and how these respond to extra- and intra-cellular signaling can determine the phenotype of the cell. *Cis*-acting RNA elements that regulate stability are often located in 3'UTRs. One particular set of elements, the AU-rich elements (AREs) are the most extensively studied (64). There are three classes of AREs. Class I has the sequence AUUUA repeated 1-3 times and a U-rich region in close proximity, class II has a nonameric sequence UUAUUUA(U/A)(U/A) embedded in a U-rich region and class III has simply a U-rich sequence. All AREs have the ability to confer instability upon the mRNAs that contain them. ARE stimulation of mRNA decay initially involves deadenylation followed by exosome-mediated degradation (29). AREs bind a large number of *trans*-acting factors that modulate their effects. The mechanisms of action of three of these proteins, HuR, AUF-1 and tristetrapolin (TTP) have been worked out in some biological situations. Binding of the first two promote mRNA stabilization (65-68) while TTP can destabilize mRNAs, for example TNFα (69). HuR appears to stabilize such RNAs by inhibiting their degradation. For example, human vascular endothelial growth factor (VEGF) has a class III ARE in its 3' UTR that binds HuR with high affinity. In hypoxia, HuR overexpression efficiently stabilizes VEGF mRNA (70). Parathyroid hormone levels are regulated in response to calcium and phosphate. AUF1 binds elements in the 3'UTR of the parathyroid hormone mRNA that are known to regulate transcript stability under changing concentrations of calcium and phosphate (68). TTP has been shown to negatively regulate TNFα in mice *in vivo* via signaling action from the p38 MAPK pathway (71). In addition there is also emerging evidence that ARE-binding proteins may also regulate translation (71, 72).

An example of 3'UTR elements that regulate gene expression but are not AREs is the highly conserved iron-response elements found in the 3'UTRs of mRNAs encoding proteins involved in iron metabolism, for example the transferrin receptor and ferritin. These elements are around 30 nucleotides in length and have the capacity to form stem loop secondary structures. The loop has the sequence CAGUGN and the stem often displays a cytosine bulge. These structural features are important for binding the *trans*-acting iron response protein (IRP). IRP binding acts to stabilize the IRE-containing mRNAs that are otherwise subject to decay, by masking a cleavage site for an RNA endonuclease component of the exosome. Under conditions of low iron concentrations the IRP conformation allows it to bind to the 3'UTR IREs resulting in stabilization of the mRNAs (73).

In general, in post-transcriptional gene regulation a theme is emerging where RNA regulatory elements act in a sequence-specific or structure-specific manner to bind proteins that regulate various steps on the mRNA biogenesis and translation pathways. This is all in response to intra- and extracellular signal pathways. As discussed above viruses make good use of strategies to interfere with or utilize the host cell RNA biogenesis and translation machinery to maximize virus gene expression. Again a common strategy involves cellular proteins binding to viral RNA *cis*-acting regulatory motifs, some of which can have significant secondary structure, to either enhance virus RNA processing or translation. The discussion that follows centers on *cis*-acting regulatory motifs that have been

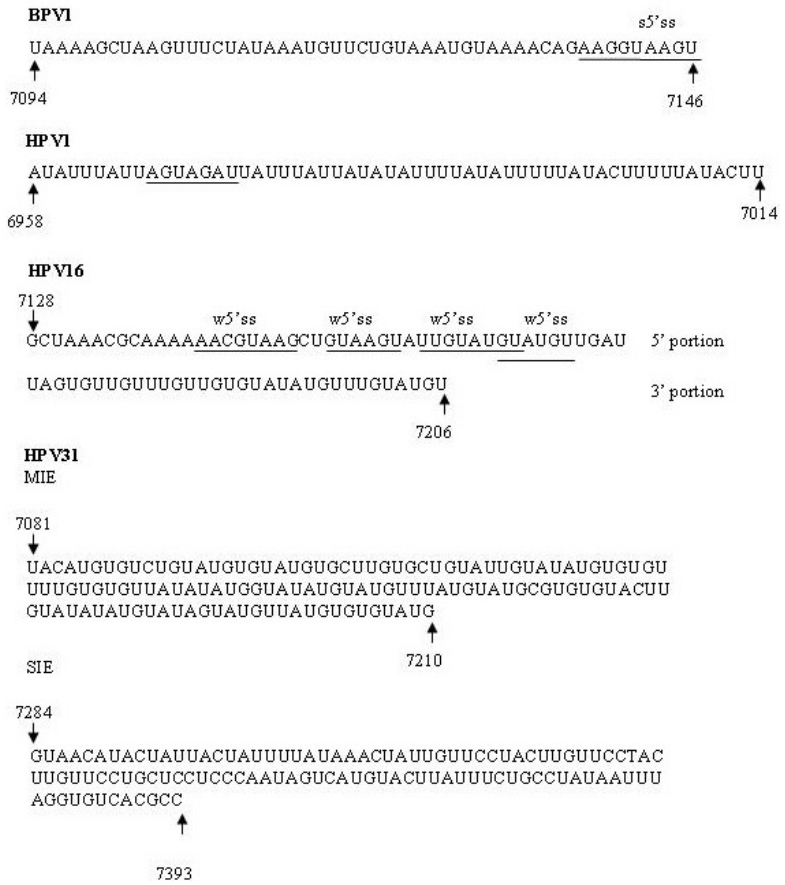

Figure 2. Sequences of the late 3' UTR regulatory elements from BPV1, HPV1, HPV16 and HPV31. Positions on the respective genomes are indicated. Strong (s) or weak (w) 5' splice site homologies are underlined for BPV1, HPV1 and HPV16. MIE: major inhibitory element, SIE: subsidiary inhibitory element.

discovered in the 3' UTR of a range of papillomaviruses, the proteins that they bind and the mechanisms by which they act, where this has been worked out.

6. THE PARADIGM: BPV1

An early observation for BPV1 late gene expression was that in BPV1-transformed C127 cells which are nonpermissive for virus late gene expression, and in fibropapillomas, late RNAs were not produced. This may be due to a switch off of the virus late promoter and a block in transcription elongation from other, early-acting promoters through to the late 3' UTR upstream of the late polyadenylation site (19). Transcripts terminated in this manner cannot be polyadenylated and so cannot be processed into mRNA. In 1991 Furth and Baker demonstrated the presence of a *cis*-acting RNA element, located in the 3' UTR of late mRNAs encoding the BPV1 capsid proteins that repressed gene expression (74). The study examined a 210 nt fragment containing the requisite signal sequences for the polyadenylation machinery to form on an RNA (i.e. AAUAAA, the polyadenylation signal, the cleavage site and the downstream GU-rich region) inserted downstream of a CAT reporter gene under control of an SV40 promoter. The discovery of the inhibitory element emerged from deletion mapping of the BPV1 late 3' UTR in the context of the report construct described above and also in the context of an expression vector for the BPV1 L1 gene. The inhibitory element was defined as 53 nts in length, sited upstream of the late poly (A) site taking up most of the short BPV1 late 3' UTR (Figure 2). Precise removal of the element increased levels of reporter mRNA by around 10-fold. It was noted that the element had the potential to form significant secondary structure. However, mutation analysis to alter the structure did not have any significant impact on its negative regulatory function (74, 75).

The observation of the impact of the BPV1 inhibitory element on cytoplasmic mRNA levels led to the hypothesis that the mechanism of action of the repressive element was at the level of mRNA stability (75). However, using the BPV1 L1 expression construct and measuring mRNA half life in actinomycin D-treated transfected NIH 3T3 cells little change in decay of mRNAs with or without the element was found. An alternative hypothesis was that the element might act as an upstream sequence element to negatively regulate RNA processing. Further deletion and site-directed mutagenesis studies revealed that the major *cis*-acting repressive sequence motif, AAG/GUAAGU, had extremely good homology with the consensus mammalian

splice donor site. This sequence motif would have the ability to bind the key splicing factor complex U1 snRNP. Interestingly, mutation of other sites upstream of the splice donor site made the element more inhibitory. One of these mutations disrupted the potential for stem loop formation in the element. So in theory, this stem loop could bind a protein that modulates activity of the negative element. Alternatively, the sequence changes introduced by mutation could result in recruitment of a cellular protein that does not normally bind the fragment and could regulate formation of a splicing complex on the splice donor site in the element. These hypotheses remain to be tested.

To prove further that the 5' splice site was key in the mechanism of action of the BPV1 inhibitory element transient transfection experiments were carried out where suppressor mutants of U1 snRNA were introduced into the BPV1-transformed C127 cells. U1 snRNA is the nucleic acid upon which U1 snRNP is formed. U1 snRNP is the complex that binds 5' splice donor sites via base pairing between the splice site and U1 snRNA to initiate splicing. Results demonstrated that base pairing of U1 snRNA to the splice donor site in the 53 nt BPV1 element was necessary and sufficient for its negative effect *in vivo*. To carry out splicing, splice donor sites must be paired with splice acceptor sites. However, none of the latter could be found downstream of the 53 nt fragment. Further, S1 nuclease analysis demonstrated that no mRNAs that utilized this splice donor site were produced from the virus genome. Thus the 5' splice site in the BPV1 3'UTR could not be active in a splicing event. Another possibility was that the element functions as an USE to regulate mRNA 3' end formation. Although usually these act in a positive manner, negative regulation could be exerted on polyadenylation by interfering with, rather than enhancing formation of the polyadenylation complex at the downstream polyadenylation site. A repressive role in binding of U1 snRNP was envisaged because there is good evidence that splicing factors can regulate polyadenylation both positively and negatively. This is probably because the processes of polyadenylation and splicing are physically linked through the CTD of RNA Pol II.

