The E2 protein of human papillomavirus type 16 is translated from a variety of differentially spliced polycistronic mRNAs

Nathalie Alloul and Levana Sherman

Department of Human Microbiology, Sackler School of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel

The major regulation protein of human papillomavirus (HPV) transcription is the viral E2 protein. Previous studies have identified a variety of alternatively spliced mRNAs containing multiple open reading frames (ORFs) encoding the E2 protein of HPV type 16. In these mRNAs the E2 ORF is contained as an internal ORF. In the present study, the translational capacities of three mRNA species starting at the p97 promoter and containing the 880/2581, 880/2708 and 226/2708 splice junctions upstream of the E2 ORF were investigated. Partial cDNAs spanning the E2 ORF and the related upstream ORFs were synthesized and assessed for E2 protein translation in vivo, in COS cells, and in vitro, in cell-free systems. Results of these analyses indicated that E2 protein was translated from all three mRNAs. Translation efficiency of E2 from the natural polycistronic templates was lower compared with that from a synthetic monocistronic control. Translation from the d-type bicistronic template (226/2708) was more efficient than that from the α-type (880/2708) and α-" (880/2561) polycistronic templates. Further investigation of the translation of proteins encoded by the ORFs preceding the E2 ORF showed that α- and α-"-type templates served for translation mainly of E7 but also of E6I, while the d-type template served for translation of E6IV. Overall, the translation data support the suggestion that the corresponding mRNAs may function as polycistronic transcripts.

Introduction

Human papillomaviruses (HPVs) are small DNA viruses which induce benign and malignant epithelial tumours. HPV type 16 is the prototype of the ‘high-risk’ group of HPVs, members of which are associated with genital pre-cancerous lesions and carcinomas (de Villiers, 1989; zur Hausen, 1989). In tumours and cell lines containing HPV-16 DNA, the viral genome is actively transcribed (Schwarz et al., 1985; Yee et al., 1985). Previous studies have defined the structure of early and late mRNAs encoded by the HPV-16 genome (Smotkin & Wettstein, 1986; Shirasawa et al., 1988; Smotkin et al., 1989; Doorbar et al., 1990; Rohlfis et al., 1991; Sherman et al., 1992; Sherman & Alloul, 1992). These studies revealed a large number of putative polycistronic mRNAs, generated by alternative splicing and encoding viral early and late gene products.

The viral p97 promoter, located upstream of the E6 open reading frame (ORF), is the major early promoter from which transcription is initiated, although initiation from intragenomic promoters has also been described (Higgins et al., 1992; Bohm et al., 1993; Nilsson et al., 1996). The majority of mRNAs generated from the p97 promoter contain ORFs of the E6 and E7 transforming genes, including the full-length E6 and the alternatively spliced E6I (226/409) and E6II (226/526) ORFs (Smotkin & Wettstein, 1986; Smotkin et al., 1989). The E6 and E7 ORFs are spliced to various downstream (3’) ORFs via three alternative splice sites at 880/3357, 880/2708 and 880/2581 (Doorbar et al., 1990; Nasseri et al., 1991; Rohlfis et al., 1991; Sherman & Alloul, 1992). Short mRNA species initiating at the p97 promoter have also been described. These include the E6III and E6IV transcripts, in which the N-terminal portion of the E6 ORF is joined to out-of-frame sequences within the E4 (226/3357) and E1 (226/2708) ORFs, respectively (Doorbar et al., 1990; Sherman & Alloul, 1992).

Little is known about the translational capacity of these HPV transcripts. In eukaryotes, most of the mRNAs are monocistronic. Translation from these mRNAs is usually cap-dependent and translation initiates at the AUG codon nearest to the 5’ end of the mRNA, conforming with Kozak’s scanning model for translation initiation (Kozak, 1989). Viral translation, on the other hand, occurs in many cases from polycistronic
mRNAs (Kozak, 1986). Translation from downstream AUGs in polycistronic mRNAs has been explained by two modifications of the scanning model: leaky scanning, by which upstream AUGs are ignored when lying in an unfavourable context, and a termination–reinitiation mechanism, by which scanning may be resumed after translation of the first cistron and, upon acquisition of the required factors, the start codon of the downstream ORF may be recognized (Kozak, 1989, 1991). In recent years, a third mechanism, internal ribosome entry, has been demonstrated for a variety of viral and some eukaryotic mRNAs (Jackson et al., 1990; Jackson & Kaminski, 1995; Sachs et al., 1997). The best documented internal ribosome entry sites (IRESs) are found in the naturally uncapped RNAs of picornaviruses (Jackson et al., 1990). These elements consist of long stretches of highly structured RNA which, in conjunction with associated proteins, guide ribosome entry directly to the initiating AUG.

