A 57-Nucleotide Upstream Early Polyadenylation Element in Human Papillomavirus Type 16 Interacts with hFip1, CstF-64, hnRNP C1/C2, and Polypyrimidine Tract Binding Protein

Xiaomin Zhao,1 Daniel Öberg,1 Margaret Rush,1 Joanna Fay,2 Helen Lambkin,2 and Stefan Schwartz1*

Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden,1 and Dublin Institute of Technology, Dublin, Ireland2

Received 18 August 2004/Accepted 15 November 2004

We have investigated the role of the human papillomavirus type 16 (HPV-16) early untranslated region (3′ UTR) in HPV-16 gene expression. We found that deletion of the early 3′ UTR reduced the utilization of the early polyadenylation signal and, as a consequence, resulted in read-through into the late region and production of late L1 and L2 mRNAs. Deletion of the U-rich 3′ half of the early 3′ UTR had a similar effect, demonstrating that the 57-nucleotide U-rich region acted as an enhancing upstream element on the early polyadenylation signal. In accordance with this, the newly identified hFip1 protein, which has been shown to enhance polyadenylation through U-rich upstream elements, interacted specifically with the HPV-16 upstream element. This upstream element also interacted specifically with Cstf-64, hnRNP C1/C2, and polypyrimidine tract binding protein, suggesting that these factors were either enhancing or regulating polyadenylation at the HPV-16 early polyadenylation signal. Mutational inactivation of the early polyadenylation signal also resulted in increased late mRNA production. However, the effect was reduced by the activation of upstream cryptic polyadenylation signals, demonstrating the presence of additional strong RNA elements downstream of the early polyadenylation signal that direct cleavage and polyadenylation to this region of the HPV-16 genome. In addition, we identified a 3′ splice site at genomic position 742 in the early region with the potential to produce E1 and E4 mRNAs on which the E1 and E4 open reading frames are preceded by both E6 and E7 AUGs. These mRNAs would therefore be more efficiently translated into E1 and E4 than previously described HPV-16 E1 and E4 mRNAs on which E1 and E4 are preceded by both E6 and E7 AUGs.

Human papillomaviruses (HPVs) are a group of nonenveloped, double-stranded DNA tumor viruses with tropism for epithelial cells (21). HPV-16 is one of the most common sexually transmitted HPV types and is also the HPV type most frequently detected in cervical cancer (42, 63). The life cycle of HPV-16 is strictly linked to the differentiation program of the infected cell (17, 21). While many of the early HPV gene products are present in all layers of the squamous epithelium, expression of the late mRNAs encoding the L1 and L2 capsid proteins is restricted to the terminally differentiated cells in the upper layers of the epithelium (29). The early and late regions are separated by the early polyadenylation signal named pAE (Fig. 1), and RNA elements that regulate the use of the pAE signal are likely to affect late gene expression. Here we have studied the role of the early untranslated region (3′ UTR) in late gene expression by investigating how deletions in the early 3′ UTR affect late mRNA levels.

The 3′ UTR is often the site for RNA elements that regulate various steps in the mRNA processing pathway, for example, mRNA transport, half-life, and translation (13). The best-studied 3′-UTR element is the group of AU-rich RNA instability elements, often containing multiple copies of AUUUA motifs (9). We and others have previously shown that late papillomavirus mRNAs contain cis-acting regulatory RNA elements in the 3′ UTR (3, 38-41). These RNA elements all act by reducing the levels of the late mRNAs. Inhibitory RNA elements in the late 3′ UTR have been found in bovine papillomavirus type 1 (BPV-1) (18), HPV-1 (51), HPV-16 (25, 26), and HPV-31 (15). The BPV-1 and HPV-16 late 3′-UTR elements both interact with U1 snRNP (14, 19), which inhibits polyadenylation (20). The HPV-16 late 3′ UTR possibly interacts with other factors (27). The HPV-1 late 3′-UTR element is an AU-rich RNA instability element containing multiple UAUUUAU motifs, which reduces mRNA half-life (45) and inhibits mRNA translation (55). This element interacts with HuR (43), hnRNP C1/C2 (44, 45), and poly(A)-binding protein (55). Interestingly, the human immunodeficiency virus type 1 (HIV-1) Rev protein, which interacts with the Crm1 nuclear export factor (34), is able to overcome the effect of the inhibitory late 3′-UTR sequences in BPV-1 (5), HPV-1 (51), and HPV-16 (50) if the Rev response element is present in cis. This suggests that the HIV-1 Rev protein competes with cellular factors that interact with the papillomavirus RNA elements and prevents expression of the late mRNAs, possibly retaining the mRNAs in the nucleus. It has been speculated that the role of these elements in the viral life cycle is to repress late gene expression at an early stage in the viral life cycle (18, 26, 41). However, the function of the elements in the context of the entire viral genome and in a differentiating environment that allows the virus to complete the life cycle remains to be experimentally determined.
The early 3′ UTR (eUTR) appears to contain sequences that affect the expression levels of the early mRNAs (22). It has been shown that the insertion of the HPV-16 early 3′ UTR in inducible transcripts rendered the mRNAs unstable in mouse fibroblasts (23). These results indicated that the eUTR contains mRNA instability elements that regulate the expression levels of the early mRNAs. Here we have studied the effect of deletions in the HPV-16 eUTR on the expression levels of early and late mRNAs. We found that the U-rich region of the early 3′ UTR had a stimulatory effect on polyadenylation at the pAE. This region interacted with hFip1, CstF-64, polypyrimidine tract binding protein (PTB), and hnRNP C1/C2, suggesting that they were involved in polyadenylation of HPV-16 early mRNAs. We also found that the HPV-16 eUTR had a weak inhibitory effect on the early mRNAs, in line with previously published data (23). The results presented here demonstrate that the early 3′ UTR encodes RNA elements that enhance polyadenylation of early mRNAs.