Elucidation of the mechanism of inhibition was achieved by Gunderson *et al.*, 1998 (61). This work was carried out entirely *in vitro* in HeLa nuclear extracts and so may not replicate the *in vivo* situation in undifferentiated bovine keratinocytes. However, this study showed that U1 snRNP bound to the inhibitory element repressed polyadenylation efficiency. The repressive component of U1 snRNP was found to be U1 70K and a direct interaction between U1 70K and poly (A) polymerase (PAP) was demonstrated. Further, this work showed that the motif in U1A, which was shown to autoregulate polyadenylation of its own mRNA was conserved in U1 70K (76). First of all, using *in vitro* polyadenylation assays the proposal of the Baker group that U1 snRNP bound to the BPV late 3' UTR 5' splice site inhibits polyadenylation was confirmed. Next it was shown that bovine PAP forms a complex with the BPV repressive element. U1 snRNP binds the repressive element and inhibits PAP activity. Since in vertebrates U1A can inhibit polyadenylation of its own mRNA (55) it was possible that this was the inhibitory component of U1 snRNP. Alternatively, it could be U1 70K, which has paired domains homologous to U1A. U1A and U1 70K proteins were purified *in vitro* and assembled onto wild type U1 snRNA or the RNA mutated in the U1A and U1 70K binding loops. The effect on activity of PAP was then tested in polyadenylation assays. U1 70K and wild type U1 snRNA yielded poor PAP activity in polyadenylation assays. However, efficient polyadenylation was achieved when U1 snRNA mutant in the U1 70K binding loop was used. U1A was not involved because the U1A-U1 snRNA complex did not inhibit polyadenylation. Finally, mutation of four U1 70K PAP inhibitory motifs, identified by comparison with U1A PAP inhibitory motifs, abolished the repressive activity of 70K on PAP. Subsequent studies investigating the precise mechanism of polyadenylation inhibition performed *in vivo* by transfection of HeLa cells or NIH3T3 cells confirmed the important role of U1 70K (60,62).

Interestingly, several other splicing related proteins, $U2AF^{65}$, Srp20 and Srp75 that are not part of U1 snRNP have very similar U1A and U1 70K inhibitory domains and would be predicted to have a similar function (61). These domains are serine-arginine-rich (SR) domains that are common in certain splicing factors (SR proteins). As mentioned previously, when tested using tethered function assays, these inhibitory domains were capable of inhibiting polyadenylation *in vivo* and specifically, *in vitro*, poly (A) polymerase (56). In contrast, the poly (A) polymerase interacting domain of $U2AF^{65}$ was able to enhance polyadenylation of the β-globin gene (57). The SR domains were specifically involved because mutation in these abrogated the regulatory effect.

The question remains whether these elegant mechanisms operate in BPV1-infected, non-transformed undifferentiated bovine epithelial cells. Although HeLa or NIH3T3 cells provide a useful tool for observing mechanisms of gene expression, they are highly transformed and highly culture adapted and it is clear that their protein profile differs from that of primary cells in tissues. It would be interesting to establish whether there were any changes in concentration of U1 snRNP components during epithelial differentiation or upon virus infection. Another layer of complexity lies in the actual numbers of the inhibitory element present in cells. At late times of infection viral DNA amplification will increase the numbers of virus genomes, as least 10-fold. If all these genomes were transcribed U1 snRNP could simply be titrated out in the differentiated keratinocytes leaving a large number of virus late transcripts that could be efficiently polyadenylated.

7. CUTANEOUS PAPILLOMAVIRUSES

To date only one human cutaneous papillomavirus, HPV1, has been studied in detail. Like BPV1, the HPV1 genome contains a late 3' UTR sequence that is inhibitory of gene expression when present in mRNA in HeLa cells (77). This element has also been shown to act at a post-transcriptional level but its

mechanisms of action and the proteins it binds are clearly different to the BPV1 element. The element was mapped to a 57 nt AU-rich and U-rich region (78) (Figure 2). This was done by observing its ability to repress expression of the HIV-1 $p17^{gag}$ gene, or the HPV1 L1 gene, both used as reporter genes in plasmid expression constructs. Interestingly, it was found that there was only a 4-fold reduction in cytoplasmic levels of RNAs containing the element. In contrast, there was a 120-fold reduction in protein levels compared to those from constructs where the element was deleted (77). The lower levels of cytoplasmic RNA in the presence of the repressive element could be due to inefficient nuclear export or reduced stability of the RNA in the nucleus or the cytoplasm. However, the very significant decrease in protein levels produced from RNAs containing the element indicates that it represses translation. In contrast, using a Vaccinia virus T7 RNA polymerase-based system for cytoplasmic expression of the $p17^{gag}$ reporter gene with or without the HPV1 repressive sequence downstream, it was shown that RNAs were equally stable and similar levels of protein were detected in each case. This indicates that cytoplasmic expression abrogates the repressive effect presumably by restricting access of the element to nuclear factors that cause nuclear retention and/or degradation of RNAs in which it is contained. Indeed, the repressive effect of the HPV1 3'UTR sequence could be reversed by coupling the reporter plasmid to an HIV-1 Rev protein/Rev responsive element (RRE) or a *cis*-acting simian virus constitutive transport element (CTE) system. Rev/RRE and CTE both promote nuclear export of unspliced HIV-1 mRNAs (79). The model proposed by Tan and Schwartz is that HPV1 mRNAs containing the repressive element are not able to access a suitable nuclear export pathway, resulting in their degradation. Providing an alternative nuclear export pathway, such as that used by Rev/RRE leads to increased cytoplasmic levels of the RNAs and efficient translation.

Like the BPV1 inhibitory element, the HPV1 repressive element contains a motif with reasonable homology to a 5' splice site. However, deletion of this motif had no effect on the repressive properties of the element and retention of the motif in the absence of the AU-rich element did not reveal an auxiliary repressive activity (77). This reveals a clear difference between BPV1 in the delta PV group and HPV1 in the mu PV group and may reflect important divergences in virus evolution and perhaps the adaptation to infection of fibroblasts and cutaneous epithelial cells respectively. Unfortunately, as yet there have been no studies on 3' UTR repressive elements in other papillomaviruses of the delta or mu groups. However, 5' splice site-like sequence have also been noted in the 3' UTRs of the alpha group papillomaviruses, HPV6B, HPV16 and HPV31 and the latter two are described in detail below.

The experiments described above indicated that the HPV1 repressive element potentially acted at a number of points on the mRNA biogenesis and translation pathway, for example, RNA stability, nuclear export and translation. A key observation was that insertion of the HPV1 ARE downstream of a CAT reporter gene reduced mRNA stability 4-fold. The repressive element contains two UAUUUAU motifs and three UAUUUUUAU sequences. RNA/protein binding studies revealed that the former motif binds HuR (see section 5.1) while the latter binds both HuR and hnRNP C1/C2. hnRNP C1/C2 proteins are expressed from the one gene but are alternative splice products differing by 13 amino acids. Like most hnRNP proteins they have been proposed to function in splicing, nuclear retention, transcript packaging and mRNA stability (80). Analysis of the binding affinities of HuR and hnRNP C1/C2 for wild type and functionally inactive, mutated HPV-1 AREs demonstrated conclusively that binding of these proteins correlates with the repressive effect of the element (81).

Intriguingly, in HeLa cells it was found that binding of nuclear HuR to the HPV1 ARE correlated with inhibition of gene expression. In contrast, binding of HuR to the element in the cytoplasm could up-regulate translation efficiency. Although HuR has a primary role in regulating mRNA stability there is good evidence that the protein can also regulate translation. This can be positively, in normal and damaged cells, in cap-dependent translation (72,82,83) and negatively, in IRES-dependent translation (84,85). The Schwartz group had noted that the HPV1 ARE appeared to enhance translatability of RNAs and ARE-mediated increased gene expression correlated with high levels of cytoplasmic HuR (86). Thus they proposed a role for HuR binding the HPV-1 ARE in control of translation. To investigate this hypothesis CAT reporter mRNAs containing the ARE or with the element mutated in the HuR binding sites were synthesized *in vitro*. Following capping and polyadenylation these RNAs were transiently transfected into cells and their translation monitored. Wild type ARE-containing mRNAs were poorly translated but translation was efficient when the HuR binding motifs were mutated (87). Further, it was demonstrate that the key translation factor, poly (A) binding protein (PABP), was capable of binding the HPV1 ARE *in vitro.* Recruitment of PABP could aid circularization of mRNAs for efficient translation or could direct loading of the mRNAs onto the polysomes. However, it remains to be seen whether PABP binding to the HPV1 ARE occurs *in vivo* in competition with the other proteins that bind the element. One further 55 kDa protein binds the ARE but it has not yet been identified (88).