We have shown previously that HPV-16 expresses a variety of alternatively spliced mRNAs containing the full-length ORF of the viral E2 transcription regulatory protein (Sherman & Alloul, 1992). The structures of these mRNAs were characterized by RT–PCR and cDNA cloning. These studies showed that the majority of E2 transcripts begin at the p97 promoter and contain multiple ORFs, with the E2 ORF present as an internal ORF. The E2 ORF is connected to various upstream 5′ ORFs via alternative splice sites at 880/2708, 880/2581 and 226/2708, which are represented by RNA species a, a′ and d, respectively (Sherman & Alloul, 1992). RNA species a is the most abundant type expressed in tumours and cell lines containing HPV-16 DNA (Sherman et al., 1992; Sherman & Alloul, 1992).

According to Kozak’s scanning model of translation (Kozak, 1989, 1991), termination–reinitiation would be expected to occur from the a-type mRNA, as there is adequate space between the termination of the upstream E1 minicistron, at position 2722 (Rohlf et al., 1991), and the E2 initiator, at position 2755. In the d- and a′-type mRNAs, the E2 AUG is out of frame and upstream of the stop codon of the preceding ORF. Such a configuration would be expected to preclude termination–reinitiation (Kozak, 1987), although exceptions from this rule have been observed (Peabody et al., 1986).

In the present study, we investigated the translational capacities of cDNA constructs carrying the full-length E2 ORF and the related 5′ ORFs of the differentially spliced a-, a′- and d-type mRNAs. Translation experiments in vitro and in vivo indicated that all three constructs served as templates for translation of E2 protein, although at different efficiencies. The efficiency of translation of E2 from the authentic polycistronic templates was lower compared with the synthetic monocistronic E2 construct. Unexpectedly, E2 protein was more efficiently translated from the d-type construct (226/2708), despite the unfavourable configuration of the upstream ORF. Translation of E2 from the a- and a′-type templates was at similar levels. Finally, we also show that proteins encoded by the 5′ ORFs were co-translated with E2, including the E6I and E7 proteins encoded by RNA types a (880/2708) and a′ (880/2581) and the E6 IV protein encoded by RNA type d, indicating the potential of the corresponding transcripts to function as polycistronic templates.

Methods

Construction of recombinant clones. CDAs spanning the E2 and upstream ORF regions of HPV-16 mRNA species a, a′ and d (Fig. 1) were synthesized in a stepwise manner, employing RT–PCR as previously described (Sherman et al., 1992; Sherman & Alloul, 1992). First, CDAs spanning the variable 5′ regions of these mRNAs were synthesized from RNA from CaSki cells. The primer E6s-101 (which also included an XhoI restriction site at the 5′ terminus) was used as a sense primer for synthesis of all three CDAs. The antisense primer E1as-2930, downstream of the 880/2708 and 226/2708 splice sites, was used for the synthesis of the a- and d-type CDAs. The antisense primer E1as-2684, downstream of the 880/2581 splice site, was used for synthesis of the a′ species.

To generate full-length CDAs of species a and d, their 5′-region CDAs (RT–PCR products) were fused to the full-length E2 ORF region by PCR. An AseI fragment spanning nucleotides 2565–3928 was excised from a genomic clone of HPV-16 and combined with the RT–PCR products by PCR. The overlapping region between the CDAs and the E2 fragment served as a ‘megaprimer’ (Herlitz & Koenen, 1990) in the first cycle of the cDNA extension. Amplification of the full-length fused CDAs was carried out using the E6s-101 primer and the E2as-3852 primer (which included the sequence of the E2 ORF stop codon and a HindIII restriction site) or the E2as-3852 Tag primer, which included the sequence of the six amino acid AU1 epitope (Lim et al., 1990) downstream of and in-frame with the sequence of the E2 ORF, followed by the termination codon and a HindIII restriction site. The resulting CDAs were cut with XhoI and HindIII and cloned into a modified pSG5 vector (Stratagene), pJS55 (Sparkowski et al., 1994), cut with the same enzymes.

