HHS Public Access

Author manuscript

*Virology*. Author manuscript; available in PMC 2017 January 15.

Published in final edited form as:

*Virology*. 2016 January 15; 488: 81–87. doi:10.1016/j.virol.2015.10.031.

# Alternative RNA splicing of KSHV ORF57 produces two different RNA isoforms

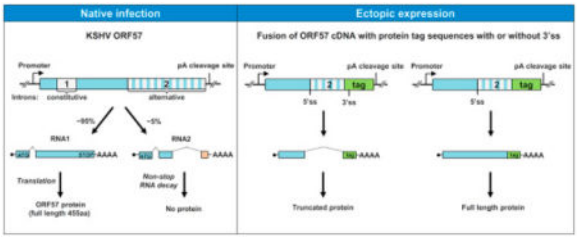
**Vladimir Majerciak** and **Zhi-Ming Zheng**

Tumor Virus RNA Biology Section, Gene Regulation and Chromosome Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, 21702 MD, USA

## Abstract

In lytically infected B cells Kaposi sarcoma-associated herpesvirus (KSHV) ORF57 gene encodes two RNA isoforms by alternative splicing of its pre-mRNA, which contains a small, constitutive intron in its 5′ half and a large, suboptimal intron in its 3's half. The RNA1 isoform encodes full-length ORF57 and is a major isoform derived from splicing of the constitutive small intron, but retaining the suboptimal large intron as the coding region. A small fraction (<5%) of ORF57 RNA undergoes double splicing to produce a smaller non-coding RNA2 due to lack of a translational termination codon. Both RNAs are cleaved and polyadenylated at the same cleavage site $CS_{83636}$. The insertion of ORF57 RNA1 into a restriction cutting site in certain mammalian expression vectors activates splicing of the subopitmal intron and produces a truncated ORF57 protein.

## Graphical Abstract




### Keywords

KSHV; ORF57; alternative splicing

## Introduction

Kaposi sarcoma-associated herpesvirus (KSHV, human herpesvirus 8 or HHV8) open reading frame (ORF) 57 is a posttranscriptional regulator of viral gene expression [for

Corresponding author: zhengt@exchange.nih.gov (Zhi-Ming Zheng).

review see (Majerciak & Zheng, 2015)] essential for virus gene expression and replication (Majerciak, Pripuzova et al., 2007;Han & Swaminathan, 2006;Verma, Li et al., 2015). A full-length ORF57 protein, naturally occurring as a homodimer (Majerciak, Pripuzova et al., 2015), is composed of 455 amino acid (aa) residues. In KSHV-infected cells, ORF57 is encoded by an abundant monocistronic transcript initiated from an inducible promoter $P_{82003}$ controlled by KSHV transactivator RTA (ORF50) (Lukac, Renne et al., 1998). The nascent transcripts are terminated at a polyadenylation cleavage site (CS) at nt 83636 in the virus genome, downstream of ORF57 coding region (Fig. 1A) (Majerciak, Yamanegi et al., 2006;Majerciak, Ni et al., 2013;Arias, Weisburd et al., 2014). The maturation of an ORF57 mRNA requires RNA splicing of a small intron (intron 1) in size of 108 nts from nt 82117 (5′donor splice site) to nt 82226 (3′ acceptor splice site) at the 5′ half of the pre-mRNA (Bello, Davison et al., 1999;Majerciak, Yamanegi et al., 2006).

A recent study by RNA-seq using epithelial iSLK cells harboring recombinant KSHV.219 (iSLK-219) suggests the possible presence of a novel, unrecognized 571-nt long intron (intron 2) in the 3′ half of ORF57 transcript (Arias, Weisburd et al., 2014) (Fig. 1A). The predicted double spliced ORF57 RNA was proposed to encode a novel, 299-aa long ORF57A protein which consists of the N-terminal 266-aa residues of ORF57 fused with 33-aa residues from a newly-created off-frame exon 3. While being extrapolated based on the observation of ribosome occupancy at the exon 3, the existence of ORF57A protein expression remained questionable. First, the expression of ORF57A would require the double spliced ORF57 RNA to read through the polyadenylation $CS_{83636}$ site to utilize a translation termination (STOP) codon further downstream (Fig. 1B). Second, the double spliced ORF57 RNA was detected at low frequency only in the iSLK-219 cells infected by a recombinant virus with the insertion of a GFP/RFP (green and red fluorescent proteins) reporter cassette downstream of ORF57 in its KSHV genome (Vieira & O'Hearn, 2004), but not in the cells infected by a native KSHV virus (Arias, Weisburd et al., 2014). Thus, this splicing was suspected as a result of undesirable activation of cryptic ORF57 RNA splicing by the insertion of a reporter GFP/RFP downstream of ORF57 gene.

## Results and Discussion

### ORF57 RNA undergoes alternative splicing during KSHV lytic infection in PEL cells

To investigate the possible presence of authentic double splicing of ORF57 during native KSHV infection, we performed RT-PCR analysis of ORF57 transcripts extracted from three primary effusion lymphoma (PEL) cell lines, including JSC-1, BCBL-1 and engineered BCBL1-TREx-RTA (TREx) (Renne, Zhong et al., 1996;Cannon, Ciufo et al., 2000;Nakamura, Lu et al., 2003), with undergoing KSHV lytic life cycle. To detect splicing of the putative intron 2, a forward primer (oVM96) located upstream of the putative ORF57 intron 2 was combined with a reverse primer (oNP18) immediately upstream of the $CS_{83636}$ site (for details see Fig. 1A and Table 1). As shown in Fig. 1C, we detected two distinct RT-PCR products from all three cell lines: a major product in size of ~932 nts with intron 2 retention representing RNA1 and a minor product in size of ~361 nts representing the RNA2 with intron 2 exclusion (lanes 1–6). No amplification was detected in the absence of reverse transcription (RT−), confirming RNA origin of the detected amplicons. The amount of the

RNA2 derived from double splicing varied among individual cells lines with JSC-1 cells showing the lowest level. DNA sequencing of the gel-purified RNA2 product revealed the exon 2-exon 3 junction at position nt 82974 to 83546 of KSHV genome (Fig. 1D), consistent with the proposed intron 2 splice sites identified in iSLK.219 cells (Arias, Weisburd et al., 2014). No other alternatively spliced isoform of ORF57 RNA was detected using others primer sets (data not shown). Thus, our data indicate the presence of an authentic RNA2 as a result of double splicing of ORF57 RNA in PEL cells with KSHV lytic infection.