MATERIALS AND METHODS

Plasmid constructions. pBearly was generated by PCR amplification of nucleotides 3400 to 4325 of the HPV-16 genome with oligonucleotides E4S (61) and L2START-Anti and cloning into the pCRII-TOPO vector (Invitrogen). For oligonucleotide sequences, see Table 1. Numbers refer to nucleotide positions in the HPV-16R sequence (6). The major p97 promoter and the differentiation dependent promoter p670 are indicated. 5′ Splice sites (SD), 3′ splice sites (SA), and the early and late polyadenylation signals pAE and pAL, respectively, are shown. The structures of the pBearly97, pBearly, and pBEL (61) expression plasmids are shown. The pBEL-derived plasmids contain a deletion in the late UTR that removes previously identified negative elements (25, 26) in order to increase the chances of obtaining detectable levels of late mRNAs (61). CMV, human cytomegalovirus immediate-early promoter.

FIG. 1. The HPV-16 genome and the subgenomic expression plasmids pBearly97, pBearly, and pBEL (61). Schematic representation of the HPV-16 genome. Boxes indicate the protein-coding regions. Numbers refer to nucleotide positions in the HPV-16R sequence (6). The major p97 promoter and the differentiation dependent promoter p670 are indicated. 5′ Splice sites (SD), 3′ splice sites (SA), and the early and late polyadenylation signals pAE and pAL, respectively, are shown. The structures of the pBearly97, pBearly, and pBEL (61) expression plasmids are shown. The pBEL-derived plasmids contain a deletion in the late UTR that removes previously identified negative elements (25, 26) in order to increase the chances of obtaining detectable levels of late mRNAs (61). CMV, human cytomegalovirus immediate-early promoter.
into pBearly using SalI and KpnI, thereby creating pBearly97. pBELDPU and pBELMDPU were generated by PCR amplification of nucleotides 4300 to 5658 of the HPV-16 genome with oligonucleotides 4272 ZHAO ET AL. J. VIROL.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2START Anti</td>
<td>CTTCGACGAGTAGGTGCACTGATTCTCCTAGT</td>
</tr>
</tbody>
</table>
| poly(A)/MLUI(AS) | CTCCGATTTCAAGATTAACTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
[32P]UTP, as previously described (47). Briefly, recombinant protein or cell extract was incubated with radiolabeled RNA (10^5 cpm) in 1× binding buffer (60 mM KCl, 10 mM HEPES [pH 7.6], 3 mM MgCl₂, 1 mM dithiothreitol, 5% glycerol, and 5 μg of heparin/μl) in the absence or presence of unlabeled competitor RNA. The radiolabeled RNAs were purified by phenol-chloroform extraction and ethanol precipitation and were resuspended in water. UV cross-linking and synthesis of radiolabeled RNA was performed as previously described (47). Radiolabeled RNA (10^5 cpm) was used in each UV cross-linking reaction. HeLa cell nuclear extracts were prepared according to the method of Dignam et al. (16). His-tagged hFip1 and hnRNP C1 were purified by using a HiTrap chelating column according to the manufacturer’s instructions (Pharmacia Biotech). Purified glutathione S-transferase (GST)-PTB proteins were prepared by using glutathione-Sepharose beads according to the manufacturer’s instructions (Pharmacia Biotech), as previously described (47). Twenty micrograms of nuclear extract or 50 ng of each recombinant his-hFip1, his-hnRNP C1, and GST-PTB were used for UV cross-linking.

RESULTS

Analysis of subgenomic plasmids expressing the early region of the HPV-16 genome identified E1 and E4 mRNAs. To investigate the role of the HPV-16 early 3’ UTR in HPV-16 gene expression, we generated plasmids pBearly97 and pBearly expressing the early region of HPV-16 under control of the human immediate-early CMV promoter (Fig. 1). The CMV promoter was inserted either in the place of the p97 promoter or the p670 promoter, resulting in pBearly97 and pBearly, respectively (Fig. 1). To determine which mRNAs these plasmids produced after transfection, they were first transfected into HeLa cells and total RNA was extracted. RNA from pBearly97 was analyzed by Northern blotting using the multiple probes indicated in Fig. 2A. All plasmids were included in at least three independent transfections with similar results. The variation between duplicates or triplicates in one transcription series was less than 20%. The CMV probe detected all mRNAs produced from the plasmids, because it hybridized to the leader sequence derived from the CMV promoter (Fig. 2A). The other probes, named probes #1, #2, and #3 (Fig. 2A), hybridized to sequences in the E6 and E7 regions that were either included or excluded on the various alternatively spliced HPV-16 mRNAs (Fig. 2A). The CMV probe detected at least four bands, of which the lowest migrating band was the most abundant species (Fig. 2B). This mRNA did not hybridize to the leader sequence (Fig. 2B). This mRNA was also detected by RT-PCR, but at least one of these mRNAs was seen in the Northern blots (the upper band in Fig. 2B) and by the E1-specific probe (61 and data not shown). The 5’ splice site at position 1302 and the 3’ splice site at 2582 were not used (data not shown).

Analysis of the RT-PCR products obtained with primers A and C (Fig. 2A) revealed an additional band, marked with an asterisk, of a size that could not be accounted for by the use of the known splice sites (Fig. 2C, lane 2). Cloning and sequencing of the RT-PCR product identified a 3’ splice site in the early region at genomic position 742. The intron ends with an AG as expected and has a suboptimal poly(A) tract with interrupted runs of pyrimidines, of which the longest consecutive pyrimidine sequence is 6 nucleotides (Table 2). This splice site is used less frequently than the 3’ splice site at 409 (Fig. 2C, lane 2). An mRNA using the 3’ splice site at position 742 followed by splicing from the 5’ splice site at positions 880 to 3558 is the only mRNA on which the E4 open reading frame (ORF) is used (Fig. 2A). The other ORFs would be hindered by the presence of both E6 and E7 ORFs in the E1 and E4 mRNAs in which the E1 and E4 ORFs are preceded only by the E6 AUG (Fig. 2A, mRNA 12). The E6 AUG deviates from the Kozak sequence for translational start sites and is located very close to the 5’ end, two properties that render it suboptimal and allow leaky scanning and translation initiation at the second E4 ORF (48, 62). If the mRNA that is spliced to the 3’ splice site at position 742 remains unspliced until the pAE, an mRNA is created on which the E1 ORF is preceded only by the E6 AUG (Fig. 2A, mRNA 13). In conclusion, this 3’ splice site generates E1 and E4 mRNAs in which the E1 and E4 ORFs are preceded only by the suboptimal E6 AUG, whereas all previously identified E1 and E4 mRNAs contain both the E6 and the E7 AUGs upstream of their ORFs (2). Therefore, the E1 and E4 mRNAs using the 3’ splice site at 742 would be more efficient producers of E1 and E4 protein, respectively, compared to the previously identified mRNAs on which translation of the E1 and E4 ORFs would be hindered by the presence of both E6 and E7 AUGs.