To generate the full-length cDNA of species a′, the 5′ region cDNA (RT–PCR product) was first elongated by the same procedure described above, but using the E2as-2930 primer. The resulting cDNA was then cut with XhoI and BstXI (BstXI cuts at position 2891) and fused with the E2 ORF or the AU1-tagged E2 ORF by replacement of the XhoI–BstXI fragment in the d-type construct.

All clones were sequencing using the Sequenase II kit (USB). Only clones containing the correct sequence of the E2 ORF region were used for further analysis. Nucleotides changes were detected in the a-type cDNA at positions 131 (G for A) and 659 (G for A) and in the a′-type cDNA at positions 131 (G for A) and 2608 (T for G).

To obtain the monocistronic E2 construct, we used a subgenomic clone, pRC/RSV-E22, containing the AseI fragment of HPV-16 from nucleotides 2565–3928 cloned in the vector pRC/RSV (Invitrogen). The genomic fragment spanning nucleotides 2731–3928 was excised from the pRC/RSV-E2 construct by AsuII and BamHI digestion. The resulting fragment, containing the full-length E2 ORF, was made blunt by filling in with Klenow DNA polymerase and cloned into the pJS55 vector at the Smal site. The cloning orientation with respect to the T7 or T3 RNA polymerase promoter was confirmed by nucleotide sequencing. All the E2 CDAs described above were also subcloned into the pGEM vector (Promega) cut with BamHI and HindIII.

To produce isogenic mutant clones containing a premature termination codon in the E2 ORF, the E2 CDAs cloned into the pJS55 vector were cut with BstXI, blunted with T4 DNA polymerase and self-ligated, thus creating a stop codon at position 2984.
Expression of the E2 protein in bacteria and production of antibodies. To enable expression in bacteria, the C-terminal region of the E2 ORF, spanning nucleotides 3697–3852, was cloned into the bacterial expression vector pET-17xb (Novagen). The E1C ORF (Sherman & Alloul, 1992) was synthesized and cloned into the pJS55 vector (A. Gluchov, unpublished data).

Expression in COS cells, metabolic labelling and immunoprecipitation. Ten µg DNA of the various expression vectors was transfected into COS cells grown in 9 cm dishes by using a calcium phosphate co-precipitation procedure as previously described (Sherman & Schlegel, 1996). At 72 h post-transfection, cells were analysed for protein expression. Cells were metabolically labelled for 4 h with 570 Ci in a commercial mixture of [35S]methionine (90%) and [35S]cysteine (10%) (Express, NEN) and then lysed in 1 ml RIPA buffer (50 mM Tris–HCl, pH 8, 150 mM NaCl, 0.5% sodium deoxycholate, 1% NP40, 0.25% SDS) containing 100 µg PMSF and 10 µg aprotinin. Protein extracts from metabolically labelled cells or cell-free systems were pre-cleared with non-immune rabbit serum (diluted 1:50) and immunoprecipitated with the specific antibody (diluted 1:50 or 1:100) as described previously (Shally et al., 1996). Rabbit antibodies prepared against bacterially expressed fusion proteins E6I, E6IV and E7 (Shally et al., 1996; Barbosa et al., 1991) were used for immunoprecipitation of the corresponding proteins, expressed either from the native or epitope-tagged ORFs. Immunoprecipitates were resolved on 15% or 10% SDS–PAGE gels. Were dried and exposed to X-ray films at −70 °C for various lengths of time.

Results

In order to investigate the translational capacities of the a-, a’- and d-type polycistronic mRNAs of HPV-16 (Fig. 1), cDNAs spanning the full-length E2 ORF and the related 5’ ORFs were synthesized and cloned into the pJS55 vector. Constructs were assessed for expression of the E2 protein and

pJS55 vectors carrying monocistronic cDNAs spanning the individual ORFs E7 and also E6I and E6IV tagged with the AU1 epitope were described previously (Sherman & Schlegel, 1996). A cDNA spanning the E1C ORF (Sherman & Alloul, 1992) was synthesized and cloned into the pJS55 vector (A. Gluchov, unpublished data).