To determine whether or not the RNA2 runs over the identified $CS_{83636}$ site to reach a translation termination codon, we performed an RT-PCR using the same forward primer (oVM96) in combination with a reverse primer (oVM280) immediately downstream of the $CS_{83636}$ site corresponding to the Rev-C primer used in Arias's study (Arias, Weisburd et al., 2014). Using this set of primers we also detected RNA1 from all three cell lines with lytic KSHV infection, with TREx BCBL1-RTA cells being the most (Fig. 1C, lanes 7–12). However, the amounts of RNA1 products amplified from this primer set (Fig. 1C, lanes 7–12) were substantially lower than the amount amplified from the primer set (oVM96+oNP18) with a reverse primer oNP18 immediately upstream of the $CS_{83636}$ site (see Fig. 1C, lanes 1–6). Using this primer set the RNA2 was not detectable, but could be amplified with more cycles (data not shown). In contrary to Arias et al., we were also able to detect other longer RNA products by using a reverse primer (Rev-D) further downstream from K9 coding region (data not shown). Together, our data suggest that the majority of ORF57 RNA isoforms are cleaved at the $CS_{83636}$ site for polyadenylation and only a minimal fraction in BCBL-1 and JSC-1 cells could run through the $CS_{83636}$ site. Why the ORF57 RNA1 could run through the $CS_{83636}$ site more often in TREx BCBL-1-RTA cells than its parent BCBL-1 cells remains to be further investigated.

## Both ORF57 RNA1 and RNA2 are polyadenylated from the $CS_{83636}$ site

To examine whether another possible alternative polyadenylation site is used by a small fraction of the ORF57 RNA isoforms running over the $CS_{83636}$ site, we carried out nested 3′RACE analysis on total RNA from TREx BCBL1-RTA cells (Nakamura, Lu et al., 2003) with high level of KSHV lytic induction. For this purpose, two forward primers (oVM96/oVM12) located upstream of the intron 2 5′ splice site (Fig. 1A) were chosen for this study. The primers (oVM67/oVM36, Fig. 1A) within the intron 2 previously used for ORF57 cleavage site mapping served as a positive control (Majerciak, Yamanegi et al., 2006). In accordance with our previous report, we found that the intron 2-specific primers amplified only a single 3′ RACE product corresponding to ORF57 RNA1 (Fig. 1E, lane 1). However, the primer upstream the intron 2 led to amplify two 3′ RACE products with the size corresponding to the RNA1 and RNA2, respectively (Fig. 1E, lane 2). DNA sequencing of all three 3′RACE products (Fig. 1E, lanes 1 and 2) revealed that they are all cleaved at a single cleavage site, $CS_{83636}$ site (Majerciak, Ni et al., 2013)(Majerciak, Yamanegi et al., 2006). We did not identify any other alternative polyadenylation site for ORF57 RNA polyadenylation. Thus, the $CS_{83636}$ site identified in our previous report (Majerciak, Ni et al., 2013)(Majerciak, Yamanegi et al., 2006) is the only polyadenylation site for all isoforms of ORF57 transcripts during KSHV lytic infection. We conclude that the transcripts running over the $CS_{83636}$ are rare and if any, might be polyadenylated by using a non-canonical pA

signal or cleavage site and even so, might be accompanied by RNA-degradation pathways (Wyers, Rougemaille et al., 2005). Lack of defined polyadenylation sites downstream of the reported $CS_{83636}$ site was further supported by the observation that the level of the amplified ORF57 transcripts was gradually decreased along with increased distance of the reverse primers from the $CS_{83636}$ site (data not shown). This is consistent with the proposed torpedo model for termination of RNA polymerase II transcription (Logan, Falck-Pedersen et al., 1987;Connelly & Manley, 1988). Finally, the usage of the reported $CS_{83636}$ site for polyadenylation of ORF57 RNA2 indicates that intron 2 splicing does not affect the $CS_{83636}$ selection as observed in other mammalian transcripts (Tian, Pan et al., 2007).

## Double-spliced ORF57 RNA2 represents a minor form of ORF57 transcripts

To determine the exact level of ORF57 intron 2 splicing during lytic KSHV infection, we next examined the expression level of ORF57 RNA2 in KSHV-infected cells. We re-analyzed previously published mapping of ORF57 transcripts in JSC-1 cells (Majerciak, Yamanegi et al., 2006) using a $^{32}$P-labeled antisense RNA-probe (see Fig. 1A) spanning over the intron 2 and exon 3 regions from the ORF57 RNA 3′end by other two common methods. First, by Northern blotting with adjusted exposure, we clearly detected two isoforms of ORF57: a previously reported, abundant RNA1 of ~1.5 kb in size and the newly identified RNA2 with predicted size of ~1 kb (Fig. 1F). Both isoforms were present only in samples with lytic, but not latent infection (Fig. 1F, compare lanes 2 and 4 to lanes 1 and 3) and were enriched in polyadenylated RNA fraction (lane 4). This is in accordance with ORF57 RNA being a lytic and polyadenylated transcript (Bello, Davison et al., 1999;Majerciak, Yamanegi et al., 2006;Majerciak & Zheng, 2009). Based on signal intensity of each ORF57 mRNA isoform, we found that the RNA2 accounts for only <5%, whereas the RNA1 accounts for ~95% of the polyadenylated, total ORF57 RNA. The low abundance of the RNA2 was also confirmed by RNase–protection assays using the same 3′ end probe on the same RNA from the activated JSC-1 cells with KSHV lytic infection. As shown in Fig. 1G, we again detected only a trace amount of the RNA2, with the majority of protected probe hybridized to the RNA1 (lane 2). The presence of two protected fragments for both RNA1 and RNA2 represents the heterogeneity in selection of the polyadenylation cleavage site $CS_{83636}$ due to its suboptimal nature of a downstream GU-rich region (Majerciak, Yamanegi et al., 2006). Consistent with our 3′RACE results in Fig. 1E, RPA did not detect any notable ORF57 transcripts being polyadenylated at another cleavage site further downstream.