It is clear that the HPV1 3' UTR repressive element is a member of the large class of AREs present in vertebrate mRNAs that regulate mRNA stability. The best known of these is that present in the 3' UTR of the c-*fos* gene whose protein product, together with Jun protein, makes up the AP-1 family of transcription factors. Indeed, the HPV1 ARE displays some similarities with this element and the mRNAs may be processed in a similar manner in the cell. Interestingly, it has been noted that levels of c-*fos* increase during epithelial differentiation in the same way as levels of HPV1 late mRNAs (78). Perhaps, if these RNAs are co-regulated, there is some functional connection between c-*fos* protein activities and late events in the virus life cycle. It is fascinating that this papillomavirus may have evolved to utilize a common

cellular RNA regulatory pathway for precise regulation of expression of its late genes. Moreover, the difference in the mechanisms of action of the HPV1 and BPV1 repressive elements is very striking.

8. MUCOSAL PAPILLOMAVIRUSES

3' UTR late regulatory elements (LRE) have been studied in two highly related mucosal papillomaviruses, HPV16 and HPV31, the former in much more detail than the latter. Both regulatory elements are longer and more complex than the BPV1 and HPV1 elements and the HPV31 element is bipartite and particularly complex. The HPV16 element overlaps the 3' end of the L1 coding region and extends into the 3' UTR. The HPV31 element begins in the 3' UTR and extends much further than the HPV16 element to beyond the late polyadenylation site. An early analysis of polyadenylation signals in the HPV16 late 3'UTR revealed not only three putative polyadenylation signals but a negative regulatory element located upstream (89). The element was discovered due to its repression of CAT reporter gene expression in transient transfection studies designed to test function of the three polyadenylation signals. The boundaries of the element were determined subsequently (90, 91) giving an element 79 nts in length. The element can be divided into two distinct regions according to the nucleotide sequence. The 5' portion of 47 nts contains 4 weak 5' splice sites (75) while the 32 nt 3' portion is U-rich (90) (Figure 2). Functional studies on the element indicated that both portions were required for its repressive activity, at least in HeLa cells (91). While deletion of the entire 79 nt element reduced reporter gene expression 40-fold, the 5' portion alone reduced expression by 10-fold in agreement with a previous study (75) and the 3' portion by only 3-fold. Analysis of the contribution of the four weak 5'splice sites in the 5' portion of the element revealed that all four were required for full repression when assayed in a reporter gene construct with the 3' portion of the LRE deleted. However, mutation of these only partially alleviated repression in the context of the entire 79 nt element (91). Site-directed mutagenesis was carried out changing five nucleotides in turn across the element yielding 15 LRE mutants which were cloned into the reporter gene construct. None of these simple mutations alleviated the repressive effect. This confirmed that the length and complexity of the element may be essential for function.

The HPV31 element was initially identified by homology with the HPV16 LRE but turned out to be significantly more complex. This is despite close homology (70% sequence identity) between these two alpha genus, species 9, papillomaviruses (92). An HPV31 element with significant sequence and structural homology to the HPV16 LRE was delineated to 101 nts and termed the negative regulatory element-like element (NLE). The first 60 nts of the NLE display 68% identity with the HPV16 LRE as expected. The region contains three weak consensus 5' splice sites (compared to four in HPV16) and a 3' GU-rich region that is more extensive than the similar sequence motif in HPV16. However when the 101 nt NLE was deleted from the HPV31 3'UTR in a reporter gene construct and tested for alleviation of repressive activity, only a 2-3 fold increase in gene expression was obtained compared to the 20-40-fold increase observed when the HPV16 LRE was deleted in a similar reporter gene expression plasmid. 3' UTR deletion and mutation analyses revealed that the HPV31 NLE major inhibitory element (MIE) was in fact 130 nts in length. Surprisingly, a 110 nt subsidiary inhibitory element (SIE) was also discovered which when deleted gives a three-fold stimulation of gene expression (92) (Figure 2). This is located downstream of the late polyadenylation site and the GU/U-rich downstream sequence element that regulates polyadenylation efficiency.

8.1. Proteins that bind the elements

Identification of cellular proteins that bind the HPV16 LRE and the HPV31 MIE/SIE was carried out in HeLa and W12 cell extracts. HeLa cells are highly transformed but are phenotypically basal cervical epithelial cells and so can provide a convenient model system for investigating protein/RNA interactions. However, for HPV16 a much better system is the W12 cell system. W12 cells are cervical basal epithelial cells that are immortalized but not transformed. The cell line was established from a biopsy from a patient with a low grade cervical lesion (93). The 20863 clone of these cells (94) contains around 100 copies of the HPV16 genome. These copies are maintained as episomes, and little integration is observed, if the cells are grown at low passage (<p18). The most useful property of these cells is that they can be induced to differentiate to cells that express markers of granular epithelial cells, amplify virus genomes and produce virus late proteins, including the major capsid protein L1 (9, 95). The CIN612 9E HPV31-positive cell line has similar properties (18). Using HeLa nuclear extracts in electrophoretic mobility shift assays (EMSA) and UV-crosslinking studies a number of protein complexes and individual proteins were found to bind the HPV16 LRE and a subset of these also bind the HPV31 MIE and SIE elements. Indeed, RNA probes for each element can compete for protein binding in EMSA (92). For HPV31 the MIE and SIE can also compete with each other for binding a subset of cellular protein complexes. This may indicate that the SIE, which would not be part of a stable mRNA following RNA processing, as it is downstream of the poly (A) site, might help load important cellular factors onto the upstream MIE. Alternatively, the SIE could act as a downstream sequence element to sterically hinder polyadenylation complex formation or to inhibit the cleavage step in 3' end formation. Using W12 and CIN612 extracts we have confirmed binding of these protein complexes and identified a number of components as described below (Table 1). However, it is worth noting that the relative proportions of the proteins that bind differ from the profiles obtained with HeLa cells (Cumming and Graham, unpublished data).

The HPV16 regulatory element displays four weak consensus 5' splice sites. So we set out to prove that they bound U1 snRNP in a same manner as the BPV1 inhibitory element. EMSA studies indicated that a major large protein complex bound these elements specifically

Table 1. Proteins that bind papillomavirus late 3' UTRs

Virus element	Proteins/complexes known to bind
BPV1 inhibitory element	U1 snRNP
HPV1 AU-rich element	HuR
	hnRNP C1/C2
	p55 (identity unknown)
	PABP
HPV16	U1 snRNP
	U2AF
	SF2/ASF (indirect binding)
	HuR
	hnRNP A1
	CstF-64 (binding is weak)
HPV31	HuR
	U2AF
	CstF-64

and with high affinity (25, 91, 92). UV crosslinking revealed two major proteins of around 60 and 50 kDa and a range of other larger and smaller proteins at around 80, 70, 35, 30 and 20 kDa (90, 91). EMSA studies with HPV16 LRE wild type and 5' splice site mutant RNA probes demonstrated that binding of most of these proteins was significantly reduced when all four splice sites were mutated such that they would be functionally inactive (91). In addition, EMSA showed that the major large RNA/protein complex that forms upon the LRE was also significantly depleted on this mutated RNA probe. Finally, antibody supershift EMSA proved that Sm proteins, key functional components of snRNPs, bound the LRE and affinity chromatography showed binding of major U1 snRNP components, Sm proteins (20-30 kDa), U1A (35 kDa) and U1 snRNA to the wild type LRE but not to the 4x5' splice site-mutated LRE. Interestingly, we were unable to detect binding of U1 70K to the LRE possibly ruling out the prospect that the LRE has a mechanism of action equivalent to the BPV1 inhibitory element. It is likely, though not yet proven, that U1 snRNP binds the HPV31 MIE because protein complexes we have identified as containing U1 snRNP also bind this related element.

The auxiliary splicing factor, U2AF was one of the first proteins demonstrated to bind the HPV16 LRE (90). U2AF, a dimer of 35 and 65 kDa subunits, binds a polypyrimidine tract upstream of 3' splice acceptor sites and aids loading of U2 snRNP that marks the 3' end of introns for splicing. Proteins were affinity-selected from HeLa extracts by binding to riboprobes synthesized from PCR-amplified templates homologous to 3' portion of the LRE. A major 65-70 kDa protein was isolated. The GU-rich 3' LRE had sequence similarity to a cognate binding site for the U2AF 65 kDa subunit. So U2AF was depleted from HeLa extracts by incubation with poly (U) sepharose to which the protein binds with very high affinity. Then $U2AF^{65}$ expressed in bacteria was added back to the extracts. Increased LRE RNA/proteins complexes were observed with increasing amounts of $U2AF^{65}$ proving that this protein binds the element. Moreover, mutations in the GU-rich region designed to knock-out binding of the protein reduced binding of the 65 kDa band. Subsequent studies demonstrated binding of U2AF by antibody depletion EMSA (25) and affinity chromatography (95). As expected the HPV31 inhibitory element also binds U2AF through its GU-rich region (92).

The key early splicing complex that forms across splice junctions in pre-mRNA comprises U1 snRNP, U2AF and a bridging SR protein, such as SF2/ASF, to modulate activity of the complex (Figure 3). As we had found that the 5'portion of the LRE bound U1 snRNP, the 3' portion binds the auxiliary splicing factor U2AF (25,90) and both the 5' and 3' portions of the element were required for full activity (91) we determined whether the LRE was capable of binding a full early splicing complex. Co-immunoprecipitation, EMSA and affinity chromatography demonstrated that SF2/ASF bound the LRE indirectly through interaction with $U2AF^{65}$ as predicted (95). We do not yet know if SF2/ASF also binds the HPV31 element. Thus a major protein complex that binds the LRE makes it resemble not just a 5' splice site as in the case of the BPV1 inhibitory element but juxtaposed 5' and 3' splice sites as would occur in splicing during intron removal.

