Available online at www.sciencedirect.com
SCIENCE DIRECT
Virus Research 118 (2006) 46–54
Virus Research
www.elsevier.com/locate/virusres

The intron 1 of HPV 16 has a suboptimal branch point at a guanosine

Marco Antonio De la Rosa-Rios [a], Martha Martínez-Salazar [a], Martha Martínez-Garcia [b], César González-Bonilla [c], Nicolás Villegas-Sepúlveda [a,*]

[a] *Depto. Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados, México*
[b] *Lab Bioquimica Molecular, UBIPRO, FES Iztacala, UNAM, México*
[c] *Unidad de Investigación Médica en Inmunología e Infectologia, Hospital de Infectología "Dr Daniel Méndez-Hernández", Centro Médico "La Raza ", IMSS, México*

Received 30 July 2005; received in revised form 11 November 2005; accepted 11 November 2005
Available online 15 December 2005

Abstract

The branch point sequence (BPS) of intron 1 of the HPV-16 was determined via RT-PCR in a cell free system, using lariat intermediates obtained by in vitro splicing reactions. We used synthetic E6/E7 transcripts and HeLa nuclear protein extracts to obtain the splicing intermediates. Then, a divergent oligonucleotide primer set, pairing on the lariat RNA that encompassed the 2′–5′ phosphodiester bond formed between the 5′ end of the intron and the BPS, was used for cDNA synthesis and PCR amplification. Subsequent RT-PCR assays revealed four splicing intermediates, made up of a major intermediary corresponding to the BPS and four cryptic branched sequences. Only intermediates bound at the 5′ end of the intron are probably the authentic branch point sequence, and all of them branch at guanosine 328 instead of the typical adenosine. Unusually, the BPS of intron 1 of HPV-16 is a suboptimal sequence (AGUGAGU) that differs from the eukaryotic consensus BPS, which correlates with the splicing profile observed for early transcripts of HPV-16 in tumors and tumor derived cell lines. The implications of this unusual branch point sequence for splicing of the HPV-16 pre-mRNA are discussed.

Keywords: HPV-16; Branch point; Splicing; BPS; Guanosine

1. Introduction

In eukaryotes, introns are removed from pre-mRNAs by a multi-protein complex of proteins and small nuclear RNAs (snRNAs) called the spliceosome (Kramer, 1995; Hastings and Krainer, 2001). Introns excision proceeds via a two step mechanism, of two sequential trans-esterification reactions: in the first step of the reaction, cleavage occurs at the donor site and an intron/exon 2 lariat is generated. In this reaction, the 5′ end of intron is linked by a 2′–5′ phosphodiester bound to an internal adenosine residue. The nucleotides neighbouring this adenosine are named the branch point sequence (BPS). In the second step of the reaction, the 3′ exon is cleaved from intron/exon 2 lariat and covalently linked to the 5′ of exon 1, yielding the spliced RNA. The excised intron remains as a lariat molecule (Keller, 1984; Ruskin et al., 1984) until its degradation by de-branching enzymes (Moore et al., 1993).

Before the first trans-esterification reaction occurs, the 5′ and 3′ ends of the intron are recognized by spliceosome factors, the BPS is bound and recognized by SF1 prior to binding of the U2 snRNP factor, a complex of several proteins and the U2 snRNA (Peled-Zahavi et al., 2001). During splicing, the ante-penultimate adenosine in the BPS is bulged out of the pairing to U2 snRNA, a condition that is essential for its reactivity (Pascollo and Seraphin, 1997). Thus, U2snRNP and other snRNPs (U4, U6 and U1) contribute to provide the structural requisites of RNA for recognition and nucleophilic attack to the 5′ end of the intron (Sharp et al., 1987). Finally, the branch point identification also depends on the recognition of the 5′ end of the intron (Michaud and Reed, 1993). The optimal consensus in yeast is 5′-UACURAC-3′ (Rain et al., 1998), while in mammals the stringency is lower, since BPS is 5′-YURAC-3′; Y: pirimidines; R: purines (Zhuang et al., 1989). Surprisingly, however, the yeast sequence remains optimal in both systems.

* Corresponding author at: CINVESTAV, Av IPN # 2508, Zacatenco, México City, Apdo Postal 14-740, CP.07360, México City, México.
Tel.: +5255 5061 3800x5023; fax: +5255 5061 3800x5015.
E-mail address: nvillega@cinvestav.mx (N. Villegas-Sepúlveda).

0168-1702/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2005.11.010

Human papillomaviruses (HPVs) are small DNA viruses containing a circular, double stranded DNA genome. More than 100 HPV genotypes that infect both cutaneous and mucosa epithelia have been identified so far. Since all HPVs possess genes that are transcribed as polycistronic messengers (Baker et al., 1987; Sherman and Alloul, 1992; Sherman et al., 1992), even though polycistronic mRNAs are inefficiently translated in eukaryotic cells (Kozak, 1987), it has been proposed that translation of the next ORFs downstream of the first HPV gene (E6), probably depends of processing of the polycistronic pre-mRNA by an intragenic splicing mechanism (Stacey et al., 1995). This is supported by the observation that splicing of the intron 1 of HPV-16 is necessary to allow efficiently translation for the second ORF that codify for the E7 protein (Zheng et al., 2004).