Nevertheless, our data show that native KSHV infection does produce a minor, double spliced ORF57 RNA2 that is polyadenylated at the $CS_{83636}$. As a result, the RNA2 would be non-coding due to lack of a functional translational stop codon and thus would subject to non-stop RNA decay [for review see (Lykke-Andersen & Bennett, 2014)]. Though a small, undetectable fraction of the RNA2 might escape the $CS_{83636}$ to use another non-canonical cleavage site further downstream for its polyadenylation, the existence of a possible ORF57A protein translated from such a rare RNA2 is negligible and its biological significance would require further experimental validation.

Homologous EB2 from Epstein-Barr virus shares the similar gene structure with KSHV ORF57. EB2 protein is also expressed from a monocistronic mRNA generated by splicing of a small constitutive intron at the 5′-half of nascent EB2 transcript (Batisse, Manet et al., 2005). Despite that bioinformatics analysis of the EB2 sequence could predict a potential alternative intron at the EB2 3′-half, we were unable to detect, by RT-PCR, any double spliced EB2 RNA from activated JSC-1 cells co-infected with KSHV and EBV or Raji cells infected with EBV only (data not shown).

## Selective activation of ORF57 alternative splicing from mammalian expression vectors

Given that the intron 2 3′ splice site of ORF57 RNA is out of frame by one nt after the ORF57 stop codon UAA, it is generally removed by cloning the coding region of ORF57 RNA1 into a mammalian expression vector (Fig. 2A). However, by RT-PCR we unexpectedly detected the spliced ORF57 RNA species expressed from some of several expression vectors in HEK293 cells using a pair of primers, one upstream of the ORF57 intron 2 5′ splice site and one downstream of the protein tag coding region (Fig. 2A and 2B). There was no detectable RNA splicing from untagged ORF57 or ORF57 with a C-terminal FLAG tag (Fig. 2B, lanes 2 and 6). A high level of RNA splicing was seen in the ORF57 fused with a C-terminal V5-His tag, c-myc tag, or HA tag (Fig. 2B, lanes 4, 8, and 10). As a result, the latter three vectors failed to efficiently express a full-length ORF57 protein, but instead express mostly a truncated form of ORF57 protein (Fig. 2C, lanes 3–5). Analysis of exon-exon junctions in the spliced RNA from those vectors showed that the splicing took place from the intron 2 5′ splice site of ORF57 to a new 3′ splice site within the vector multiple clone site sequences (Fig. 3A). The vectors expressing no detectable ORF57 splicing (untagged ORF57 and ORF57-FLAG) lack a notable 3′ splice site at this location (Fig. 2B, lanes 2 and 6). However, in the case of ORF57-GFP vector where GFP was fused to the ORF57 C-terminus (Fig. 2D) via a *Bam* HI site in a pEGFP-N1 vector, this clone site in pEGFP-N1 does not bear a 3′ splice site, but evident splicing of ORF57-GFP RNA was identified in HEK293 cells (Fig. 2D). By sequencing of the spliced product, we found this fusion transcript was spliced from the intron 2 5′splice site of ORF57 to a cryptic 3′ splice site in the GFP coding region (Fig. 2D and 2E, Fig. 3B). This splicing happens in-frame and leads to express a chimeric, truncated ORF57-GFP protein which contains the N-terminal 266-aa residues of ORF57 and the C-terminal 235-aa residues of GFP. Thus, the chimeric, truncated ORF57 can be recognized both by an anti-GFP antibody and an anti-ORF57 N-terminal antibody recognizing the epitope upstream of the intron 2, but not by an antibody against the ORF57 C-terminal epitope resided in the intron 2 (Fig. 2D and 2F).

## Splicing efficiency of the ORF57 intron 2 is regulated by the strength of 3′ splice site

Most of eukaryotic introns are GU-AG introns, with GU dinucleotide being a 5′ (donor) splice site and AG dinucleotide being a 3′ (acceptor) splice site. Immediately upstream of the 3′ splice sites are featured by a branch point sequence and a polypyrimidine track which are recognized first by cellular splicing machinery to facilitate intron excision (splicing) from a nascent transcript (Zheng, 2004). The first intron in the 5′ half of ORF57 RNA is a constitutive intron, with a small (114 nts), capped exon immediately upstream and thereby, its splicing can be efficiently promoted by the capping machinery cross over the exon (Konarska, Padgett et al., 1984;Lewis, Izaurralde et al., 1996;Zheng, Tao et al., 2004). In