Finally, the polyadenylation factor cleavage stimulatory factor (CstF) 64 has also been shown to bind the HPV16 and 31 inhibitory elements directly using similar techniques employed to demonstrate binding of U2AF (25,92). This protein is one of three subunits of CstF that binds GU-rich regions downstream of polyadenylation sites as part of the 3' end formation complex (27). One theory is that if it bound these 3' UTR elements proper formation of the polyadenylation complex may be inhibited. This is because CstF would be positioned upstream of CPSF, the second essential polyadenylation factor complex, rather than downstream. However, it seems that binding of CstF-64 to the HPV16 LRE may be weak as its binding can easily be completed by other LRE-binding proteins (25).

Although all these proteins have been shown to bind the LRE, direct evidence that they regulate RNA processing of virus late transcripts has been elusive to date. Overexpression or siRNA knockdown studies are required to demonstrate such an effect in conjunction with use of LRE mutations in the 5' splice sites and GU-rich regions. Previously, it has been hard to carry out these experiments due to difficulties transfecting epithelial cells, especially differentiated epithelial cells. Improvements in protocols utilizing lentivirus vectors for delivery of small hairpin RNA should make this much easier in the future.

8.2. Mechanisms of action of the HPV16 LRE – RNA processing.

As in the case of the BPV1 inhibitory element, no 3' splice site can be identified downstream of the LRE 5' splice site sequences and RT-PCR did not reveal any transcripts arising from this region (unpublished data). However, as discussed above it is clear that the LRE alone can mimic juxtaposed 5' and 3' splice sites. In contrast to the BPV1 element the HPV16 LRE did not decrease polyadenylation efficiency *in vitro* (89) nor *in vivo*, at least in HeLa cells extract (unpublished data). So instead of U1 snRNP/LRE binding regulating polyadenylation, an alternative role of the HPV16 LRE in post-transcriptional regulation of gene expression could be that it modulates terminal exon definition, as proposed by Furth *et al.*, 1994. In RNA processing exons are normally defined by splice

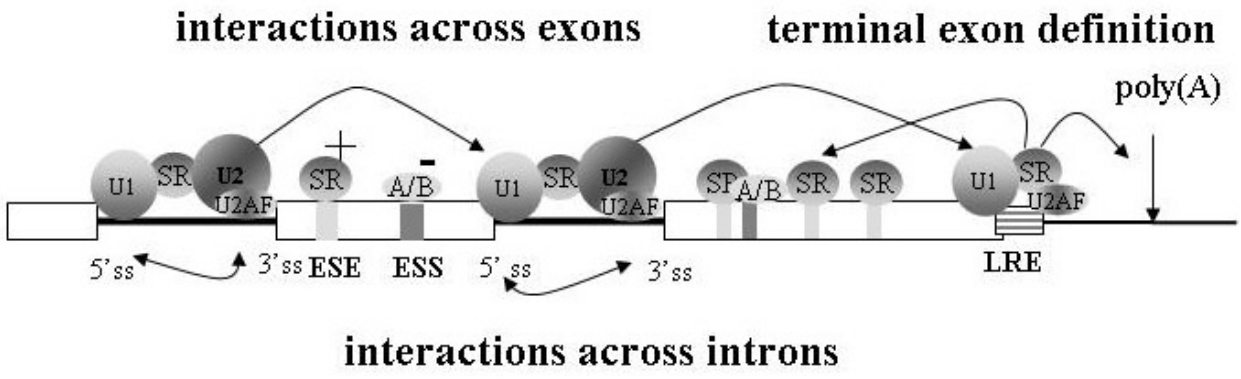

Figure 3. A schematic diagram of possible protein/protein and RNA/protein interactions involved in exon definition. Splice complexes interact across introns in splicing. However, they also interact across exons to define these. Exon definition is enhanced or silenced by exonic sequences binding SR proteins or hnRNP A/B proteins respectively. The possible mechanisms of action of the HPV16 LRE are indicated as discussed in the text. U1: U1 snRNP, U2: U2 snRNP, SR, serine/arginine-rich protein, U2AF: dimer of $U2AF^{35}$and $U2AF^{65}$, A/B: hnRNP A or hnRNP B protein, lined box, LRE: HPV16 late regulatory element, poly(A): polyadenylation site, 5'ss: 5' splice donor site, 3'ss: 3' splice acceptor site, open boxes, exons, thick lines: introns, thin line: 3' UTR, light grey boxes: exonic splicing enhancers, dark grey boxes, exonic splicing silencers, +: positive regulatory interaction, -: negative regulatory interaction. Arrows indicate proposed regulatory interactions (34,96,112).

complexes forming at the 3' and 5' exon/intron junctions. This is regulated positively and negatively by SR proteins and hnRNP proteins binding exonic sequence enhancers and silencers respectively (96). The 5' end of the 5'-most exon in an RNA is defined by proteins binding the cap complex. The 3' most exon is defined by the polyadenylation complex. These 5' and 3' protein complexes then interact with splice complexes across the terminal exons to define them (Figure 3). However, it is clear that often terminal exons require additional *cis*-acting sequences, and the proteins they bind, for efficient splicing. For example, U1 snRNP binding to a good consensus 5' splice site in the 3' UTR of an mRNA leads to poor terminal definition. In contrast, U1 snRNP binding to a weak 5' splice site links splicing and polyadenylation of the message and efficient terminal exon definition (97). Terminal exon definition may be very important in papillomavirus gene expression because the L1 coding region is around 1.5 kb in length while the mean length of a human 3' terminal exon is only around half this length (98). The four weak consensus 5' splice sites located in the 5' portion of the HPV16 LRE, two of which are located within the L1 coding region, could bind U1 snRNP and bridge to the 3' splice site at the 5' end of the L1 coding region resulting in an increase in splicing efficiency of the L1 exon. For BPV1 the inhibitory element is fully within the 3' UTR and is much closer to its polyadenylation sites in comparison to the location of the HPV16 element. This proximity could mean that the BPV1 element, with its single strong 5' splice site binding U1 70K, acts mainly on the polyadenylation process, rather than on splicing efficiency.

8.3. Mechanisms of action of the HPV16 LRE – RNA stability and translation.

The U-rich portion of the LRE resembles a class III HuR binding site and HuR was shown to bind the element using HeLa cell extracts (25). Recently we have demonstrated HuR binds the LRE 3' portion directly in W12 cell extracts. Interestingly, binding was much stronger in differentiated as opposed to undifferentiated W12 cells (manuscript in preparation). To establish a function for HuR binding the protein was overexpressed in undifferentiated W12 cells and siRNA used to knock out expression in differentiated cells. Overexpression resulted in induction of L1 protein expression in the basal epithelial cells and siRNA knock-down of HuR reduced L1 expression when the cells were differentiated. This suggests that HuR is a major regulator of HPV16 late gene expression, repressing late gene expression in undifferentiated epithelial cells but enhancing it when the cells differentiate.

Our results lead to speculation about the role of the splicing complex that forms upon the LRE and any competition with HuR. Although we have been unable to demonstrate a role for the LRE and the proteins complexes it binds in inhibiting polyadenylation it is tempting to suggest that the splicing complex forms upon the LRE as it is synthesized in undifferentiated cells. Such complex formation could restrict access or activity of poly (A) polymerase leading to transcripts that are not fully processed (as we have demonstrated (9)) and are therefore degraded in the nucleus. Splice complex affinity for the LRE is high in undifferentiated epithelial cell extracts so perhaps HuR cannot compete for binding at this stage of cell differentiation. The splicing complex appears to form in both undifferentiated and differentiated W12 cells. However, in the latter the complex may be different in that its SF2/ASF component is hyperphosphorylated perhaps leading to altered complex formation (95). In differentiated cells, such an altered splicing complex could form transiently on the LRE to direct efficient terminal exon definition during co-transcriptional RNA processing then dissociate upon completion of mRNA synthesis. Thus when the late mRNAs leave sites of transcription HuR could gain

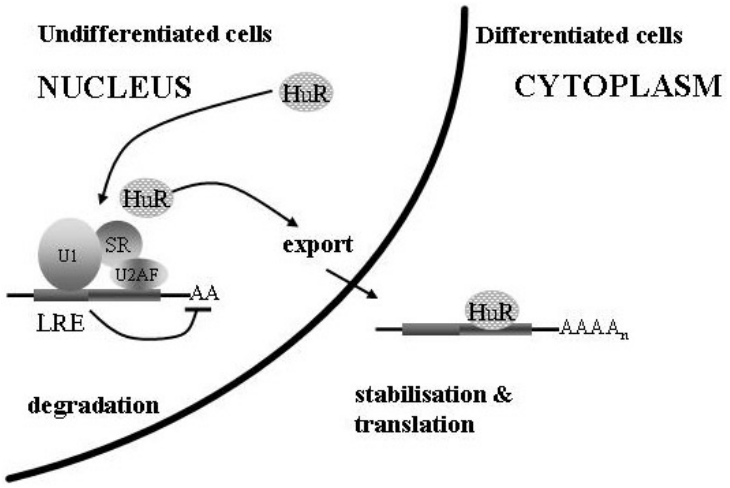

Figure 4. A model for the mechanism of action of the HPV16 late regulatory element (LRE). In undifferentiated epithelial cells late, nonpolyadenylated RNAs can be detected but these are unstable and late mRNAs are not produced and not exported to the cytoplasm. This may be due to the early splicing complex (U1 snRNP, U2AF and SF2/ASF) binding the LRE and inhibiting polyadenylation. When cells differentiate HuR is somehow free to access the LRE and the splicing complex cannot compete with HuR for binding. HuR stabilizes the late transcripts and facilitates export to the cytoplasm for translation.

access to the LRE to stabilize and export them to the cytoplasm and maintain their stability in the cytoplasm (Figure 4). Indeed, HuR binding to the LRE is most efficient in the cytoplasm of differentiated W12 cells (manuscript in preparation). Determining how the splicing complex changes with respect to binding the LRE in differentiated virus-infected epithelial cells will be a major step forward in elucidating mechanisms of action of this element in a differentiation stage-specific manner.