Expression of the E2 protein in bacteria and production of antibodies. To enable expression in bacteria, the C-terminal region of the E2 ORF, spanning nucleotides 3697–3852, was cloned into the bacterial expression vector pET-17xb (Novagen). The E1C ORF (Sherman & Alloul, 1992) was synthesized and cloned into the pJS55 vector (A. Gluchov, unpublished data).

Expression in COS cells, metabolic labelling and immunoprecipitation. Ten µg DNA of the various expression vectors was transfected into COS cells grown in 9 cm dishes by using a calcium phosphate co-precipitation procedure as previously described (Sherman & Schlegel, 1996). At 72 h post-transfection, cells were analysed for protein expression. Cells were metabolically labelled for 4 h with 570 Ci in a commercial mixture of [35S]methionine (90%) and [35S]cysteine (10%) (Express, NEN) and then lysed in 1 ml RIPA buffer (50 mM Tris–HCl, pH 8, 150 mM NaCl, 0.5% sodium deoxycholate, 1% NP40, 0.25% SDS) containing 100 µg PMSF and 10 µg aprotinin. Protein extracts from metabolically labelled cells or cell-free systems were pre-cleared with non-immune rabbit serum (diluted 1:50) and immunoprecipitated with the specific antibody (diluted 1:50 or 1:100) as described previously (Shally et al., 1996). Rabbit antibodies prepared against bacterially expressed fusion proteins E6I, E6IV and E7 (Shally et al., 1996; Barbosa et al., 1991) were used for immunoprecipitation of the corresponding proteins, expressed either from the native or epitope-tagged ORFs. Immunoprecipitates were resolved on 15% or 10% SDS–PAGE gels. Were dried and exposed to X-ray films at −70 °C for various lengths of time.

Results

In order to investigate the translational capacities of the a-, a’- and d-type polycistronic mRNAs of HPV-16 (Fig. 1), cDNAs spanning the full-length E2 ORF and the related 5’ ORFs were synthesized and cloned into the pJS55 vector. Constructs were assessed for expression of the E2 protein and

Vaccines and immunisation experiments were carried out as previously described (Shally et al., 1996). The resulting [35S]-labelled translation products were separated on SDS–PAGE gels or immunoprecipitated prior to gel analysis (described below). Intensities of the protein bands were quantified by densitometry with the biological imaging system 202 (Dinco & Rhenium).
proteins encoded by the 5' ORFs by using in vitro transcription and translation systems and transfection into COS cells. Monocistronic constructs containing cDNAs of the individual ORFs were used as controls.

**In vitro analysis of the d-, a- and a'-type templates for protein expression**

The efficiency of translation of E2 protein in the in vitro systems, WGE and RTL, was first evaluated by testing translation from the monocistronic E2 construct (pJS55-E2). Various amounts of the in vitro transcription product (0.5% and 5% of the product obtained from transcription of 1 µg cDNA) were used for translation. A protein with the expected size for E2 (about 42 kDa) was produced in both RTL and WGE systems. The amount of translated protein correlated with the amount of RNA applied (data not shown). Translation of E2 protein was much more efficient in RTL compared with WGE (Fig. 2a). The translated product was immunoprecipitated specifically with two different antibodies directed against the E2 protein (Fig. 2b). Immunoprecipitation was more efficient with the antibody raised against the C-terminal end of the E2 protein (E2C) (described in Methods). Consequently, further immunoprecipitations were carried out with the E2C antibody.

Translation of the d-type cDNA, containing the E6IV and E2 ORFs (clone E6IV–E2), in the RTL or WGE systems resulted in the production of both proteins. The larger product co-migrated with E2 protein produced from the E2 monocistronic control and was immunoprecipitated specifically with the E2C antibody (Fig. 3a), confirming its identity. This product was not translated from a mutant d-type cDNA (d-mut) that carried a stop codon in the E2 ORF (Fig. 3a). The amount of E2 protein translated from the bicistronic cDNA was lower than that translated from the monocistronic control. The second product obtained from translation of the E6IV–E2 clone was identified as the E6IV protein. This product was translated more efficiently in WGE. Fig. 3(b) shows that the E6IV protein translated in WGE co-migrated with the product of the monocistronic E6IV control (pJS55-E6IV) and was immunoprecipitated specifically by the E6IV antibody.