addition to the suboptimal feature of the intron 2, a large size (749 nts) of the internal exon 2 upstream of the intron 2 may also prevent recognition of the intron 2 5′ splice site by the cellular splicing machinery It has been well documented that the size of an internal exon affects cross-talk (exon definition) of a 3′ splice site and a 5′ splice site over the exon and this cross-talk is limited by an oversized exon (>500 nts) (Robberson, Cote et al., 1990;Sterner, Carlo et al., 1996;Zheng, 2004;Ajiro, Jia et al., 2012). Moreover, the intron 2 3′ splice site AG in ORF57 is preceded by unfavorable A and may also contribute to a low level of intron 2 splicing during native KSHV infection (Fig. 3A) since a nucleotide preceding the AG has striking influence on selection of a 3′ splice site. The order of competitiveness in the splicing scanning model has a nucleotide preference of CAG congruent to UAG > AAG > GAG at this position (Smith, Chu et al., 1993). By insertion of ORF57 ORF into an *Xba*I site of a pFLAG-myc-CMV-22 vector or a pcDNA3.0-HA vector, the strength of 3′ splice site is increased from AAG to UAG, leading to efficient splicing of ORF57 RNA (Fig. 3A). While insertion of genomic ORF57 region into a pcDNA3.1/V5-His TOPO vector remains the 3′ splice site AAG unchanged (Fig. 3A), but the majority of the RNA was spliced even though a small fraction of the full-length ORF57 protein can be expressed (Fig. 2C),. One interpretation is that a U upstream of the AAG in the ORF57 intron 2 3′ splice site was missing from the inserted ORF57 intron 2 in the TOPO vector and the nucleotide U is suppressive to RNA splicing (Tanaka, Watakabe et al., 1994)(Zheng, He et al., 1997). In addition, other splicing enhancer sequence(s) from the TOPO vector is presumably also involved in the enhanced splicing. In the case of the pFLAG-CMV-5.1 or pEGFP-N1 vector, insertion of ORF57 ORF into its *BamH*I site disrupts the intron 2 3′splice site. Although a full-length ORF57-FLAG is expressed from the pFLAG-CMV-5.1 vector, insertion of ORF57 ORF into a pEGFP-N1 vector activates a cryptic 3′ splice site in the 5′ half of GFP ORF region downstream (Fig. 3B) and thereby leads to partial production of truncated ORF57-GFP fusion protein. Altogether, our data indicate that the accidental creation of a potential splice site should be avoided in designing ORF57 expression in a mammalian expression system.

## Materials and methods

### Cell lines, transfection and virus induction

KSHV-transformed primary effusion lymphoma (PEL) cell lines JSC-1 (Cannon, Ciufo et al., 2000), BCBL-1 (Renne, Zhong et al., 1996), and BCBL1-TREx-RTA (Nakamura, Lu et al., 2003) were cultivated in RPMI1640 medium supplemented with 10% fetal bovine serum (FBS, GE Healthcare) and 1× Penicillin-Streptomycin-Glutamine (PSG, Thermo Scientific). KSHV lytic replication was induced by treatment of JSC-1 cells with 3 mM sodium butyrate, BCBL-1 cells with 1 mM sodium valproate, or BCBL1-TREx-RTA cells with 1μg/μl of doxycycline, for period 24–48 h. HEK293 cells were grown in complete DMEM medium with 10% FBS and 1× PSG. All transfections were performed by LipoD293™ DNA In Vitro Transfection Reagent (SignaGen Laboratories) as recommended.

### Plasmids

All ORF57 expression vectors were generated by cloning of ORF57 genomic or cDNA fragments into various eukaryotic expression vectors. The vector pVM1 expressing C-

terminally V5-His-tagged ORF57 was generated by TOPO cloning of genomic region of ORF57 into pcDNA3.1D/V5-His-TOPO vector (Invitrogen). The pcDNA3.0 (Invitrogen) containing ORF57 cDNA was used for the expression of untagged ORF57, pFLAG-CMV-5.1 (Sigma) for the expression of C-terminal FLAG-tagged ORF57 (pVM7), pFLAG-myc-CMV-22 for the expression of both N-terminal FLAG- and C-terminal myc-tagged ORF57 (pVM78), pcDNA3-HA for the expression of C-terminal HA-tagged ORF57 (pVM79), and pEGFP-N1 (Clontech) for the expression of ORF57-GFP fusion protein.

### Two-step RT-PCR

Total RNA prepared by TRIpure reagent (Roche) was first treated with TURBO DNA-free Kit (Ambion) to remove residual contaminant DNA before RT-PCR. RT was carried out on 1μg of total RNA using random hexamers with MuLV Reverse Transcriptase (Applied Biosystems) at 42° C for 1 h followed by PCR with a pair of gene-specific primers using GeneAmp RNA PCR Amplification kit (Applied Biosystems). The sequences and position of primers used in the study are summarized in Table 1.

### Northern blot

Poly A-selected RNA was prepared form total RNA using illustra QuickPrep mRNA Purification Kit (GE Healthcare). Total and polyA-selected RNA were separated on formaldehyde containing 1.2% agarose gel and transfer to a nylon membrane. KSHV ORF57 transcripts were detected by a $^{32}$P-labeled RNA antisense probe prepared by in vitro transcription on PCR DNA template using Riboprobe System-T7 (Promega).

### RNase-protection assay (RPA)

RNase-protection assay was carried out using RPAIII kit (Ambion). Total RNA was co-precipitated with a $^{32}$P-labeled RNA antisense probe. The pelleted RNA was dissolved in 10 μl on hybridization buffer and incubated overnight at 42° C. The unprotected RNA probe by hybridization was removed by RNase A/T1 digestion for 1 h at 37° C. The protected RNA probe during hybridization was separated on a denaturing 8% urea PAGE gel, with $^{32}$P-labeled 100bp DNA Ladders being size markers.

### 3′RACE

RNA polyadenylation sites were determined using SMARTer RACE 5′/3′ Kit (Clontech) using 1μg of total RNA as recommended by manufacturer. The RACE products from the first amplification were diluted 1:50 and used as template for the second amplification with nested primers. The resulted RACE products were separated on an agarose gel and purified for direct Sanger sequencing.