Analysis of the expression plasmid named pBearly (Fig. 1 and 3A) by Northern blotting and multiple probes (Fig. 3A) revealed a simpler picture, with one dominating mRNA and additional but minor species barely detectable by Northern blotting and RT-PCR (Fig. 3B and C). Northern blotting with CMV, E1, and E4 probes (Fig. 3A) indicated that the major band was spliced from the 5’ splice site at position 880 to the 3’ splice site at position 3558 (Fig. 3B). The E2 mRNA in Fig.
3B can be seen in longer exposures of the filter (data not shown). Analysis of cytoplasmic RNA revealed that the mRNAs produced from pBearly were exported to the cytoplasm (Fig. 3B). The results were confirmed by cloning and sequencing of the major RT-PCR product (Fig. 3C). Based on the simpler expression pattern, we decided to use pBearly to study the effect of deletions in the early 3’/H11032 UTR on the early mRNA levels.

Deletion of the HPV-16 early 3’/H11032 UTR caused a modest increase in early mRNA levels. The early 3’/H11032 UTR can be divided into a 5’ and a 3’ half, the latter being U rich (Fig. 4A). Removal of either the entire early 3’ UTR or the U-rich 3’ half alone resulted in a modest increase in early mRNA levels (Fig. 4B). A quantitation of the 880/3358 major early mRNA seen by Northern blotting with the CMV probe revealed that removal of the entire early 3’/H11032 UTR resulted in a 1.6-fold increase in mRNA steady-state levels (Fig. 4B and C). Similar results were obtained with CMV and E4 probes (Fig. 4B). Although this effect was relatively small, the results were in agreement with previously published data which showed an mRNA-destabilizing effect of the early 3’/H11032 UTR (22, 23). Deletion of the U-rich region alone had a smaller effect on mRNA levels (Fig. 4B). The quantitations shown in Fig. 4B were made in a phosphorimager. We concluded that the early 3’/H11032 UTR had a modest but
reproducible 1.6-fold inhibitory effect on the steady-state levels of the early HPV-16 mRNAs in HeLa cells.

**Mutational inactivation of the HPV-16 early polyadenylation signal induces late gene expression.** The early 3' UTR sequence could possibly also act on the early polyadenylation signal. To test this, the various 3'-UTR deletions were inserted into the previously described pBEL plasmid (61) (Fig. 1 and SA). The deletions were also introduced into pBELM (61), a
derivative of pBEL which produces late spliced L1 mRNAs as a result of the mutational inactivation of splicing silencers in L1 (61) (Fig. 5A). In these plasmids, the effects on polyadenylation of the early 3'-UTR deletions could be measured by probing for late mRNAs using the L1 probe (Fig. 5A). First we investigated if inactivation of the pAE itself induced late gene expression. The polyadenylation signal in pBEL was destroyed by specific mutations in the canonical AAUAAA, resulting in pBELDP (Fig. 5A), or by a deletion that removed the

<table>
<thead>
<tr>
<th>Genomic position</th>
<th>Sequence of the 3' splice site#</th>
</tr>
</thead>
<tbody>
<tr>
<td>409---------------</td>
<td>AAACCGUUGUGUGAUUGUAUUAGUGG</td>
</tr>
<tr>
<td>526---------------</td>
<td>ACCCGUGAUAUGUGUGUGCAGUC</td>
</tr>
<tr>
<td>2709--------------</td>
<td>AAGAACUGAAUUCCUUUCUAAGGAC</td>
</tr>
<tr>
<td>3358--------------</td>
<td>AUAAUUGCUCAUCUGUGUUUAGCAG</td>
</tr>
<tr>
<td>742---------------</td>
<td>ACAAAUUAUGUAACCUUUGUCAGAGGU</td>
</tr>
</tbody>
</table>