Finally, the LRE also binds hnRNP A1 in the nucleus of differentiated W12 cells directly via a good consensus binding site spanning the 5' and 3' portions of the LRE (Cheunim *et al*, in press). hnRNP A1 is the antagonistic counterpart of SF2/ASF in splicing but it can also regulate nuclear export, mRNA stability and translation (99). Whether this protein can bind the LRE in competition with the splicing complex and HuR protein is not yet known. The role of hnRNP A1 in regulation of HPV16 late gene expression is currently being elucidated. hnRNP A1 has already been shown to bind an exonic splicing silencer within the L1 gene and it clearly has a role in regulation of splicing events in the virus late region (100). hnRNP A1 binds cooperatively to exons in a 3' to 5' direction and SF2/ASF blocks this binding (101-103). Perhaps hnRNP A1 is recruited to the LRE as a donation site for loading onto the L1 exon in competition with SR proteins. The distance between the LRE and the L1 coding region hnRNP A1 binding sites does not preclude their cooperation: although the cooperative mechanism has not been worked out as yet, it is clear that hnRNP A1 binding sites need not be adjacent in exons and can indeed be some distance apart (104). Levels of hnRNP A1 and SF2/ASF are known to be essential for certain steps in splicing and for alternative splicing (105). These increase in concert in the differentiated, infected epithelial cell consistent with their having a role in regulating virus late gene expression.

9. EVOLUTION OF REGULATORY ELEMENTS

It is becoming apparent that papillomaviruses of several genera and species contain late region 3' UTR regulatory elements and a number of these have gene regulatory activity (88). Apart from the examples described in detail above, two cutaneous "low risk" types, HPV1 (genus mu) and HPV2 (genus alpha, species 4), one mucosal "low risk" type, HPV6b (genus alpha, species 10), HPV41 (genus nu) and HPV61 (genus alpha, species 3) have been shown to contain such elements. Presumably, many other papillomaviruses have regulatory 3' elements that help control appropriate and efficient expression of the virus capsid proteins. This is because they all complete their replication cycle in differentiating epithelia and have to suit regulation of their gene expression to respond to signals from the host cell. In addition, all such elements have been demonstrated to act at one or more post-transcriptional levels to inhibit late gene expression in undifferentiated epithelial cells where expression of the late proteins would abrogate the infectious life cycle. However, each element may utilize a different range of post-transcriptional mechanisms to achieve appropriate late gene expression (Figure 5). Indeed, those examined in detail so far bind different sets of proteins (Table 1). Even comparing the regulatory elements from HPV16 and HPV31, two highly related papillomaviruses (both genus alpha, species 9) it is clear that although there is considerable overlap in the factors they bind, a subset of RNA-binding proteins is unique to each (92). The BPV1 inhibitory element and the HPV16 element both bind U1 snRNP. For BPV1, the mechanism is clear; this RNA/protein interaction inhibits polyadenylation but this is not the case for the HPV16 element. For HPV1 the mechanism also seems clear. This repressive element may act as a classical ARE to destabilize late RNAs in undifferentiated epithelial cells and/or to stop efficient nuclear export of any RNAs synthesized in undifferentiated

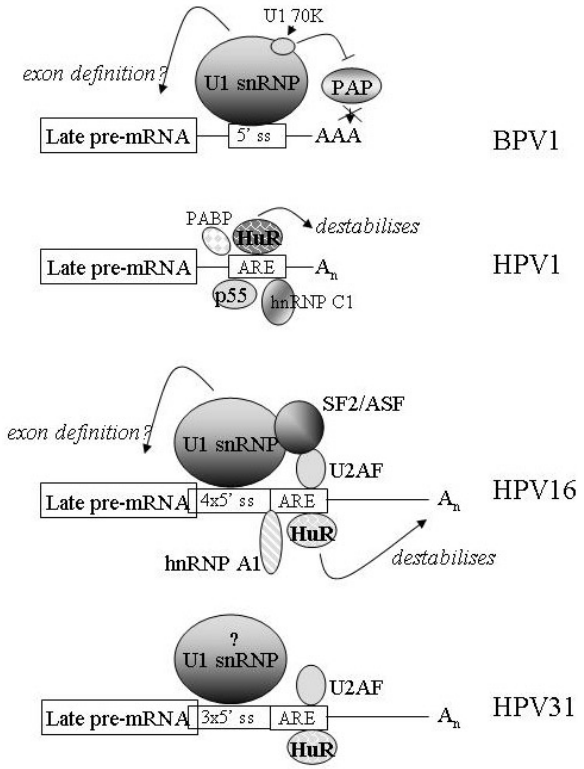

Figure 5. Known RNA/protein interactions on the BPV1, HPV1, HPV16 and HPV31 late regulatory elements and possible mechanisms by which they may act. Open boxes: late pre-mRNAs, stippled boxes, late regulatory elements, AAA: immature poly (A) tail, A_n fully formed poly (A) tail, ARE: AU-rich element, 5' ss: 5' splice site, PAP, poly (A) polymerase.

cells. Its effect in differentiated epithelial cells has not been examined. The HPV16 element, through binding HuR can have a positive and negative effect depending upon the differentiation status of the epithelial cell. We hypothesize that in undifferentiated epithelial cells it destabilizes late RNAs while in differentiated epithelial cells it stabilizes late mRNAs and/or increases their translation efficiency.

The HPV16 LRE is more complex than the BPV1 or HPV1 elements but displays features similar to both. Like BPV1, the HPV16 LRE binds U1 snRNP, although it is not a complete complex as it appears to lack U1 70K (91). We have been unable to prove that the HPV16 LRE inhibits polyadenylation, at least using HeLa cell extracts. It might be reasoned that this is due to lack of U1 70K in the complex. However, the complex does contain two proteins from the same SR-domain-containing family, U1A and U2AF^{65} (61). We have shown indirectly that late RNA polyadenylation is inhibited in undifferentiated W12 cells (9) and may need to revisit the *in vitro/in vivo* polyadenylation experiments using cell extracts from undifferentiated W12 cells, instead of HeLa cells, to investigate if these proteins do indeed inhibit polyadenylation. Furth *et al.*, 1994 argued that U1 snRNP binding to the BPV1 inhibitory element might regulate terminal exon definition with U1 snRNP present. The longer and more complex HPV16 LRE can accommodate a full early splicing complex that might be predicted to more efficiently promote terminal exon definition (106) but this has not yet been investigated. The LRE also bears similarities to the HPV1 ARE. It is not a classical ARE but does contain significant U-rich regions that bind HuR in a similar manner to the HPV1 ARE. It is likely that they operate by the same mechanisms. It is not yet understood why the HPV16, and to a great extent, the HPV31 elements have evolved to be much more complex, utilizing a series of different RNA regulatory mechanisms to achieve correct differentiation stage-specific late gene expression. One explanation could be that, in contrast to BPV1 and HPV1, these viruses complete the late stage of their life cycle in the very uppermost layer of the epithelium (8) and tighter control of gene expression is required to achieve this restriction and to efficiently express the capsid proteins in relatively few cells. It will be of interest to examine other cervical mucosa-infective human papillomaviruses to discover if these also contain complex LREs.

10. PROSPECTS FOR ANTIVIRAL INTERVENTIONS

Although many countries have now licensed the anti-HPV16/18 subunit vaccines for delivery there is a major cohort of women who are infected and may develop cervical lesions in the future. In addition, although it is hoped that the vaccine will be very effective, it will be many decades before this can be properly assessed and it is possible that the vaccine uptake may be lower than hoped for. Thus, development of new anti-viral therapies is required. All papillomaviruses studied so far have late inhibitory elements, and there seems to be a theme of key known RNA-binding proteins that interact with these elements to mediate their function. So it may be possible to design RNA-based drugs that would be active against a range of papillomaviruses to disrupt RNA/protein interactions (107). Such drugs could be aptamers, which are chemically stabilized RNAs selected *in vitro* to bind proteins with extremely high affinity, for example HuR. If these could be formulated to be topically applied to the epithelia surface of an HPV16-iduced lesion it may be possible to reduce L1 protein expression and so decrease the production of virus particles from the lesion. Topical application of preparations to inhibit virus gene expression events would also be deleterious to host gene expression. However, this may not be a significant problem because differentiated cells in the uppermost epithelial layers are beginning to shut down nuclear function (108). It remains to be seen if disruption of late events in the virus life cycle may have feedback effects on earlier events, or indeed have a deleterious effect, for example by tipping the balance towards integration of virus genomes into the host genome.