### Western blotting

Total protein extract was prepared by direct cell lysis in Laemmli SDS protein sample buffer supplemented with 5% (vol/vol) of 2-Mercaptoethanol and was separated on a gradient SDS-PAGE gel. After transfer to a nitrocellulose membrane, ORF57 and its fusions were blotted with an anti-ORF57 N- or C-terminal antibody (Majerciak, Kruhlak et al., 2010) (Majerciak, Pripuzova et al., 2015) or an anti-GFP (Clontech) antibody.

## Acknowledgments

This work was supported by the intramural research programs of the Center for Cancer Research, National Cancer Institute, National Institutes of Health.

## References

Ajiro M, Jia R, Zhang L, Liu X, Zheng ZM. Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*I and E7 expression. PLoS ONE. 2012; 7:e46412. [PubMed: 23056301]

Arias C, Weisburd B, Stern-Ginossar N, Mercier A, Madrid AS, Bellare P, Holdorf M, Weissman JS, Ganem D. KSHV 2.0: a comprehensive annotation of the Kaposi's sarcoma-associated herpesvirus genome using next-generation sequencing reveals novel genomic and functional features. PLoS Pathog. 2014; 10:e1003847. [PubMed: 24453964]

Batisse J, Manet E, Middeldorp J, Sergeant A, Gruffat H. Epstein-Barr virus mRNA export factor EB2 is essential for intranuclear capsid assembly and production of gp350. J Virol. 2005; 79:14102–14111. [PubMed: 16254345]

Bello LJ, Davison AJ, Glenn MA, Whitehouse A, Rethmeier N, Schulz TF, Barklie CJ. The human herpesvirus-8 ORF 57 gene and its properties. J Gen Virol. 1999; 80( Pt 12):3207–3215. [PubMed: 10567653]

Cannon JS, Ciufo D, Hawkins AL, Griffin CA, Borowitz MJ, Hayward GS, Ambinder RF. A new primary effusion lymphoma-derived cell line yields a highly infectious Kaposi's sarcoma herpesvirus-containing supernatant. Journal of Virology. 2000; 74:10187–10193. [PubMed: 11024147]

Connelly S, Manley JL. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. 1988; 2:440–452. [PubMed: 2836265]

Han Z, Swaminathan S. Kaposi's sarcoma-associated herpesvirus lytic gene ORF57 is essential for infectious virion production. J Virol. 2006; 80:5251–5260. [PubMed: 16699005]

Konarska MM, Padgett RA, Sharp PA. Recognition of cap structure in splicing in vitro of mRNA precursors. Cell. 1984; 38:731–736. [PubMed: 6567484]

Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj IW. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5′ splice site. Genes Dev. 1996; 10:1683–1698. [PubMed: 8682298]

Logan J, Falck-Pedersen E, Darnell JE Jr, Shenk T. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. Proc Natl Acad Sci USA. 1987; 84:8306–8310. [PubMed: 3479794]

Lukac DM, Renne R, Kirshner JR, Ganem D. Reactivation of Kaposi's sarcoma-associated herpesvirus infection from latency by expression of the ORF 50 transactivator, a homolog of the EBV R protein. Virology. 1998; 252:304–312. [PubMed: 9878608]

Lykke-Andersen J, Bennett EJ. Protecting the proteome: Eukaryotic cotranslational quality control pathways. J Cell Biol. 2014; 204:467–476. [PubMed: 24535822]

Majerciak V, Kruhlak M, Dagur PK, McCoy JP Jr, Zheng ZM. Caspase-7 cleavage of Kaposi sarcoma-associated herpesvirus ORF57 confers a cellular function against viral lytic gene expression. J Biol Chem. 2010; 285:11297–11307. [PubMed: 20159985]

Majerciak V, Ni T, Yang W, Meng B, Zhu J, Zheng ZM. A viral genome landscape of RNA polyadenylation from KSHV latent to lytic infection. PLoS Pathog. 2013; 9:e1003749. [PubMed: 24244170]

Majerciak V, Pripuzova N, Chan C, Temkin N, Specht SI, Zheng ZM. Stability of structured Kaposi's sarcoma-associated herpesvirus ORF57 protein is regulated by protein phosphorylation and homodimerization. J Virol. 2015; 89:3256–3274. [PubMed: 25568207]

Majerciak V, Pripuzova N, McCoy JP, Gao SJ, Zheng ZM. Targeted disruption of Kaposi's sarcoma-associated herpesvirus ORF57 in the viral genome is detrimental for the expression of ORF59, K8alpha, and K8.1 and the production of infectious virus. J Virol. 2007; 81:1062–1071. [PubMed: 17108026]

Majerciak V, Yamanegi K, Zheng ZM. Gene structure and expression of Kaposi's sarcoma-associated herpesvirus ORF56, ORF57, ORF58, and ORF59. J Virol. 2006; 80:11968–11981. [PubMed: 17020939]

Majerciak V, Zheng ZM. Kaposi's sarcoma-associated herpesvirus ORF57 in viral RNA processing. Front Biosci. 2009; 14:1516–1528.