The intron 1 of HPV-16, is situated between the first two ORFs (E6 and E7). This intron is complex and has one donor site and three different acceptor sites. The splicing the intron 1, produces at least four polycistronic transcripts from these E6/E7 ORFs: the full unspliced E6/E7 pre-mRNA, two of these spliced transcripts encompassed the amino coding region of E6 gene and the complete E7 coding sequence named E6*I and E6*II RNAs (Baker and Calef, 1996) and the fourth, a spliced transcript named E6Ê7 (Ordoñes et al., 2004; Zhao et al., 2005). As the E6 and E7 oncogenes are expressed in all tumors and tumor-derived cell lines, this suggests an important role for these genes in host cell transformation (for reviews see zur Hausen, 2000). Therefore, production of all unspliced and spliced transcripts from this E6/E7 transcript is important for the process of cell transformation and/or tumorigenesis. The fully unspliced E6/E7 transcript is the source for translation of the E6 oncogene (Stacey et al., 1995), while two of the spliced transcripts (E6*I and E6*II) are the source of mRNA for the E7 oncogenic protein (Smotkin and Wettstein, 1986; Smotkin et al., 1989; Zheng et al., 2004). Finally, the third spliced transcript putatively codify for a new protein formed from the amino terminal end of the E6 protein, fused in frame to the carboxyl terminal end of the E7 protein (Ordoñes et al., 2004). Both the function and whether this protein is expressed remains unknown.

HPV-16 spliced transcript profiles are heterogeneously produced in tumor cells (Schwarz et al., 1985; Smotkin and Wettstein, 1986; Baker et al., 1987; Shirasawa et al., 1991; Sherman et al., 1992), probably because the three acceptor sites may be differential used in tumor cells and the difference between these spliced RNAs probably results in different levels of E6 and E7 oncogenes. Nevertheless, the mechanism underlying the recognition of these acceptor sites are yet not understood. We would like to understand the rules of usage of acceptor sites, particularly to identify all the sequences involved in the recognition and splicing of the intron 1. However, branch point sequences for the HPV-16 intron I have not yet identified. Thus, we used a very sensitive PCR technique (Vogel et al., 1997) to mapping and detected the probably branch point sequences present in this intron. Our results revealed a unusual non-consensus BPS located around 82 nt upstream of the first of the acceptor sites of intron 1 of HPV-16. Moreover, the branch point is at guanosine instead of the usual adenosine found for pre-mRNAs from other genes. The localization of this BPS and the unusual sequence suggest an explanation as to why it is difficulty to detect some of the early HPV-16 transcripts (Lehn et al., 1985; Sherman et al., 1992) and their splicing profile in tumor derived cell lines.

2. Materials and methods

2.1. Production of nucleic acids

An E6/E7 DNA fragment (nt 73–758) including the complete E6/E7 ORFs, was amplified from CaSki cells DNA (GenBank^{TM} accession no. U89348), using PCR and E6/E7 specific oligonucleotide primers (RHPV-16 and HPV-16-E7 end), subsequently this DNA fragment was cloned in the plasmid blue-script SKII$^+$ (Stratagene, USA) and named pHPV16-E6/E7. As this DNA fragment contains exon 1, intron 1 and exon 2 of HPV-16, other DNA fragments could be amplified by PCR using the appropriate primers. These were: to amplify the complete E6/E7 ORFs, the primer RHPV-16 (nt 73–95) 5′-CAGACATTTATGCACCAAAGA-3′ and the reverse primer HPV-16-E7 end (nt 834–858) 5′-TTATGGTTTCTGAGAACAGATGGG-3′ were used; to amplify the E6 fragment containing the first acceptor site, the above mentioned forward primer RHPV-16 and the reverse primer 3′AssHPV-16 (nt 440–463) 5′-GTCCAGATGTCTT-TGCTTTTCTTC-3′ were used. To amplify PCR DNA fragments suitable for in vitro transcription, an additional forward primer (RHPV-16 + T7, 5′-TAATACGACTCACTATAGG-CAGACATTTATGCACCAAAGA-3′) was used. This primer was longer than RHPV-16, and also contains the sequence of the T7 promoter. All synthetic RNAs were prepared using the T7 RNA polymerase and the mMessage mMachine^{TM} kit (Ambion Inc.).

2.2. Production of cDNA lariat intermediates, RT-PCR and sequencing

Lariat cDNAs were generated in vitro by splicing reactions using the HeLa nuclear protein extracts (HeLa NE) purchased from Promega Co. Briefly, splicing assays were performed in 50 μl of splicing buffer (Promega, RNA splicing system): 5 mM Hepes pH 7.9, 0.6% PVA, 20 mM Creatine Phosphate, 10% glycerol and 0.4 mM ATP 40% HeLa NE, and 1 μg of each synthetic E6/E7 RNA. Splicing reactions were incubated for 2 h at 30 °C and inactivated for 15 min incubation to 70 °C. Splicing reactions were then halted by phenol–chloroform extraction and ethanol precipitated at −20 °C. Precipitated splicing intermediates were subsequently dissolved in 20 μl of nuclease free water. The first strand cDNA synthesis was carried out for 2 h at 42 °C according to Vogel et al. (1997). Briefly, the total products of the splicing reaction were mixed with 10 μM antisense primer BPas (5′-CCCATCTCTATATACTATGCCATA-3′), and 200U of Superscript II RNAse H$^-$ reverse transcriptase (Gibco BRL) in 50 μl of reaction buffer, strictly following the manufacturer's protocol. Amplification of the 2′–5′ branched intermediates was carried out by PCR, involving a fiftieth of cDNA mixture as a template in