In a similar approach to that applied to the d-type cDNA, the a- and a'-type cDNAs were examined first for E2 protein expression. Each translation assay included the monocistronic E2 template (pJS55-E2) as a control. Fig. 4(a) shows analysis on a 10% SDS–PAGE gel of the products obtained from translation of the RNA templates in RTL. A band with the expected size for the E2 protein (42 kDa), which co-migrated with the translation product of the monocistronic E2 RNA, was translated from both a- and a'-type RNAs. The identity of the translated products was confirmed by immunoprecipitation with the E2C antibody (data not shown). Amounts of the protein obtained from translation of equal amounts of RNA of the a- and a'-type templates were similar and were much lower than that obtained from the monocistronic E2 RNA. This protein was not translated from mutant a- and a'-type RNAs that carried stop codons in the E2 ORF at position 2984 (data not shown). A strong band with the expected size for E7 was also detected among the a- and a'-type translation products produced in RTL (Fig. 4a). The a- and a'-type cDNAs were also analysed for the translatable ability of the ORFs upstream of E2, including E6, E7 and E1C. Translation was carried out in WGE, as it has been shown to support the translation of the E6 splice-variant proteins (E6I–E6IV) more efficiently (Shally et al., 1996). Fig. 4(b) shows 15% SDS–PAGE analysis of translation products obtained in WGE. The translation patterns of the a- and a'-type RNAs were identical. Translation of both RNAs yielded proteins with the apparent molecular masses of E7 and E6, as evident from their parallel migration with products of the respective monocistronic controls. The identities of the proteins were verified by immunoprecipitation with specific antibodies. Fig. 5 shows immunoprecipitation of E7 and E6I proteins obtained from translation of the a-type template. Both proteins were also immunoprecipitated from the a'-type RNA translation extract (data not shown). No band that might represent the putative 9 kDa product of the E1C ORF was detected in the a'-type RNA translation extract, despite the fact that such a band was detected after translation of the monocistronic RNA (data not shown). A small peptide of 40 amino acids that might result from translation of the E1 minicistron (Rohlf et al., 1991) was also not detected. These results suggest that neither the E1C ORF in the a'-type mRNA nor the E1 minicistron in the a-type mRNA were translated from the monocistronic mRNAs. In all the translation experiments using a- and a'-type templates in both WGE and RTL, the E7 protein was expressed correctly and at higher levels than the E2 and E6I proteins.

**Protein expression in COS cells from the d-, a- and a'-type cDNAs**

Further support for the results obtained from translation experiments carried out in cell-free systems was obtained by in
**vivo** expression in **vivo** expression in **vivo** expression in COS cells. **vivo** translation experiments conducted in COS cells transiently transfected with the d-, a- and a'-type cDNAs revealed expression of the E2 protein from all three polycistronic templates. Translation of E2 was confirmed in cells transfected with plasmids encoding the wild-type (Fig. 6a, lanes 1–4) or the AU1-tagged (Fig. 6a, lanes 5–8) E2 protein by probing with the E2C or AU1 antibodies, respectively. E2 expression from the polycistronic cDNAs was much lower than that obtained from the E2 monocistronic construct (Fig. 6a, lane 2). Expression of E2 from the d-type cDNA (Fig. 6a, lane 6) was higher than that from the a- or a'-type constructs. The possibility that E2 expression from the a'- and a-type cDNAs resulted from mRNA splicing that gave rise to the d-type transcript was excluded by RT–PCR analysis. Using the primer pair E6s-101 and E1as-2930, no product with the expected size for the E6IV splice (226/2708, 248 bp) was detected (data not shown). In addition to the E2 protein, expression of E7 protein was readily detected in COS cells transfected with the a- and a'-type cDNAs, as well as in cells transfected with the E7 monocistronic control (Fig. 6c). Similarly, expression of E6IV was detected in cells transfected with the d-type cDNA and the monocistronic control (Fig. 6b). Immunoprecipitation with the E6I antibody failed to detect expression of E6I protein in COS cells transfected with the a- or a'-type cDNAs or the monocistronic control (data not shown). These findings are consistent with previous studies, which failed to detect the E6I protein in **vivo** (Androphy et al., 1987; Schneider-Gadicke et al., 1988), although expression in COS cells of an epitope-tagged E6I protein from a monocistronic construct was detected with the AU1 antibody (Sherman & Schlegel, 1996).