Majerciak V, Zheng ZM. KSHV ORF57, a Protein of Many Faces. Viruses. 2015; 7:604–633. [PubMed: 25674768]

Nakamura H, Lu M, Gwack Y, Souvlis J, Zeichner SL, Jung JU. Global changes in Kaposi's sarcoma-associated virus gene expression patterns following expression of a tetracycline-inducible Rta transactivator. J Virol. 2003; 77:4205–4220. [PubMed: 12634378]

Renne R, Zhong W, Herndier B, McGrath M, Abbey N, Kedes D, Ganem D. Lytic growth of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in culture. Nat Med. 1996; 2:342–346. [PubMed: 8612236]

Robberson BL, Cote GJ, Berget SM. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol Cell Biol. 1990; 10:84–94. [PubMed: 2136768]

Smith CW, Chu TT, Nadal-Ginard B. Scanning and competition between AGs are involved in 3′ splice site selection in mammalian introns. Mol Cell Biol. 1993; 13:4939–4952. [PubMed: 8336728]

Sterner DA, Carlo T, Berget SM. Architectural limits on split genes. Proc Natl Acad Sci USA. 1996; 93:15081–15085. [PubMed: 8986767]

Tanaka K, Watakabe A, Shimura Y. Polypurine sequences within a downstream exon function as a splicing enhancer. Mol Cell Biol. 1994; 14:1347–1354. [PubMed: 8289812]

Tian B, Pan Z, Lee JY. Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. Genome Res. 2007; 17:156–165. [PubMed: 17210931]

Verma D, Li DJ, Krueger B, Renne R, Swaminathan S. Identification of the Physiological Gene Targets of the Essential Lytic Replicative Kaposi's Sarcoma-Associated Herpesvirus ORF57 Protein. J Virol. 2015; 89:1688–1702. [PubMed: 25410858]

Vieira J, O'Hearn PM. Use of the red fluorescent protein as a marker of Kaposi's sarcoma-associated herpesvirus lytic gene expression. Virology. 2004; 325:225–240. [PubMed: 15246263]

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell. 2005; 121:725–737. [PubMed: 15935759]

Zheng ZM. Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. J Biomed Sci. 2004; 11:278–294. [PubMed: 15067211]

Zheng ZM, He PJ, Baker CC. Structural, functional, and protein binding analyses of bovine papillomavirus type 1 exonic splicing enhancers. J Virol. 1997; 71:9096–9107. [PubMed: 9371566]

Zheng ZM, Tao M, Yamanegi K, Bodaghi S, Xiao W. Splicing of a cap-proximal human Papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5′ cap. J Mol Biol. 2004; 337:1091–1108. [PubMed: 15046980]

## Highlights

- A native KSHV ORF57 transcript has a small constitutive intron and a large suboptimal intron
- RNA splicing to include or exclude the suboptimal intron results in two ORF57 isoforms
- The RNA1, a major isoform derived from splicing of the constitutive intron, encodes full-length ORF57 protein
- The RNA2 derived from double splicing is a non-coding RNA
- The suboptimal intron in RNA1 can be activated to splice in certain expression vectors

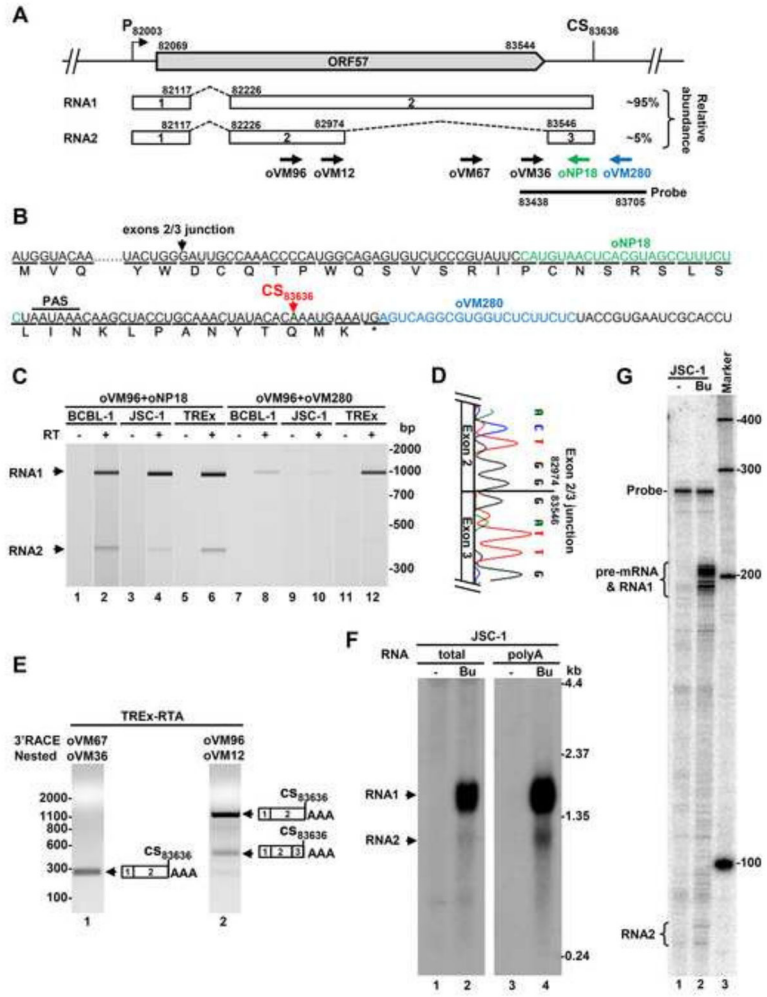



**Fig. 1.**

KSHV ORF57 gene encodes two isoforms of alternatively spliced RNA. (A) Diagram of the KSHV ORF57 locus in the virus genome showing the ORF57 coding region (big grey arrow), the transcriptional starting site (P) at nt 82003 and the polyadenylation cleavage site (CS) at nt 83636. The alternatively spliced RNA transcripts (RNA1 and RNA2) are shown below. The numbered boxes represent exons separated by introns (dashed lines). On the right is the expression level of individual ORF57 transcripts determined from the Northern blots shown in Fig. 1F determined by ImageJ analysis (http://imagej.nih.gov/ij/). Arrows below the RNA2 are primers used in this study, with a thick line for an antisense RNA-probe used for Northern blot and RPA. The nucleotides positions are based on the sequence of the KSHV genome (GenBank Acc. No. U75698.1). (B) Amino acid composition of the hypothetical ORF57A encoded by RNA2 derived from double splicing. Arrows mark the positions of the exon-exon junction and ORF57 polyadenylation cleavage site (CS), respectively. The classical polyadenylation signal (PAS) AAUAAA of ORF57 and reverse primers (colored) upstream or downstream of the $CS_{83636}$ site are also indicated in this panel. (C) Alternative splicing of ORF57 exon 2 during KSHV lytic infection in three KSHV-infected primary effusion lymphoma cell lines. The viral lytic cycle was induced by