Alternatively the drug could be a small molecule selected from a combinatorial library or from collections of compounds from pharmaceutical companies to bind the LRE with high affinity in a similar manner to what has been achieved for HIV1 Tat/TAR RNA interaction (109,110). Such binding should inhibit complex formation upon the inhibitory element and so abrogate its function.

The papillomavirus late regulatory elements have the potential to form significant secondary structure. This is not involved in any of the regulatory mechanism because targeted disruption of the structures had no effect on reporter gene expression (75, 77, 91). This sort of structure would be ideal for selection of small molecules that interacted with high affinity and that may block binding of late regulatory element-binding proteins. Such small molecules or compounds could be systemically or topically applied. More recently work has focused on the possible uses of siRNAs, miRNAs and shRNAs in drug discovery. There is evidence that these small RNAs can gain access to cells and retain function (111). Our recent observations of the positive and negative effects of modulation of HuR proteins concentrations in cells on HPV16 L1 expression open the way for rational design of such therapies.

11. PERSPECTIVES

It is 17 years since the first report of a papillomavirus late regulatory element (89). In the intervening time much has been elucidated regarding the structure of such elements, the proteins they bind and the mechanisms of action they exert on gene expression. One of the major problems in carrying out such studies is lack of availability of good culture systems for studying papillomaviruses due to their complete reliance on a differentiating epithelium for their replication cycle. Most of the experiments to date have been carried out in transformed cell lines, such as HeLa cells. Of course, papillomaviruses infect normal basal epithelial cells and complete their life cycle in fully differentiated keratinocytes. So it could be argued that studies to date, although very informative, may have missed important information. The use of cell lines such as W12 and CIN612 where the cells can be differentiated in monolayer and organotypic raft culture has started to reveal nuances of the differentiation stage-specific LRE-responsive regulation not previously observed in lines such as HeLa cells. Charting the profiles of proteins that bind the various elements in undifferentiated and differentiated normal keratinocytes and the pattern of expression of the RNA-binding proteins in normal tissue, infected tissue and cancer tissues will prove extremely useful in future studies.

12. ACKNOWLEDGEMENTS

I would like to thank previous members of my research group, especially Maria McPhillips, Sarah Cumming and Thanaporn Cheunim, for their input into the theories we have developed and tested over the last 10 years. I am grateful to Professor Saveria Campo for reading and commenting on the manuscript. I would like to dedicate this review to the memory of my late colleague Professor Barklie Clements who gave me the chance to research RNA-based mechanisms of gene expression in human papillomaviruses. Our studies were funded by the BBSRC and the Wellcome Trust.

13. REFERENCES

1. Munoz,N., Bosch,F.X., de Sanjosé,S., Herrero,R., Castellsague,X., Shah,K.V., Snijders,P.J.F., Meijer,C.J. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.*, 348, 518-527 (2003)

2. Howley,P.M. In Fields,B.N., Knipe,D.M., Howley,P.M. (eds), Fields Virology. Lippincott-Raven Publishers, Philadelphia, Pa., pp. 2045-2076 (1996)

3. Bosch,F.X., Manos,M.M., Munoz,N. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J. Natl. Cancer Inst.*, 87, 796-802 (1995)

4. Stanley,M.A. Human papillomavirus vaccines. *Rev. Med. Virol.*, 16, 139-149 (2006)

5. Giroglou, T., Florin, L., Schäfer, F., Streeck, R. E., Sapp, M. Human papillomavirus infection requires cell surface heparan sulfate. *J. Virol.* 75, 1565-1570 2001.

6. Evander M., Frazer, I. H., Payne, E., Qi, Y M., Hengst, K., McMillan, N. A. Identification of the alhpa6 integrin as a candidate receptor for papillomaviruses. *J. Virol.* 71, 2449-2456. (1997)

7. Longworth,M.S., Laimins,L.A. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol. Mol. Biol. Rev.*, 68, 362-372 (2004)

8. Peh,W.L., Middleton,K., Christensen,N., Nicholls,P., Egawa,K., Sotlar,K., Brandsma,J., Percival,A., Lewis,J., Liu,W.J., Doorbar,J. Life cycle heterogeneity in animal models of human papillomavirus-associated disease. *J. Virol.*, 76, 10411-10416 (2002)

9. Milligan, S. G., Veerapraditsin, T., Ahamat, B., Mole, S., Graham, S. V. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* 360, 172-181 (2007)

10. Baker,C.C., Howley,P.M. Differential promoter utilization by the bovine papillomavirus in transformed cells and productively infected wart tissues. *EMBO J.*, 6, 1027-1035 (1987)

11. Choe,J., Vaillancourt,P., Stenlund,A., Botchan,M.R. Bovine papillomavirus type 1 encodes two forms of a transcriptional repressor: structural and functional analysis of new viral cDNAs. *J. Virol.*, 63, 1743-1755 1989)

12. Smotkin,D., Wettstein,F.O. Transcription of human papillomavirus type 16 early genes in a cervical cancer and a cancer-derived cell line and identification of the E7 protein. *Proc. Natl. Acad. Sci. USA*, 83, 4680-4684 (1986)

13. Smotkin,D., Prokoph,H., Wettstein,F.O. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. *J. Virol.*, 63, 1441-1447 (1989)

14. Doorbar,J., Parton,A., Hartley,K., Banks,L., Crook,T., Stanley,M.A., and Crawford,L. Detection of novel splicing patterns in a HPV-16-containing keratinocyte cell line. *Virology*, 178, 254-262 (1990)

15. Grassman,K., Rapp,B., Maschek,H., Petry,K.U., Iftner,T. Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. *J. Virol.*, 70, 2339-2349 (1996)

16. Ozbun,M.A., Meyers,C. Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b. *J. Virol.*, 72, 2715-2722 (1998)

17. Steger,G., Corbach,S. Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *J. Virol.*, 71, 50-58 (1997)

18. Hummel,M., Hudson,J.B., Laimins,L.A. Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes. *J. Virol.*, 66, 6070-6080 (1992)
19. Baker,C.C., Noe,J.S. Transcription termination between bovine papillomavirus type 1 (BPV-1) early and late polyadenylation sites blocks late transcription in BPV-1-transformed cells. *J. Virol.*, 63, 3529-3534 (1989)
20. Öberg,D., Collier,B., Zhao,X., and Schwartz,S. Mutational inactivation of two distinct negative RNA elements in the human papillomavirus type 16 L2 coding region induces production of high levels of L2 in human cells. *J. Virol.*, 77, 11674-11684 (2003)
21. Terhune,S.S., Milcarek,C., Laimins,L.A. Regulation of human papillomavirus type 31 polyadenylation during the differentiation-dependent life cycle. *J. Virol.*, 73, 7185-7192 (1999)
22. Terhune,S.S., Hubert,W.G., Thomas,J.T., Laimins,L.A. Early polyadenylation signals of human papillomavirus type 31 negatively regulates capsid gene expression. *J. Virol.*, 75, 8147-8157 (2001)
23. Stoler,M.H., Wolinsky,S.M., Whitbeck,A., Broker,T.R., Chow,L.T. Differentiation-linked human papillomavirus types 6 and 11 transcription in genital condylomata revealed by *in situ* hybridisation with message-specific RNA probes. *Virology*, 172, 331-340 (1989)
24. Beyer-Finkler, E., Stoler, M. H., Girardi, F., Pfister, H. J. Cell differentiation-related gene expression of human papillomavirus 33. *Med. Microbiol. Immunol*. 179, 185-192 (1990)
25. Koffa,M.D., Graham,S.V., Takagaki,Y., Manley,J.L., Clements,J.B. The human papillomavirus type 16 negative regulatory element interacts with three proteins that act at different posttranscriptional levels. *Proc. Natl. Acad. Sci. USA*, 97, 4677-4682 (2000)
26. Wang,J., Manley,J.L. Regulation of pre-mRNA splicing in metazoa. *Curr. Opin. Gen. Dev.*, 7, 205-211 (1997)
27. Shatkin,A.J., Manley,J.L. The ends of the affair: Capping and polyadenylation. *Nat. Struct. Biol.*, 7, 838-842 (2000)
28. Cullen B. Nuclear RNA export. *J. Cell .Sci.*, 116, 587-597 (2007)
29. Mitchell,P., Tollervey,D. mRNA stability in eukaryotes. *Curr. Opin. Gen. Dev.*, 10, 193-198 (2000)
30. Proud,C.G. In Chapman,K.E., Higgins,S.J. (eds), Regulation of Gene Expression. Portland Press, London, Vol. 37, pp. 97-108 (2001)
31. Bentley,D. Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell Biol.*, 11, 347-351 (1999)
32. Hirose, Y., Manley, J. L. RNA polymerase II and the integration of nuclear events. *Genes Dev*. 14, 1415-1429. (2000)
33. Black,D.L. Mechanisms of alternative pre-messenger RNA splicing. *Annu .Rev. Biochem.*, 72, 291-336 (2003)
34. Smith,C.W.J., Valcárcel,J. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.*, **25,** 381-388 (2000)
35. Colgan D.F., Manley, J. L. Mechanisms and regulation of mRNA polyadenylation. *Genes Dev*. 11, 2755-2766 (1997)
36. Hieronymus, H., Silver, P. A. A systems view of mRNP biology. *Genes Dev.* 18, 2845-2860 (2004)
37. Singh, G., Lykke-Andersen. New insights into the formation of active nonsense-mediated decay complexes. *Trends Biochem. Sci.* 28, 464-466 (2003)
38. Wilusz, C., Wang, W., Peltz, S. W. Curbing the nonsense: the activation and regulation of mRNA surveillance. *Genes Dev.* 15, 2781-2785 (2001)
39. Sandri-Goldin, R. M. Interaction between a herpes simplex virus regulatory protein and cellular mRNA processing pathways. *Methods* 16, 95-104 (1998)
40. Malik,P., Schirmer,E.C. The Kaposi's sarcoma-associated herpesvirus ORF57 protein: a pleurotropic regulator of gene expression. *Biochem. Soc Trans.*, **34,** 705-710 (2006)
41. Belsham,G.J., Jackson,R.J. In Sonenberg,N., Hershey,J.W.B., and Mathews,M.B. (eds), Translational control of gene expression. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, pp. 869-900 (2000)
42. Ropers, D., Ayadi, L., Gattoni, R., Jacquenet, S., Damier, L., Branlant, C., Stevenin, J. Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sistes of HIV-1 RNA. *J. Biol. Chem.* 279(29), 29963-29973 (2004)
43. Ashe, M. P., Pearson, L. H., Proudfoot, NJ. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J.* 16, 5752-5763 (1997)
44. Klasens, B. I. F., Thiesen, M., Virtanen, A., and Berhout, B. The ability of the HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure. *Nucleic Acids Res.* 27, 446-454 (1999)
45. Gilmartin,G.M., Fleming,E.S., Oetjen,J., Graveley,B.R. CPSF recognition of an HIV-1 mRNA 3' processing enhancer: multiple sequence contacts involved in poly(A) site definition. *Genes Dev.*, 9, 72-83 (1995)
46. Moriera,A., Wollerton,M., Monks,J., Proudfoot,N. Upstream sequence elements enhance poly(A) site efficiency of the C2 complement gene and are phylogenetically conserved. *EMBO J.*, 14, 3809-3819 (1995)
47. Brackenridge,S., Proudfoot,N. Recruitment of a basal polyadenylation factor by the upstream sequence element of the human lamin B2 polyadenylation signal. *Mol. Cell. Biol.*, 20, 2660-2669 (2000)
48. DeZazzo,J.D., Imperiale,M.J. Sequences upstream of AAUAAA influence poly(A) site selection in a complex transcription unit. *Mol. Cell .Biol.*, 9, 4951-4961 (1989)
49. Prescott,J., Falck-Pedersen,E. Sequence elements upstream of the 3' cleavage site confer substrate strength to the adenovirus L1 and L3 polyadenylation sites. *Mol .Cell. Biol.*, 14, 4692-4693 (1994)
50. Russnak,R., Ganem,D. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes Dev.*, 4, 764-776 (1990)
51. Sanfaçon,H., Brodmann,P., Hohn,T. A dissection of the cauliflower mosaic virus polyadenylation signal. *Genes Dev.*, 5, 141-149 (1991)
52. Lutz,C.S., Alwine,J.C. Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of