Overall, the in **vivo** and **vivo** translation experiments with d-, a- and a'-type polycistronic cDNAs support the conclusion that these cDNAs may serve as templates for translation of the E2 protein from an internal ORF. However, comparing the relative translation efficiencies of the proteins encoded by each of the transcripts, the d-type mRNA appears to serve as template for both E6I and E2 proteins whereas the a- and a'-type mRNAs appear to function mainly for translation of E7.

**Quantitative analysis of E2 protein translation from the d-, a- and a'-type templates**

The results described above indicate that E2 was translated from all three polycistronic mRNAs. To determine the effect of differential splicing and configuration of the upstream ORFs on the efficiency of translation of E2 protein more accurately, a controlled quantitative analysis of E2 protein translation from
Fig. 5. Immunoprecipitation analysis of proteins translated in WGE from the polycistronic a-type RNA. Translation products were divided into three aliquots. The first aliquot was directly electrophoresed (lanes 1) and the other two were immunoprecipitated with normal rabbit serum (lanes 2) or a specific antibody (lanes 3) prior to electrophoresis on a 15% SDS–PAGE gel. Results of immunoprecipitation with the E7 antibody (a) and the E6I antibody (b) are shown. Short arrows indicate the position of E6I (a) and E7 (b).

Fig. 6. Expression of the d-, a- and a’-type cDNAs in COS cells and immunoprecipitation of the products. COS cells were transfected with 10 µg DNA and analysed for protein expression by metabolic labelling and immunoprecipitation. (a) E2 expression from cDNAs encoding the wild-type (lanes 2–4) or epitope-tagged (lanes 6–7) E2 protein. Lanes 1–4 and 5–8 show immunoprecipitations with the E2C or AU1 antibodies, respectively, from cells transfected with pJS55 DNA (lanes 1, 5) or the E2 monocistronic (lane 2), a-type (lanes 3, 7), a-type (lanes 4, 8) and d-type (lane 6) cDNAs. (b) Immunoprecipitation with the E6IV antibody from cells transfected with the vector pJS55 (lane 1), the monocistronic E6IV cDNA (lane 2) and the d-type cDNA (lane 3). (c) Immunoprecipitation with the E7 antibody from cells transfected with the monocistronic E7 cDNA (lane 1) as a positive control or the a-type (lane 2) and a’-type (lane 3) cDNAs. Positions and sizes of molecular mass markers are indicated on the left.

Fig. 7. Quantitative analysis of the E2 protein translated from the d-, a- and a’-type RNAs in RTL. Different amounts of the RNA products obtained from transcription of 1 µg cDNA were translated in RTL and translation products were analysed on a 10% SDS–PAGE gel. (a) Proteins translated from 0–3 (lane 1), 0–06 (lane 2) and 0–3 µg (lane 3) RNA. (b) Densitometric scanning of the E2 protein bands shown in (a). (c) Translation from 0–3 µg of the indicated RNAs transcribed from pGEM-based vectors.

the various templates was carried out in the RTL system. The RNA obtained from transcription of 1 µg cDNA was visualized on an agarose gel and quantified by spectrophotometry. Equal amounts of the various RNAs were used for translation in the same experiment.

Quantification of translation from amounts of RNA ranging from 0–03 to 0–3 µg revealed positive dose-dependent translation efficiency (Fig. 7a). The E7 protein, which was co-translated from the a- and a’-type mRNAs, and the E6IV protein, which was co-translated from the d-type mRNA, also showed a dose-dependent relationship in translation efficiency (data not shown).