48 h treatment with 1 mM valproic acid (BCBL-1), 3 mM sodium butyrate (JSC-1) or 1 μg/ml of doxycycline (TREx BCBL1-RTA). Two-step RT-PCR was carried out on total RNA by reverse transcription (RT+) with random hexamers followed by 25 cycles of PCR amplification with the indicated gene-specific primers. The size and concentration of obtained PCR products were determined by Bioanalyzer 2100 (Agilent). The reactions without RT (−) were used as a negative control. (D) A chromatograph shows the exon-exon junction of ORF57 intron 2 splicing determined by Sanger sequencing of spliced RT-PCR products. (E) Polyadenylation of ORF57 transcripts at the $CS_{83636}$. The nested 3′RACE was carried out on total RNA from lytic TREx BCBL1-RTA cells. The primers oVM96 and oVM67 were used for the first and oVM12 and oVM36 for the second (nested) amplification. The resulting products were analyzed by agarose gel electrophoresis. The polyadenylation cleavage site was determined by Sanger DNA sequencing of each nested 3′RACE product. (F and G) Expression profile of ORF57 transcripts in KSHV-infected cells. Total or poly A-selected RNA from JSC-1 cells with latent (−) and lytic (Bu-butyrate treated) infection used for ORF57 Northern blot in our previous publication (Majerciak, Yamanegi et al., 2006) was re-probed by an antisense RNA-probe as described in the panel A (F). Total cell RNA from JSC-1 cells with or without butyrate treatment for lytic infection was also examined by RNase-protection assay (RPA) by the same antisense RNA probe described in the panel A (G).

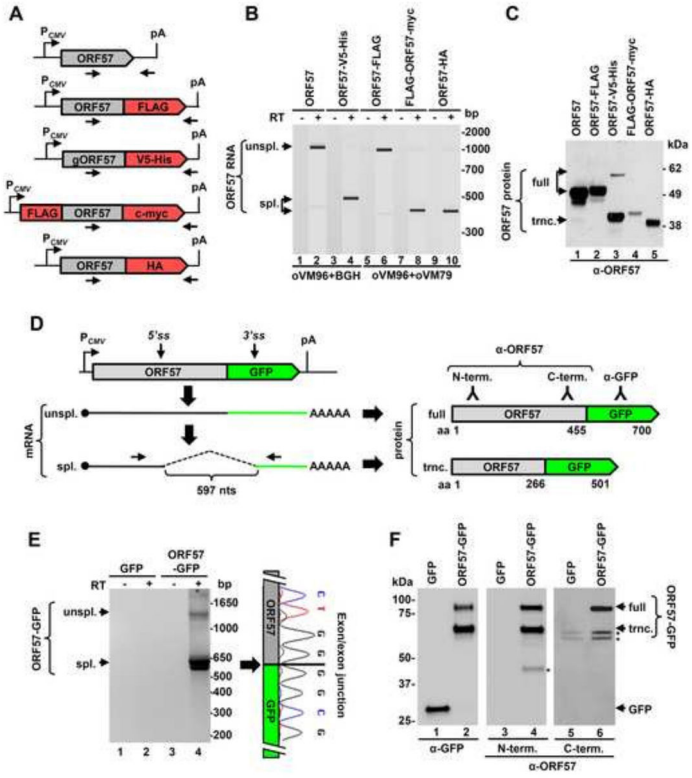



**Fig. 2.**

Activation of an alternative RNA splicing during ORF57 ectopic expression. (A) Diagrams of various ORF57 expression vectors generated by cloning the ORF57 (grey) coding region with (gORF57) or without intron 1 into various eukaryotic expression vectors (for details see Materials and Methods). The protein tags fused to ORF57 proteins are shown in red. The small arrows below each diagram are the primers used in RT-PCR. $P_{CMV,}$ cytomegalovirus early promoter; pA, polyadenylation site. (B) Detection of ORF57 transcripts expressed from various expression vectors shown in the panel (A). RT-PCR was performed on total RNA isolated form HEK293 cells 24 h after plasmid transfection with the primer sets (Table 1) indicated below. The reactions without RT (−) were used as a negative control. The resulting products were analyzed by Bioanalyzer 2100. (C) Expression of truncated (trnc) ORF57 protein from spliced ORF57 RNA. Western blot analysis of ORF57 protein expression from the indicated vector in HEK293 cells was performed 24 h after transfection. The protein samples were analyzed by SDS-PAGE and blotted by an anti-ORF57 N-terminal antibody. (D) Diagrams of the chimeric ORF57-GFP fusion expression, RNA splicing, and the resulted protein isoforms. ORF57 cDNA was inserted into pEGFP-N1 at *Bam*HI site to express an ORF57-GFP fusion. Unspliced and spliced chimeric RNAs were detected by RT-PCR with the indicated primer set (small arrows). Dashed lines are the activated cryptic intron in size of 597 nts in this expression vector. On the panel right are the fused proteins expressed from individual transcripts, with its length in amino acid residues shown below. Shown above the protein diagrams are the relative epitope positions for anti-ORF57 (N- or C-terminal) or anti-GFP antibody recognition. (E–F) Alternative splicing of chimeric ORF57-GFP RNA and protein expression in HEK293 cells. The cells transfected

with a GFP or chimeric ORF57-GFP expression vector were harvested 24 h after transfection for analysis of total cell RNA (E) and proteins (F). Splicing of the chimeric ORF57-GFP RNA was detected from total cell RNA by RT-PCR using a primer pair of oVM96 and oZMZ296 and the exon-exon junction was determined by DNA sequencing (E). Western blot of total cellular extracts was performed by an anti-GFP, anti-ORF57 N-terminal, or anti-ORF57 C-terminal antibody (F). *, unknown proteins.