# The longest uninterrupted pyrimidine stretch is underlined.

FIG. 3. (A) Structure of the pBearly expression plasmid. The locations of the CMV, E1, and E4 probes (61) are indicated. Locations of primers 757S (60) and E4A (61) used for RT-PCR are shown. Schematic representation of all predicted splice variants of the mRNAs produced by pBearly are shown. SD, 5' splice site; SA, 3' splice site. (B) Northern blots on total or cytoplasmic RNA extracted from HeLa cells transfected with pBearly. Filters were hybridized with the CMV, E1, and E4 probes (61) (Fig. 3A). The E1 mRNA and the mRNAs spliced from 880 to 2709 or 3358 are indicated on the right. (C) RT-PCR with primers 757S (60) and E4A (61) on total RNA extracted from HeLa cells transfected with pBearly or pBEL (61). The position of the mRNA spliced from 880 to 3358 is indicated. Lane 1, RT-PCR on sample lacking RNA. The right-hand gel shows the same RNA samples as in the left-hand gel analyzed by RT-PCR with the same primer pair in the absence of RT.
AAUAAA and 57 nucleotides of the early 3' UTR, resulting in pBELDPU (Fig. 5A). The plasmids were transfected into HeLa cells, and the late mRNA levels were monitored with the L1 probe (Fig. 5A). The results revealed that late mRNA levels produced by pBEL were undetectable, as previously described (Fig. 5B) (61), whereas plasmids pBELDP and pBELDPU, which contained the mutations in the polyadenylation signal, produced high levels of late mRNAs (Fig. 5B). Both the L2/L1 mRNAs and the spliced L1 mRNAs were induced by the removal or inactivation of the early polyadenylation signal (Fig. 5B). These results demonstrated that inactivation of the early polyadenylation signal induced late gene expression. Introducing the same two mutations in pBELM (61), generating plasmids pBELMDP and pBELMDPU, also resulted in an increase in late mRNA levels (Fig. 5B). The late mRNAs produced from pBELM were primarily spliced L1 mRNAs (Fig. 5B), because the splicing silencers in L1 had been inactivated by point mutations, as previously described (61). Because pBELM produced detectable levels of late mRNAs, the fold induction of late mRNAs caused by the pAE mu-
tations could be calculated. The results revealed that alteration of the AAUAAA sequence of the pAE to ACGCGU caused a 7.3-fold increase in late mRNA levels (Fig. 5B). The early and late mRNAs produced from the various plasmids with and without early UTRs were also found in the cytoplasmic fraction when cytoplasmic RNA was analyzed (Fig. 5B and data not shown). The results demonstrated that inactivation of the pAE resulted in induction of late gene expression and suggested that the early 3′ UTR may affect expression levels of the late mRNAs.

The U-rich region of the HPV-16 early 3′ UTR enhances recognition of the early polyadenylation signal. Next, the entire early 3′ UTR or the U-rich 3′ half was deleted in the presence of a functional polyadenylation signal, resulting in

---

FIG. 5. (A) Schematic drawing of the deletions in the early 3′ UTR introduced in pBEL (61) to create the indicated plasmids. The pBEL-derived plasmids contain a deletion in the late UTR that removes previously identified negative elements (25, 26) in order to increase the chances of obtaining detectable levels of late mRNAs (61). The same deletions were introduced into pBELM, in which the splicing silencer in the L1 coding region had been inactivated (61). The AAUAAA-to-ACGCGU mutation of the pAE in pBELDP is indicated. Nucleotide numbers refer to the HPV-16R sequence (6). The location of the L1 probe in pBEL is indicated. The locations of the RT-PCR primers 757S (61) and E5stops (Table 1) are shown. The two late mRNAs encoding L2 and L1 are displayed, and the location of the previously identified splicing silencer in L1 is indicated. (B and C) Northern blots on total or cytoplasmic RNA extracted from HeLa cells transfected with the indicated plasmids. Filters were hybridized to the L1 probe (61) (Fig. 5A). The L2/L1 and the L1 mRNAs are indicated. The fold induction of late mRNAs were calculated on the data obtained with pBELM-derived plasmids. The levels of late mRNAs were quantified in a phosphorimager and were divided with the late mRNA levels detected in the pBELM lane to yield fold induction. (D) Northern blots on total RNA extracted from HeLa cells transfected with the indicated plasmids. Filters were hybridized to the CMV probe or the E4 probe (61) (for location of probes, see Fig. 3A). The early mRNAs, the L2/L1 and the L1 mRNAs, are indicated.
pBELDUTR and pBELDU, respectively (Fig. 5A). Interestingly, these deletions also induced late gene expression, but to a lower extent than the mutational inactivation of the pAE (Fig. 5B). These results demonstrated that the early 3' UTR contains sequences that enhance polyadenylation at the early poly(A) signal. Analysis of the same mutations in the background of pBELM (Fig. 5A) yielded similar results (Fig. 5B); however, as the starting plasmid pBELM produced measurable levels of late mRNAs, the fold induction of late gene expression could be determined (Fig. 5B). Deletion of the entire 3' UTR resulted in a 3-fold induction of late gene expression, whereas deletion of the U-rich 3' half caused a 2.4-fold induction (Fig. 5B), indicating that the U-rich half of the early 3' UTR contained the major enhancing element. Deletion of the 5' half of the early 3' UTR affected late mRNA levels to a lesser extent (Fig. 5C). We concluded that the U-rich region in the early 3' UTR acts as an upstream sequence element (USE) that enhances polyadenylation at the pAE.

**Mutational inactivation of the HPV-16 early polyadenylation signal induces cryptic upstream polyadenylation signals.**

The RNA samples extracted from HeLa cells transfected with plasmids pBEL, pBELDU, pBELDUTR, and pBELDPU were also hybridized to the CMV probe (Fig. 2A), which detects both early and late mRNAs. This allowed us to determine the extent of the effect of the mutations in the early poly(A) signal and early 3' UTR by measuring the ratios between early and late mRNAs. As can be seen, all plasmids produced primarily early mRNAs, including pBELDPU, in which the polyadenylation signal had been inactivated (Fig. 5D). Hybridization with the E4 probe confirmed that the dominant lower bands on the Northern blots represented early mRNAs (Fig. 5D). The late mRNAs were less abundant in all cases (Fig. 5D). The early mRNA produced from pBELDPU migrated faster than the early mRNAs produced from pBEL, which had a functional pAE (Fig. 5D). These results suggested that an upstream cryptic polyadenylation signal could have been activated.

To determine how early mRNAs could be produced when the early poly(A) signal had been inactivated, 3' RACE was performed on the mRNA samples from pBEL and pBELDPU RNA extractions. Two different sense primers (757S and
E5stops) were used (Fig. 5A, Table 1). A major band was seen when 3' RACE was performed with primer 757S (60) and oligo(dTGC) (31) on total RNA from pBELDPU-transfected cells (Fig. 6A). Cloning and sequencing of the 3' RACE product revealed that this band represented an mRNA that was cleaved and polyadenylated upstream of the mutant pAE at genomic position 3820/3821 (Fig. 6B). In addition, a minor band was seen upstream of this cryptic early mRNA in the Northern gel (Fig. 5D). This second band was amplified with primers E5stops and oligo(dTGC) (Fig. 5A, Table 1) and represented a collection of minor mRNA species cleaved and polyadenylated at positions 4110/4111 (1 of 9), 4114/4115/4116 (2 of 9), 4146/4147 (1 of 9), 4148/4149/4150 (5 of 9) (Fig. 6C).