the SV40 late polyadenylation signal. *Genes Dev.*, 8, 576-586 (1994)
53. Lutz,C.S., Murthy,K.G.K., Schek,N., O'Connor,J.P., Manley,J.L., Alwine,J.C. Interaction between the U1 snRNP-A protein and the 160 kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency *in vitro*. *Genes Dev.*, 10, 325-337 (1996)
54. Moriera,A., Takagaki,Y., Brackenridge,S., Wollerton,M., Manley,J.L., and Proudfoot,N.J. The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms. *Genes Dev.*, 12, 2522-2534 (1998)
55. Boelens, W. C., Jansen, E. J. R., van Venrooij, W. J., Stripecke, R., Mattaj, I. M., Gunderson, S. I. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* 72, 881-892 (1993)
56. Ko, B. and Gunderson, S. I. Identification of new poly(A) polymerase-inhibitory proteins capable of regulating pre-mRNA polyadenylation. *J. Mol. Biol.* 318, 1189-1206 (2002)
57. Millevoi, S., Geraghty, F., Idowu, B., Tam, J. L. Y., Antoniou, M., Vagner, S. A novel function for the U2AF 65 splicing factor in promoting pre-mRNA 3'-end processing. *EMBO Reports* 3, 869-874 (2002)
58. Phillips, C., Jung, S., and Gunderson, S. I. Regulation of nuclear poly(A) addition controls the expression of immunoglobulin M secretory mRNA. *EMBO J.* 20, 6443-6452 (2001)
59. Phillips, C., Pachikara, N., Gunderson, S. I. U1A inhibits cleavage at the immunoglobulin M heavy-chain secretory poly(A) site by binding between the two downstream GU-rich regions. *Mol. Cell. Biol*. 24, 6162-6171 (2004)
60. Beckley, S. A., Liu, P., Stover, M. L., Gunderson, S. I., Lichtler, A. C., Rowe, D. W. Reduction of target gene expression by a modified U1 snRNA. *Mol. Cell. Biol*. 21, 2815-2825 (2001)
61. Gunderson,S.I., Polycarpou-Schwarz,M., Mattaj,I.W. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell*, 1, 255-264 (1998)
62. Ashe, M. P., Furger, A., Proudfoot, N. J. Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. *RNA* 6, 170-177 (2000)
63. Vagner, S., Rûegsegger, U., Gunderson, S. I., Keller, W., Mattaj, I. W. Position-dependent inhibition of the cleavage step of pre-mRNA 3' end processing by U1 snRNP. *RNA* 6, 178-188 (2000)
64. Brennan,C.M., Steitz,J.A. HuR and mRNA stability. *Cell. Mol. Life Sci.*, 58, 266-277 (2001)
65. Peng,S.S.-Y., Cehn,C.-Y.A., Xu,N., Shyu,A.-B. RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. *EMBO J.*, 17, 3461-3470 (1998)
66. Fan,X.C., Steitz,J.A. Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the *in vivo* stability of ARE-containing mRNAs. *EMBO J.*, 17, 3448-3460 (1998)
67. Kiledjian,M., DeMaria,C.T., Brewer,G., Novick,K. Identification of AUF1 (heterogeneous nuclear ribonucleoprotein D) as a component of the a-globin mRNA stability complex. *Mol. Cell. Biol.*, 17, 4870-4876 (1997)
68. Sela-Brown,A., Silver,J., Brewer,G., Naveh-Many,T. Identification of AUF1 as a parathyroid hormone mRNA 3'-untranslated region-binding protein that determines parathyroid hormone mRNA stability. *J. Biol .Chem.*, 275, 7424-7429 (2000)
69. Carballo,E., Lai,W.S., Blackshear,P.J. Feedback inhibition of macrophage tumour necrosis factor-alpha production by tristetraprolin. *Mol. Cell. Biol.*, 281, 1001-1005 (1998)
70. Levy NS., Chung,S., Furneaux,H., Levy,AP. Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR. *J. Biol. Chem.*, 273, 6417-6423 (1998)
71. Hitti,E., Iakovleva,T., Brook,M., Deppenmeier,S., Gruber,A.D., Radzioch,D., Clark,A.R., Blackshear,P.J., Kotyarov,A., Gaestel,M. Mitogen-activated protein kinase-activated protein kinase 2 regulates tumor necrosis factor mRNA stability and translation mainly by altering tristetraprolin expression, stability, and binding to adenine/uridine-rich element. *Mol. Cell. Biol.*, 26, 2399-2407 (2006)
72. Kawai.T., Lal, A, Yang, X., Galban, S., Mazan-Mamczarz, K., Gorospe, M. Translational control of cytochrome *c* by RNA-binding proteins TIA-1 and HuR. *Mol. Cell. Biol.* 26, 3295-3307 (2006)
73. Rouault,TA. The role of iron regulatory proteins in mammalian iron homeostasis and disease. *Nat. Chem. Biol.*, 2, 406-414 (2006)
74. Furth,P.A., Baker,C.C. An element in the bovine papillomavirus late 3' untranslated region reduces polyadenylation cytoplasmic RNA levels. *J. Virol.*, 65, 5806-5812 (1991)
75. Furth,P.A., Choe,W.-T., Rex,J.H., Byrne,J.C., Baker,C.C. Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol. Cell. Biol.*, 14, 5278-5289 (1994)
76. Gunderson,S.I., Beyer,K., Martin,G., Keller,W., Boelens,W.C., Mattaj,I.M. the human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell*, 76, 531-541 (1994)
77. Tan,W., Schwartz,S. 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 (1995)
78. Sokolowski,M., Zhao,C., Tan,W., Schwartz,S. AU-rich mRNA instability elements on human papillomavirus type 1 late mRNAs and c-*fos* mRNAs interact with the same cellular factors. *Oncogene*, 15, 2303-2319 (1997)
79. Tan,W., Felber,B.K., Zolotukhin,A.S., Pavlakis,G.N., Schwartz,S. Efficient expression of the human papillomavirus type 16 L1 protein in epithelial cells by using rev and the rev-response element of human immunodeficiency virus or the *cis*-acting transactivation element of simian retrovirus type 1. *J. Virol.*, 69, 5607-5620 (1995)
80. Krecic, AM. Swanson, MS. hnRNP A1 complexes: composition, structure, and function. *Curr. Opin. Cell Biol*. 11, 363-371 (1999)