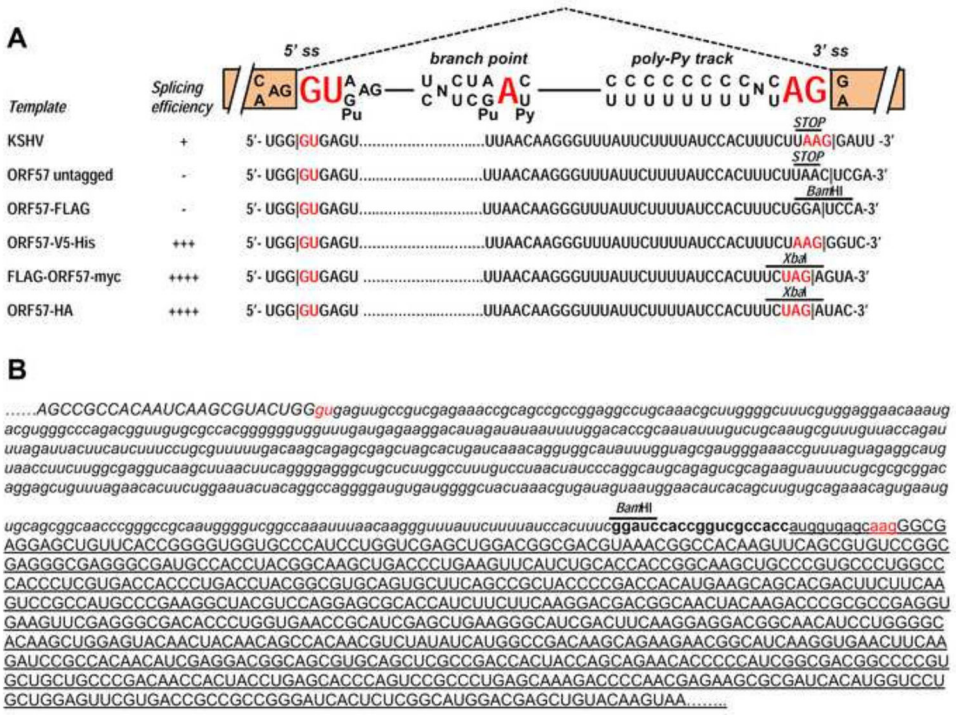



**Fig. 3.**

Sequence structure of the ORF57 intron 2 and its adjacent exons. (A) Diagram of an eukaryotic GU-AG intron and its important features (5′ splice site [5′ ss], branch point sequence, polypyrimidine track, and 3′ splice site [3′ ss]). Below the diagram are the ORF57 suboptimal intron 2 and its native 3′ splice site or an artificially created 3′ splice site derived from the cloning site of individual expression vectors. Stop indicates a stop codon UAA in the native ORF57 ORF. In addition to the 5′ splice site GU, the 3′ splice site AG, and the branch point A in red color, a nucleotide preceding the AG dinucleotide in a 3′ splice site is also colored in red to show its importance in selection of a 3′ splice site. (B) Partial sequence view of the chimeric ORF57-GFP fusion region, with exon sequences in uppercase, intron sequences in lowercase, ORF57 in italic, and vector in bold letters. The GFP sequences are underlined. The 5′ splice site GU and 3′ splice site AG and its preceding A are red colored. A *Bam*HI site used for insertion of ORF57 ORF into the pEGFP-N1 vector is shown.

## TABLE 1

Primers used in this study.

| Name | Position | Sequence | Assay |
|---|---|---|---|
| *oVM12* | 82699-82718 | 5′-CTCAGACTCCCTGCGAGCAT-3′ | 3′RACE |
| *oVM36* | 83438-83454 | 5′-TGGAACATCACAGCTTG-3′ | 3′RACE, RPA[a] |
| *oVM67* | 83254-83274 | 5′-GCATGTAACCTTCTTGGCGAG-3′ | 3′RACE |
| *oVM90* | 83705-83687 | 5′-taatacgactcactatag/GGCGGTCTGGTGTTGTGTT-3′ | RPA[a] |
| *oVM96* | 82675-82695 | 5′-tactcagaattcaccatg/GGTGTGTCTGACGCCGTAAAG-3′ | RT-PCR, 3′RACE |
| *oNP18* | 83606-83583 | 5′-atcgtgttaac/GAGAAAGGCTACGTGAGTTACATG-3′ | RT-PCR |
| *oVM280* | 83667-83647 | 5′-GAGAAGAGACCACGCCTGACT-3′ | RT-PCR[b] |
| *oZMZ296* | N/A | 5′-gcatggcggacttgaagaa-3′ (vector) | RT-PCR |
| *oVM79* | N/A | 5′-gggcactggagtggcaac-3′ (vector) | RT-PCR |
| *BGH* | N/A | 5′-tagaaggcacagtcgagg-3′ (vector) | RT-PCR |

Primer positions are based on KSHV genome sequence GenBank Acc. No. U75698.1. The lower case marks non-KSHV sequences.

[a] primers used to prepare DNA template for RNA probe by in vitro transcription, with oVM90 attached by a T7 promoter sequence;

[b] primer identical to Rev-C from (Arias, Weisburd et al., 2014). BGH, bovine growth hormone sequence from the vector.