In conclusion, all alternatively polyadenylated early HPV-16 mRNAs produced in cells transfected with the pAE mutant pBELDPU were cleaved and polyadenylated upstream of the mutated pAE. The 3' RACE reactions on total RNA from pBELDPU-transfected HeLa cells. (D) The cleavage sites used by the wild-type pAE. Cleavage sites were identified by 3' RACE on RNA from pBEL-transfected HeLa cells. Numbers refer to nucleotide positions in the HPV-16R sequence (6). Two or three equally plausible cleavage sites are given when the poly(A) tail is preceded by one or two As.

FIG. 6. (A) 3' RACE on total RNA from HeLa cells transfected with pBELDPU and pBEL (Fig. 5A). The primer used for pBELDPU in 3' RACE was 757S (61) (for the location of the primer, see Fig. 5A) and oligo(dTGC) (31). The primers used for pBEL in 3' RACE were E5stops (Table 1) and oligo(dTGC) (31). The arrow indicates the major amplification product which was cloned and sequenced. The right-hand gel shows the same RNA samples as those in the left-hand gel analyzed by RT-PCR with the same primer pair in the absence of RT. (B) The major cleavage site identified by 3' RACE on RNA from pBELDPU-transfected cells is indicated. (C) Positions of multiple minor cleavage sites identified by 3' RACE on RNA from pBELDPU-transfected HeLa cells. (D) The cleavage sites used by the wild-type pAE. Cleavage sites were identified by 3' RACE on RNA from pBEL-transfected HeLa cells. Numbers refer to nucleotide positions in the HPV-16R sequence (6). Two or three equally plausible cleavage sites are given when the poly(A) tail is preceded by one or two As.
position 4232/4233/4234, 12 nucleotides downstream of the 
pAE, indicating that this was the major cleavage site of the 
wild-type pAE. In conclusion, mutational inactivation of the 
pAE activated cryptic polyadenylation signals. All of these 
were located upstream of the pAE, indicating that additional 
strong regulatory elements were present downstream of the 
pAE and that these elements directed polyadenylation to this 
region in the genome.

The HPV-16 early 3' UTR interacts with CstF-64, hnRNP 
C1/C2, and PTB. Having established that the U-rich region in 
the HPV-16 early 3' UTR enhances early polyadenylation, we 
wished to investigate if the early 3' UTR interacts with cellular 
actors. Three radiolabeled RNAs, named TUP, TDU, and 
TDP, were synthesized in vitro (Fig. 7A). One RNA contained 
the entire 3' UTR and the pAE (TUP), one RNA contained a 
deletion that removed the U-rich region in the early 3' UTR 
(TDU), and one RNA contained the entire 3' UTR and a 
mutationally inactivated AAUAAA (TDP) (Fig. 7A). UV 
cross-linking of the radiolabeled RNAs to nuclear extracts 
identified four proteins, named A, B, C, and D, that 
were specifically competed for by the HPV-16 early 3' UTR 
(Fig. 7E). These factors did not interact with an unrelated 
RNA derived from the L1 coding region (Fig. 7B). All factors 
that interacted with the U-rich region in the TUP probe also 
bound to the TDP probe, indicating that they interacted with 
the U-rich region independently of the AAUAAA hexanucle-
otide (Fig. 7B). To confirm that proteins cross-linked specifically 
to the TUP RNA, competitions with the TUP or TDU 
RNA showed that the TUP RNA competed for the factors that 
bound to the TUP RNA, whereas the TDU deletion mutant 
failed to compete for the proteins binding to the U-rich region 
(Fig. 7B). The majority of the proteins that cross-linked to the 
early 3' UTR also did so in the presence of the TDP compet-
tor RNA (Fig. 7B), confirming that most proteins 
cross-linked to the U-rich 3' half of the early 3' UTR.

Next we performed UV cross-linking to binding sites of 
hnRNP C1/C2 (probe T2C1), HuR (probe T2B2), and CstF-64 
(probe TcStF64) in parallel, because these factors bind to 
U-rich sequences and have been proposed to play various roles 
in polyadenylation (60). The sequences of the transcribed reg-
ions in pT2C1, pT2B2, and pTCstF64 RNAs are shown in Fig. 
7A. Interestingly, protein B, detected by the HPV-16 early 3' 
UTR probe TUP, migrates at the same position as the CstF-64 
protein (Fig. 7B), and protein C migrates at the same positions 
as hnRNP C (Fig. 7C). We tested if the TUP and TDU RNAs 
competed with the CstF-64 binding site for CstF-64. As can be 
seen, the HPV-16 early 3' UTR RNA TUP competed for 
CstF-64, whereas the TDU RNA did not (Fig. 7D). The 
CstF-64 binding site competed with itself (Fig. 7D), as ex-
pected. Therefore, CstF-64 binds specifically to the U-rich 
region in the HPV-16 early 3' UTR. To investigate if the 
smaller protein C represented hnRNCP, we tested if the T2C1 
RNA, which contains six hnRNCP C1/C2 binding sites, com-
peted with the HPV-16 3' UTR probe TUP. The results re-
vealed that the T2C1 RNA competed specifically for the lower-
molecular-sized two proteins, named C and D (Fig. 7E), but 
not with the high-molecular-weight bands, demonstrating that 
they were hnRNCP C1/C2, while the T2B2 RNA which contains 
four HuR binding sites did not compete with the HPV-16 
early 3' UTR probe TUP (Fig. 7E). In addition, recombinant 
hnRNCP C1 binds to the HPV-16 early 3' UTR and the TUP 
RNA competes efficiently with the probe, whereas the TDU 
RNA did not (Fig. 7F), confirming that hnRNCP C1 binds to 
the U-rich region of the 3' UTR. In addition to hnRNCP C1/C2 
and CstF-64, PTB has also been shown to bind to U-rich USE 
elements and to stimulate polyadenylation (30). PTB may also 
compete with CstF-64 for RNA substrates, in this case inhib-
iting polyadenylation (8). The preferred binding site of PTB is 
UCUU, and there is one UCUU motif present in the U-rich 
region of the HPV-16 early 3' UTR. Cross-linking of GST-
PTB to the TUP and TDU probes revealed that PTB interacted 
preferentially with the UP probe, which contains the 
UCUU sequence (Fig. 7G), suggesting that PTB may affect 
the polyadenylation efficiency of the early HPV-16 mRNAs. HuR 
does not bind to this region. In conclusion, hnRNCP C1/C2, 
CstF-64, and PTB interact specifically with the 57-nucleotide 
USE element upstream of the HPV-16 early polyadenylation 
signal, suggesting a role for these proteins in the polyadenyl-
ation of the early HPV-16 mRNAs.