After densitometric scanning of the E2 protein bands (Fig. 7b), several observations could be made. Translation of E2 protein was most efficient from the pJS55-E2 monocistronic control. Of the polycistronic templates, translation from the d-type was most efficient (48% of the amount of E2 protein obtained from translation of pJS55-E2). Translation of E2 protein from the a- and a’-type RNAs was similar, with slightly
more protein translated from the a’-type RNA (16% and 24% of the amount from the monocistronic RNA, respectively). The order of E2 translation efficiency observed was not affected by the expression vector used for cloning and transcription. Translation of equal amounts of the various RNA templates transcribed from the pGEM vector showed the same pattern of order of translation efficiency (Fig. 7c). Irrespective of the amount of RNA used for translation, the E7 protein was expressed at a consistently higher level than the E2 protein from both the a- and a’-type RNAs, consistent with the results described above.

Discussion

The E2 protein of papillomaviruses regulates viral gene transcription and replication (McBride et al., 1991; Thierry, 1996). The role of E2 in the virus replication cycle is not completely understood and there is limited information regarding the biosynthetic mechanisms involved in E2 protein expression. Previous reports have provided evidence that transcripts encoding multiple ORFs function as templates for synthesis of the full-length E2 protein of cottontail rabbit papillomavirus and HPV-11 (Barbosa & Wettstein, 1988; Rotenberg et al., 1989).

The data presented in this paper demonstrate for the first time that a number of alternatively spliced mRNAs containing multiple ORFs serve as templates for synthesis of the E2 protein of HPV-16 in cell-free systems and in COS cells. Controlled quantitative translation experiments carried out in RTL allowed the efficiency of E2 translation from the various transcripts to be determined. Unexpectedly, translation of E2 from the d-type template, representing the 226/2708 spliced mRNA, was more efficient than that from the a-type (880/2708) and a’-type (880/2581) transcripts. The terminator of E6IV is positioned downstream of the E2 initiator: thus, according to the scanning model (Kozak, 1987, 1989), translation of E2 is unlikely to occur by a termination–reinitiation mechanism, although reinitiation from a downstream ORF preceded by an overlapping short ORF has been reported (Peabody et al., 1986). Inspection of the contexts of the E6/E6IV and E2 initiation codons revealed that the E2 AUG is positioned in a more favourable context than that of E6/E6IV. The sequence around the E2 AUG agrees perfectly with the consensus at positions −3 (A) and +4 (G), which have been shown to be the major determinants of efficient initiation of translation (Kozak, 1991). Thus, in the d-type template, leaky scanning rather than reinitiation of translation appears more likely to direct E2 translation. The level of translation of E2 from the E7 ORF-containing a- and a’-type transcripts was similar, though it was slightly higher for the a’-type template, despite the different configurations of the upstream ORFs. In the a’-type template the E1C ORF overlaps the E2 ORF, while in the a-type mRNA the E2 initiator is positioned 33 nucleotides downstream of and in-frame with the preceding E1 minicistron, which could allow termination–reinitiation to occur (Kozak, 1987). The intercistronic spacing in the a-type template did not, however, increase E2 translation. Thus, it appears that in the a’- and a-type mRNAs the short cistrons upstream of E2 (the E1 minicistron or the E1C ORF, respectively) are not translated and reinitiation of translation, if it operates, occurs after termination of translation of the upstream E7 ORF. If this is the case, the greater distance between the E7 and E2 ORFs may allow translation reinitiation from both templates. This model is consistent with our inability to detect translation of either the E1C or E1 minicistron, while readily detecting the synthesis of E7.

Although we cannot exclude the possibility that E2 protein is translated from one or more of the polycistronic templates by a scanning-independent mechanism, such as internal entry (Jackson et al., 1990), several observations make this unlikely. If translation of E2 was directed by internal ribosome entry, efficient translation of E2 would be expected from the polycistronic templates. Yet, we observed that translation of E2 from all three natural polycistronic templates, a-, a’- and d-types, was less efficient than from the synthetic monocistronic control, both in RTL and in COS cells. Moreover, E2 translation was detected in WGE from all three templates, although at a lower efficiency compared with RTL. This is inconsistent with IRES-dependent translation, as no viral or cellular IRES has been reported to function in the wheat-germ system (Sachs et al., 1997). Finally, internal initiation of translation was shown to be dependent on the presence of an IRES, the secondary structure and also, presumably, tertiary structure of which play a critical role in internal ribosome entry (Jackson & Kaminski, 1995; Sachs et al., 1997). Our preliminary results (data not shown), indicating roughly equal efficiency of translation of E2 from heat-denatured (67 °C for 10 min) and renatured RNA templates, do not support a requirement for such a structure for E2 translation from any of the polycistronic templates.