81. Sokolowski,M., Furneaux,H., Schwartz,S. The inhibitory activity of the AU-rich RNA element in the human papillomavirus type 1 late 3' untranslated region correlates with its affinity for the elav-like HuR protein. *J .Virol.*, 73, 1080-1091 (1999)
82. Mazan-Mamczarz, K., Galban, S., López de Silanes, I., Martindale, J. L., Atasoy, U., Keene, J. D., Gorospe, M. RNA-binding HuR enhances p53 translation in response to ultraviolet light irradiation. *Proc. Natl. Acad. Sci. USA* 100, 8354-8359 (2003)
83. Lal, A, Kawai, T., Yang, X., Mazan-Mamczarz, K., Gorospe, M. Anti-apoptotic function of RNA-binding HuR effected through prothymosin a. *EMBO J*. 24, 1852-1862 (2005)
84. Kullman, M., Gopfert, U., Siewe, B., Hengst, L. ELAV/Hu proteins inhibit p27 translation via an IRES element in the p27 5'UTR. *Genes Dev*. 16, 3087-3099 (2002)
85. Meng, Z., King, P. H., Nabors, L. B., Jackson, N. L., Chen, C. Y., Emanuel, P. D., Blume, S. W. The ELAV RNA-stability factor HuR binds the 5'-untranslated region of the human IGF-IR transcript and differentially represses cap-dependent and IRES-mediated translation. *Nucleic Acids Res*. 33, 2962-2979 (2005)
86. Carlsson, A., Schwartz, S. Inhibitory activity of the human papillomavirus type 1 AU-rich element correlates inversely with the levels of the elav-like HuR protein in the cell cytoplasm. *Archives Virol.* 145, 491-503 (2000)
87. Wiklund,L., Sokolowski,M., Carlsson,A., Rush,M., Schwartz,S. Inhibition of translation by UAUUUAU and UAUUUUUAU motifs of the AU-rich RNA instability element in the HPV-1 late 3' untranslation region. *J. Biol. Chem.*, 277, 40462-40471 (2002)
88. Zhao, X., Rush, M., Carlsson, A., Schwartz, S. The presence of inhibitory RNA elements in the late 3'-untranslated region is a conserved property of human papillomaviruses. *Virus Res.* 125, 135-144 (2007)
89. Kennedy,I.M., Haddow,J.K., Clements,J.B. Analysis of human papillomavirus type 16 late mRNA 3' processing signals *in vitro* and *in vivo*. *J. Virol.*, 64, 1825-1829 (1990)
90. Dietrich-Goetz,W., Kennedy,I.M., Levins,B., Stanley,M.A., Clements,J.B. A cellular 65kDa protein recognizes the negative regulatory element of human papillomavirus late mRNA. *Proc. Natl. Acad. Sci. USA*, 94, 163-168 (1997)
91. Cumming,S.A., McPhillips,M.G., Veerapraditsin,T., Milligan,S.G., Graham,S.V. Activity of the human papillomavirus type 16 late negative regulatory element is partly due to four weak consensus 5' splice sites that bind a U1 snRNP-like complex. *J. Virol.*, 77, 5167-5177 (2003)
92. Cumming,S.A., Repellin,C.E., McPhillips,M., Radford,J.C., Clements,J.B., Graham,S.V. The human papillomavirus type 31 late 3' untranslated region contains a complex bipartite negative regulatory element. *J. Virol.*, 76, 5993-6003 (2002)
93. Stanley,M.A., Browne,H.M., Appelby,M., Minson,A.C. Properties of a non-tumourigenic human cervical keratinocyte cell line. *Int. J. Cancer*, 43, 672-676 (1989)
94. Jeon,S., Allen-Hoffman,B.L., Lambert,P.F. Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. *J. Virol.*, 69, 2989-2997 (1995)
95. McPhillips,M.G., Veerapraditsin,T., Cumming,S.A., Karali,D., Milligan,S.G., Boner,W., Morgan,I.M., Graham,S.V. SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during epithelial differentiation. *J. Virol.*, 78, 10598-10605 (2004)
96. Berget,S. Exon recognition in vertebrate splicing. *J. Biol. Chem.*, 270, 2411-2414 (1995)
97. Wassarman,K.M., Steitz,J.A. Association with terminal exons in pre-mRNAs: a new role for the U1 snRNP? *Genes Dev.*, 7, 647-659 (1993)
98. Zhang,M.Q. Statistical features of human exons and their flanking regions. *Human Mol. Gen.*, 7, 919-932 (1998)
99. Izaurralde,E., Jarnolowski,A., Beisel,C., Mattaj,I.W., Dreyfuss,G., Fischer,V. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J. Cell Biol.*, 137, 27-35 (1997)
100. Zhao,X., Rush,M., Schwartz,S. Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene. *J. Virol.*, 78, 10888-10905 (2004)
101. Zhu, J., Mayeda, A., Krainer, A. R. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell* 8, 1351-1361 (2001)
102. Damgaard, CR. and Tange, TØ. Kjems J. hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure. *RNA* 8, 1401-1415 (2002.)
103. Marchand, V., Méreau, A., Jacquenet, S., Thomas, D., Mougin, A., Gattoni, R., Stévenin, J., Branlant, C. A Janus splicing regulatory element modulates HIV-1 *tat* and *rev* mRNA production by coordination of hnRNP A1 cooperative binding. *J. Mol. Biol.* 323, 629-652 (2002)
104. Blanchette, M., Chabot, B. Modulation of exon skipping by high affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J.* 18, 1939-1952 (1999)
105. Misteli,T. RNA splicing: what has phosphorylation got to do with it? *Curr. Biol.*, 9, R198-R200 (1999)
106. Fortes, P., Cuevas, Y., Guan, F., Liu, P., Pentlicky, S., Jung, S. P., Martinez-Chantar, M. L., Prieto, J., Rowe, D., Gunderson, S. I. Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small nuclear RNAs targeted to terminal exons of pre-mRNA. *Proc. Natl. Acad. Sci. USA* 100, 8264-8269 (2003)
107. Afshar,M., Prescott,C.D., Varani,G. Structure-based and combinatorial search for new RNA-binding drugs. *Curr. Opin. Biotechnol.*, 10, 59-63 (1999)
108. Fuchs,E., Byrne,C. The epidermis: rising to the surface. *Curr. Opin. Gen. Dev.*, 4, 725-736 (1994)
109. Hamy, F., Felder, ER., Heizmann, G, Lazdins, J., Aboul-ela, F., Varani, G., Karn, J., Klimkait, T. An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc. Natl. Acad. Sci. USA* 94, 3548-3553 (1997)
110. Hwang, S, Tamilarasu, N., Ryan, K., Huq, I., Richter, S., Still, WC., Rana, T. M. Inhibition of gene expression in human cells through small molecule-RNA interactions. *Proc. Natl. Acad. Sci. USA* 96, 12997-13002 (1999)

111. De Paula, D., Bentley, M. V. L. B., Mahato, R. I. Hydrophobization and bioconjugation for enhanced siRNA delivery and targeting. *RNA* 13, 431-456 (2007)
112. Caceres,J.F., Kornblihtt,A.R. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.*, 18, 186-193 (2002)

Abbreviations UTR: untranslated region, ARE: AU-rich element, LRE: late regulatory element, PV: papillomavirus, HPV: human papillomavirus, BPV: bovine papillomavirus, HIV: human immunodeficiency virus, HSV: herpes simplex virus, RNP: ribonucleoprotein particle , snRNA: small nuclear RNA, snRNP: small nuclear ribonucleoprotein, PABP: poly(A) binding protein, PAP: poly(A) polymerase, CPSF: cleavage polyadenylation specificity factor, CstF: cleavage stimulatory factor, hnRNP: heterogeneous RNP, CTD: carboxyl terminal domain of RNA polymerase II, EJC: exon junction complex, USE: upstream sequence element, PTB: polypyrimidine tract binding protein, TTP: tristetrapolin, TNF: tumor necrosis factor, VEGF: vascular endothelial growth factor, MAPK: mitogen-activated protein kinase, IRE: iron response element, IRP: iron response protein, LTR: long terminal repeats, CAT: chloramphenicol acetyltransferase, RRE: HIV1 Rev responsive element, CTE: simian virus constitutive transport element, NLE: negative regulatory element-like element, MIE: major inhibitory element, SIE: subsidiary inhibitory element, UV: ultra violet irradiation, EMSA: electrophoretic mobility shift assay, CIN: cervical intraepithelial neoplasia, SR protein: serine/arginine-rich protein, siRNA: small interfering RNA, miRNA: microRNA, shRNA: small hairpin RNA, 5'ss: 5'splice site, 3'ss: 3' splice site,

Key Words Papillomavirus, late gene expression, RNA processing, mRNA stability 3' Untranslated Region, Post-Transcriptional control, HPV16, BPV1, HPV1, HPV31, Terminal Exon Definition, Polyadenylation, RNA-protein interactions, Review

Send correspondence to: Dr Sheila V.Graham, Institute of Comparative Medicine, Jarrett Building, Garscube Estate, University of Glasgow, Glasgow G61 1QH, Scotland, UK, Tel: 441413306256, Fax: 441413305602, E-mail: s.v.graham@bio.gla.ac.uk

http://www.bioscience.org/current/vol13.htm