The U-rich region in the HPV-16 early 3' UTR interacts with 
the recently identified CPSF-160-associated polyadenylation 
factor hFip1. hFip1 is a newly discovered polyadenylation fac-
tor which is an integral subunit of CPSF (24). It binds CPSF-
160 and PAP but also binds directly to U-rich stimulatory 
upstream RNA elements through its arginine-rich C terminus 
(24). Binding of hFip1 to U-rich upstream elements causes an 
hFip1-mediated stimulation of PAP activity, thereby enhanc-
ing polyadenylation (24). The authors speculated that hFip1 
has remained undiscovered until recently due to its low abun-
dance in the CPSF complex (24). We therefore tested binding 
of recombinant his-tagged hFip1 to the USE in the HPV-16 
early 3' UTR. Two radiolabeled RNAs, named TUP and TDU, 
were synthesized in vitro (Fig. 7A) and were cross-linked to 
his-tagged hFip1. As can be seen from the results, hFip1 in-
teracted with TUP RNA, which had the U-rich USE, but not 
with TDU, which lacks the U-rich region in the early 3' UTR 
(Fig. 8A). These results demonstrated that hFip1 bound spe-
cifically to the HPV-16 USE. To confirm this finding, a com-
petition experiment was performed. RNAs TUP and TDU 
served as competitors in cross-linkings of hFip1 to TUP. This 
experiment showed that TUP RNA competed for binding to 
hFip1, whereas TDU, which lacked the U-rich region, did not 
(Fig. 8B). These results confirmed that hFip1 binds specifically 
to the USE in the HPV-16 early 3' UTR and suggested that the stimulatory effect of the HPV-16 USE is 
mediated by the hFip1 protein.

DISCUSSION

A 3' splice site was identified at position 742 in the HPV-16 
genome. Use of this splice site, in combination with the 5' 
splice site at position 226, results in the production of an E1 
mRNA or an E4 mRNA, if splicing between the 5' splice site 
at position 80 and the 3' splice site at 3358 occurs (Fig. 2A). 
These E1 and E4 mRNAs would produce higher levels of E1 
and E4 proteins, respectively, than previously identified E1 and 
E4 mRNAs. The splicing between the 5' splice site at position 
226 and the 3' splice site at position 742 removes the E7 AUG 
and creates E1 and E4 mRNAs, on which the E1 and E4 open
reading frames are preceded only by the suboptimal E6 AUG. This AUG does not conform to the consensus AUG for optimal initiation of translation (28). The E6 AUG is further weakened by its close proximity to the 5′ end of the mRNA (48, 62). Therefore, leaky scanning at the E6 AUG allows for efficient translation of downstream open reading frames (36). A 3′ splice site at the corresponding position has not been identified in any HPV type, which raises the question of

FIG. 7. (A) Schematic drawing of the in vitro-synthesized RNAs. The sequences in plasmids TUP, TDU, and TDP represent sequences from the HPV-16 early 3′ UTR. Numbers refer to the nucleotide positions in HPV-16R (6). The AAUAAA-to-ACGCGU mutation of the pAE in pTDP is indicated. The T2B2 RNA contains four HuR binding sites (underlined) (43), the T2C1 RNA contains six hnRNP C1/C2 binding sites (underlined) (44, 45), and the TCstF64 RNA contains two CstF-64 binding sites. T7, bacteriophage T7 RNA polymerase promoter. (B) Left panel, UV cross-linking of HeLa nuclear extract to the indicated RNA probes. The L1 RNA is a non-U-rich sequence derived from the HPV-16 L1 coding region and served as control. Arrows indicate four proteins, named A, B, C, and D, detected with the TUP probe but not with the TDU probe. For the right panel, radiolabeled TUP probe was UV cross-linked to HeLa nuclear extract in the absence or presence of threefold serially diluted unlabeled TUP or TDU competitor RNAs as indicated. The molar excess of the competitor over probe was 1-, 3-, 9-, 27-, and 81-fold. (C) UV cross-linking of HeLa nuclear extract to the indicated RNA probes. HnRNP C is indicated. (D) Radiolabeled CstF64 probe was UV cross-linked to HeLa nuclear extract in the absence or presence of threefold serially diluted unlabeled TUP, TDU, or TCstF64 competitor RNA as indicated. The molar excess of the competitor over probe was 1-, 3-, 9-, 27-, and 81-fold. CstF-64 is indicated. (E) Radiolabeled TUP probe was UV cross-linked to his-tagged hnRNP C1 in the absence or presence of threefold serially diluted unlabeled TUP or TDU competitor RNAs as indicated. The molar excess of the competitor over probe was 1-, 3-, 9-, 27-, and 81-fold. His-hnRNP C1 is indicated. (G) UV cross-linking of GST-PTB to the indicated RNA probes.
whether this splice site is used in HPV-16-infected cells. During the preparation of this report, Ordonez et al. (32) published the identification of mRNAs in HPV-16-infected cells which used the same splice site as we have identified here, demonstrating that mRNAs spliced to the 3' splice site at position 742 are made in the viral life cycle.