Analysis of the proteins translated from the various cDNA templates supports the suggestion that the corresponding mRNAs may function as polycistronic templates. Since the truncated cDNA constructs used in this study lack the native mRNA sequences downstream of the E2 ORF, we were unable to evaluate the translation or effect of the E5 protein. However, previous studies have demonstrated the coordinated translation of E2 and E5 proteins from bicistronic constructs containing the full-length E2 and E5 ORFs of HPV-16 (Johnsen et al., 1995). We have shown that mRNAs containing the E7 ORF, although capable of functioning as polycistronic templates, served mainly for translation of E7. In both a- and a’-type mRNAs, the E7 ORF is preceded by the short E6I ORF. Translation of E7 from these templates is likely to occur as previously suggested, by a termination–reinitiation mechanism (Smotkin et al., 1989) or by leaky scanning (Stacey et al., 1995). E2 protein was translated with different efficiencies from the differently spliced transcripts. Efficiency of translation from the 226/2708 transcript was two to three times higher than...
from the 880/2708 and 880/2581 transcripts. Expression from all three polycistronic templates was lower than from the monocistronic control template. Monocistronic transcripts encoding E2 have so far not been detected, although evidence of the existence of promoters within the E7 ORF, which potentially could direct transcription of E2-encoding transcripts (Sherman & Alloul, 1992), has been provided (Doobhar et al., 1990; Higgins et al., 1992; Bohn et al., 1993, Nilsson et al., 1996).

The existence of multiple E2-encoding transcripts raises the possibility that alternative splicing and perhaps also alternative initiation act as mechanisms for controlling the level of E2 protein. The E2 protein is a potent activator of transcription. Trans-activation by E2 depends on it binding as a dimer to a palindromic sequence, ACCGNNNNCGGT, which is present in multiple copies in papillomavirus genomes (McBride et al., 1991). In genital human papillomavirus, E2 represses rather than activates transcription from the viral early promoters, positioned upstream from the E6–E7 ORFs (i.e. HPV-16 p97). The repression of transcription involves binding of the full-length E2 protein to E2 binding sites close to the TATA box, which apparently prevents binding of transcription factors required for formation of the transcription initiation complex (Romanczuk et al., 1990; Tan et al., 1992, 1994). Studies on regulation of the E6 promoter of HPV-18 by the homologous E2 protein have suggested that the regulatory effect may depend on the level of E2, activating at a low level and repressing as the level increases (Steger & Corbach, 1997). It is reasonable to assume that a low level of E2 will be needed at early phases of viral infection to allow early gene transcription in the suprabasal layers of the genital mucosa, while at later stages, when virus late promoters are presumably activated in the more differentiated layers, a high level of E2 may act to suppress further transcription from the p97 promoter. Consistent with this model is evidence showing the activity of a differentiation-dependent promoter located within the E7 ORF of HPV-16, which could possibly serve for transcription of an E2-encoding mRNA (Higgins et al., 1992; Nilsson et al., 1996). In HPV-18-associated keratinocytes, differences in the usage of splice-acceptor sites were observed between early infected keratinocytes and established lines. Transcripts corresponding to the d-type (E6IV, E2) and a-type (E6I, E7, E2) RNAs were only detected in the latter (Meyers et al., 1997). Further studies will be needed to establish whether expression of distinct E2-encoding transcripts possessing different translational capacities relates to stages of virus replication and cell differentiation.

We thank M. Muller and L. Gissmann (German Cancer Research Centre, Heidelberg) for their donation of antibodies. This research was supported, in part, by a grant from the Basic Research Fund, Tel-Aviv University, to Levana Sherman. The work was performed in partial fulfilment of the requirements for the PhD degree of Nathalie Alloul, Sackler Faculty of Medicine, Tel-Aviv University, Israel.

References


human papillomavirus type 16 immortalized cervical keratinocytes contain transcripts encoding E6, E7, and E2 initiated at the P97 promoter and express high levels of E7. *Virology* **184**, 131–140.


Received 11 June 1998; Accepted 15 September 1998