The early polyadenylation signal of all known HPVs is unique in that it is the only polyadenylation signal that is used...
at an early stage in the infection, whereas at later stages, usage of the downstream late polyadenylation signal is activated at the expense of the early polyadenylation signal. It is therefore not unlikely that the pAE is under the control of multiple regulatory elements. Most polyadenylation signals contain an AAUAAA sequence followed by the cleavage site and a U- or G/U-rich downstream element (DSE) (11, 54, 60). Many polyadenylation signals also contain auxiliary enhancer elements upstream and downstream of the AAUAAA (11, 54, 60). The DSE is normally located within 30 nucleotides after the cleavage site and has the consensus sequence YGUGUUY (11, 54, 60). The polyadenylation factor CstF-64 binds to the downstream element and is required for the interaction between the AAUAAA element and CPSF (11, 54, 60). There is no obvious U- or G/U-rich element within a 30-nucleotide distance of the HPV-16 pAE. The question of how CPSF recognizes the AAUAAA in the HPV-16 early polyadenylation signal is therefore of interest. For HPV-31, a close relative of HPV-16, it has been proposed that three weak CstF-64 binding sites downstream of the pAE act in concert to ensure efficient polyadenylation at the HPV-31 pAE (53). It remains to be seen if HPV-16 has similar alternative downstream elements at the pAE. However, binding of CstF-64 to the HPV-16 early 3' UTR as seen here may aid in the recognition of the AAUAAA by CPSF-160 through interaction with the other polyadenylation factors.

Some known polyadenylation signals contain auxiliary sequences that modulate the efficiency of 3' processing (11, 54, 60). One example is upstream enhancer elements (USEs) (11, 54, 60). These elements are normally located in the 3' UTR sequence and have been found primarily on viral mRNAs encoded by various unrelated viruses, including adenoviruses, herpesviruses, hepadnaviruses, polyomaviruses, and retroviruses, but also on some cellular mRNAs, such as C2 complement, lamin B2, and histone H2a mRNAs (60). Although no consensus sequence for these elements has been derived, they are all U rich. The results presented here on the USE in HPV-16 showed that the U-rich region in the early 3' UTR was
the major enhancer element in the HPV-16 early 3′ UTR. Recently, a novel component of CPSF named hFip1 was identified (24). hFip1 interacts with U-rich upstream elements and stimulates polyadenylation (24). Our finding that hFip1 interacts with the HPV-16 early 3′ UTR strongly supports a role for hFip1 in the USE-mediated enhancement of early HPV-16 polyadenylation. It remains to be proven that hFip1 stimulates polyadenylation at the HPV-16 pAE in vitro. The HPV-16 USE also interacted with CstF-64 that, similarly to hFip1, stimulates polyadenylation by interacting with CPSF (11, 54, 60). One may speculate that both CstF-64 and hFip1 enhance polyadenylation at the HPV-16 pAE. The HPV-16 USE was also found to interact with other factors (hnRNP C1/C2 and PTB) that may modulate the activity of the USE. The USE on the C2 complement mRNA binds CstF-64 and PTB, as does the HPV-16 USE (30), suggesting that these factors may enhance polyadenylation. The C2 complement USE stimulates cleavage/polyadenylation by binding to PTB and CstF-64 (30). PTB appeared to increase the levels of CstF-64 that cross-linked to the USE (30). In contrast, binding of PTB to a downstream element prevented binding of CstF-64, thereby inhibiting polyadenylation (8). It is unlikely that PTB and CstF-64 compete for the same site in the HPV-16 USE. The roles of CstF-64 and PTB in HPV-16 early polyadenylation remain to be determined.

The USE identified here in HPV-16 is also similar to the USE in the adenovirus L3 polyadenylation signal, in that it contains one or more stretches of five or more consecutive U nucleotides (35). The sequence UUCUUUUU, present in the adenovirus L3 upstream region, was shown to enhance processing efficiency as well as stable complex formation by increasing binding of CstF-64 to the RNA (35). The U-rich sequence itself acted as a binding site for the hnRNP C1/C2 proteins (35). Runs of uridylates are present in the HPV-16 USE as well, and they interacted with hnRNP C1/C2. Further experiments are needed to determine if hnRNP C1/C2 af-
ffects polyadenylation at the HPV-16 pAE. hnRNP C1/C2 has been implicated in polyadenylation of late simian virus 40 (SV40) mRNAs (57). The downstream domain of the SV40 late poly(A) addition signal has been shown to influence the efficiency of the polyadenylation reaction (57). Uridyly-rich sequences located downstream of the cleavage site were required for efficient cross-linking of hnRNP C1/C2 (56). However, there was not a strict correlation between cross-linking of hnRNP C1/C2 proteins and the efficiency of polyadenylation (56). The 5-base uridyly tract restored efficient in vitro processing to several independent poly(A) signals in which it was substituted for downstream element sequences. Although hnRNP C1/C2 proteins specifically interacted with U-rich sequences, this interaction did not appear to be required for efficient in vitro polyadenylation (10). Others noted loss of activity of the SV40 polyadenylation signal while high levels of hnRNP C1/C2 were cross-linking, raising questions about the specificity of the interaction between the hnRNP C1/C2 and polyadenylation precursor RNAs in vitro (49). Analysis of the HPV-16 pAE in an in vitro polyadenylation assay may allow us to determine if hnRNP C1/C2 binding to the upstream element is involved in polyadenylation of early HPV-16 mRNAs.

Changing the AAUAAA of the HPV-16 pAE to AGCGGU resulted in a complete abolishment of cleavage and polyadenylation at the cleavage site used by the pAE. However, cleavage and polyadenylation was induced at multiple positions upstream of the pAE. The dominating cleavage site was at genomic position 3820. In spite of the mutational inactivation of the pAE, the vast majority of the HPV-16 mRNAs were still polyadenylated in the early region. Therefore, strong regulatory RNA elements presumably direct the polyadenylation machinery to cryptic sites in the vicinity of the pAE. Because all cryptic polyadenylation sites were located upstream of the pAE, these regulating elements are probably located downstream of the pAE. Similar observations have been reported for BPV-1 (1). One article presented results on an AAUAAG to UGUAAA mutation of the BPV-1 pAE which did not result in significant read-through at the pAE (1). This was shown to be caused by the selection of cleavage and polyadenylation sites at multiple positions approximately 100 nucleotides upstream of the pAE (1). Another groups reported that deletion of the BPV-1 early polyadenylation signal activated the use of alternative polyadenylation signals in the early region of the BPV-1 genome (7). Mutational inactivation of the pAE in HPV-31 did not eliminate production of early mRNAs (52), suggesting that cryptic polyadenylation signals in the early region were activated. In conclusion, mutational inactivation of the pAE in HPV-16, HPV-31, and BPV-1 yields similar results, indicating the existence of strong and conserved regulatory elements that direct the polyadenylation machinery to this region of the papillomavirus genome. Other results on BPV-1 showed that transcripts that bypass the pAE terminate before the RNA polymerase reach the late polyadenylation signal, suggesting that the region downstream of the pAE contains elements that either cause termination or induce pausing and termination (4). Pausing downstream of a polyadenylation signal has recently been shown to enhance polyadenylation (58). We have seen polyadenylation of a cryptic polyadenylation signal in L1 in the late region (31, 37). Results obtained with HPV-31 suggested that a large sequence of 800 nucleotides in HPV-31 L2 is prerequisite for polyadenylation at the pAE (52). Preliminary data on HPV-16 from our group indicate that the L2 coding region is of paramount importance for polyadenylation at the HPV-16 pAE (data not shown). The element in the early 3′ UTR enhances polyadenylation at the pAE, but it is not necessary for polyadenylation. Because the pAE has a key role in the switch from the early to late gene expression as the cell differentiates, it is also possible that the activity of the pAE is under the control of multiple elements. It remains to be determined how the various RNA elements interact and how they regulate polyadenylation at the pAE.

The 3′ UTR often contains RNA elements that affect or regulate mRNA translation, localization, or stability (13). For example, the late 3′ UTR of HPV-1 reduces mRNAs half-life (45) and inhibits translation (55). The HPV-16 early 3′ UTR may affect RNA processing steps other than polyadenylation. Previous results have shown that integration of the HPV-16 genome into cellular chromosomes resulted in stabilization of the E6 and E7 mRNAs due to deletion of the early 3′ UTR by the integration event (22, 23). In addition, insertion of the early 3′ UTR after the human beta-globin gene in the c-fos promoter expression system caused a reduction of the mRNA half-life compared to that of the mRNA lacking the HPV-16 early 3′ UTR (23). This insertion resulted in a more than sixfold reduction in mRNA half-life. In the pBearly- and pBEL-derived constructs analyzed here, we saw an increase in mRNA steady-state levels when the early 3′ UTR was deleted. However, the effect was relatively small, in all cases less than a twofold increase in mRNA levels. Previously, we reported that the insertion of the HPV-16 early 3′ UTR downstream of the HIV-1 p17gag reporter gene had no effect on expression levels of the reporter gene, whereas the late 3′ UTR had a 50- to 100-fold inhibitory effect (59). The early 3′ UTR does not contain the classical AUAAUUAA instability elements, but long U stretches also confer short half-life (9, 46). The reason for the difference between our results and those of Jeon et al. (22, 23) is likely to lie in the different experimental systems used. While we use transient transfections of reporter constructs with the constitutively active human CMV immediate-early promoter into HeLa cells, Jeon et al. (22, 23) used more sensitive, inducible c-fos promoter-driven plasmids stably transfected into the murine fibroblast NIH 3T3 cell line (23). Sequences that affect the mRNA half-life allow for regulation of expression levels, which may be necessary to ascertain the ordered expression of the early genes during the viral life cycle. It will be of interest to investigate if any of the proteins that bind to the early 3′ UTR also affect the half-life of the mRNAs. We have previously shown that hnRNP C1/C2 interacts with RNA instability elements in the HPV-1 late 3′ UTR (45), but there was no strict correlation between hnRNP C1/C2 binding and short RNA half-life (44).

The HPV-16 early 3′ UTR is very U rich, even compared to many other HPV types. In general, the early 3′ UTRs of all papillomaviruses are relatively pyrimidine rich, indicating that the presence of CU-rich regions upstream of the pAE plays an important role in the virus life cycle. A comparative functional analysis of the importance of the upstream sequences in early 3′ UTR in polyadenylation of various HPV will reveal if the USE is conserved among the HPVs. We have also identified a number of cellular factors (hFip1, CstF-64, hnRNP C1/C2, and
PTB) that interact specifically with the early 3′ UTR. Analysis of the HPV-16 polyadenylation signal in an in vitro cleavage/polyadenylation assay will determine if these factors are involved in polyadenylation of the early HPV-16 mRNAs. As usage of the early polyadenylation signal should decrease in response to keratinocyte differentiation, one may speculate that if any of the factors identified here are involved in the regulation of early polyadenylation, the levels of this factor may change in response to differentiation. Different HPV s also show onset of late gene expression at different levels in the epithelium (33), suggesting that regulatory elements that are involved in the early-to-late switch may differ between types. The elements may interact with different factors, or they may have different affinities for the same factor caused by sequence variations in the RNA elements. It will be of interest to study the role of the HPV-16 USE in response to cell differentiation to determine its role in the viral life cycle and in an in vitro polyadenylation assay to establish the importance of the various factors that were found to interact with the HPV-16 USE.

In conclusion, we have identified a USE element that spans the U-rich region in the HPV-16 early 3′ UTR. This USE interacts with hFip1L, CstF-64, hnRNP C1/C2, and PTB and enhances polyadenylation at the early polyadenylation signal 5′EAE.

ACKNOWLEDGMENTS

We thank W. M. LeStourgeon and W. Keller for plasmids encoding his-tagged hnRNP C1 and hFip1.

This research was sponsored by the Swedish Cancer Society and by the Swedish Research Council.

REFERENCES