

## Transcriptional Origin of Kaposi's Sarcoma-Associated Herpesvirus MicroRNAs

Xuezhong Cai and Bryan R. Cullen\*

*Center for Virology and Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina 27710*

Received 27 September 2005/Accepted 5 December 2005

**Kaposi's sarcoma-associated herpesvirus (KSHV) encodes 11 distinct microRNAs, all of which are found clustered within the major latency-associated region of the KSHV genome in the same transcriptional orientation. Because the KSHV microRNAs are all expressed in latently infected cells and are largely unaffected by induction of lytic replication, it appeared probable that they would be processed out of KSHV transcripts that are derived from a latent promoter(s) present in this region. Here, we define three latent transcripts, derived from two distinct KSHV latent promoters, that function as both KSHV primary microRNA precursors and as kaposin pre-mRNAs. These activities require the readthrough of a leaky viral polyadenylation signal located at nucleotide 122070 in the KSHV genome. In contrast, recognition of this polyadenylation signal gives rise to previously identified mRNAs that encode the KSHV open reading frames (ORFs) 71, 72 and 73 proteins as well as a novel unspliced KSHV mRNA that encodes only ORF72 and ORF71. Thus, transcripts initiating at the two latent promoters present in the KSHV latency-associated region can undergo two entirely distinct fates, i.e., processing to give a kaposin mRNA and viral microRNAs on the one hand or expression as KSHV ORF71, ORF72, or ORF73 mRNAs on the other, depending on whether the viral polyadenylation site located at position 122070 is ignored or recognized, respectively.**

MicroRNAs (miRNAs) are a class of evolutionarily conserved, ~22-nucleotide-long noncoding RNAs that are expressed in all metazoan eukaryotes examined so far (1). Recently, several pathogenic DNA viruses have also been shown to encode miRNAs, including members of the herpesvirus family (4, 21, 22, 24, 28). In particular, Kaposi's sarcoma-associated herpesvirus (KSHV) encodes at least 11 distinct miRNAs, termed miR-K1 to miR-K11, all of which are expressed in latently KSHV-infected human cells.

In almost all cases, miRNAs are initially transcribed as part of a long capped, polyadenylated transcript, generated by RNA polymerase II, termed a primary miRNA (3, 15). The mature miRNA forms part of the stem of an ~80-nucleotide RNA hairpin within the primary miRNA precursor and one primary miRNA may contain a cluster of several different miRNAs (6).

The first step in miRNA processing is the nuclear excision of the upper ~60 nucleotides of this stem-loop by a heterodimer consisting of the RNase III enzyme Droscha and its cofactor DGCR8 (7, 10, 13). The resultant pre-miRNA hairpin is exported to the cytoplasm by the nuclear export factor exportin 5 (17, 30), where it interacts with a second RNase III enzyme termed Dicer. Dicer, acting together with its cofactor TRBP (5, 13) removes the terminal loop of the pre-miRNA hairpin to generate the ~22-bp miRNA duplex intermediate (6). One strand of this duplex is then incorporated into the RNA-induced silencing complex RISC, where it acts as a guide to direct RISC to complementary mRNA targets (12, 18, 27). Depending upon the degree of complementarity, RISC bind-

ing can result in the cleavage and eventual degradation of the target mRNA or can lead to translation inhibition (9, 31).

Although the initial primary miRNA transcript has the same structure as a pre-mRNA, the fact that the first step in miRNA processing involves the nuclear cleavage of the primary miRNA precursor would seem to preclude a primary miRNA from simultaneously acting as an mRNA (6). To circumvent this problem, cellular miRNAs are generally derived from polymerase II transcripts that are noncoding or, alternatively, are excised out of introns present in noncoding or coding polymerase II transcripts (3, 6, 15). The latter strategy can permit the coordinate expression of one or more miRNAs together with a protein-coding mRNA. Finally, if the primary miRNA cleavage step is inefficient, then a single polymerase II transcript could be used either as a primary miRNA or as an mRNA, depending on whether the primary miRNA is cleaved prior to nuclear export or exported prior to nuclear Droscha cleavage.

In latently infected cells, KSHV gene expression is largely restricted to four genes clustered together between coordinates 117436 and 127886 in the ~141,000-bp KSHV genome (Fig. 1A) (14, 25). In addition, all 11 KSHV miRNAs are encoded within this region and all are in the same genomic orientation, suggesting that all 11 KSHV miRNAs might derive from the same primary miRNA precursor (4, 21, 24). Several pre-mRNAs and mRNAs have been mapped to the latency-associated region of the KSHV genome (Fig. 1A), including a candidate KSHV primary miRNA that also has the potential to function as a pre-mRNA for the KSHV kaposin proteins (8, 16, 23, 26, 29). This kaposin mRNA has been proposed to initiate at coordinate 123842 and polyadenylate at 117436 and bears an intron extending from 123594 to 118799 (Fig. 1A) (16). Ten of the KSHV miRNAs map within this intron, while

\* Corresponding author. Mailing address: Duke University Medical Center, Box 3025, Durham, NC 27710. Phone: (919) 684-3369. Fax: (919) 681-8979. E-mail: culle002@mc.duke.edu.

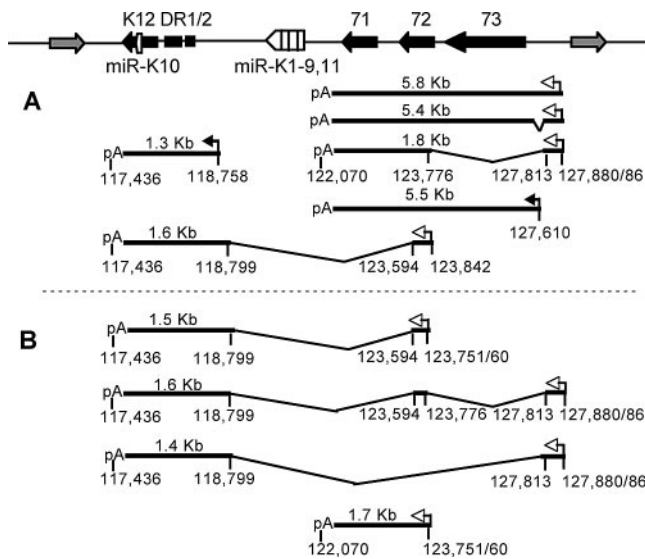


FIG. 1. Transcription of the latency-associated region of the KSHV genome. The KSHV latency-associated region encodes four proteins, indicated by black boxes, as well as 11 miRNAs (white boxes) and is flanked by lytic genes (gray boxes). In panel A, previously reported promoters, splice sites, and poly(A) addition sites are indicated. Lytic promoters are indicated by black arrows and latent promoters by white arrows. Novel KSHV transcripts identified in this report are shown in panel B. Sequence coordinates are derived from the KSHV genome sequence obtained from BC-1 cells (GenBank accession number NC\_003409).

one miRNA, miR-K10, is located within a kaposin open reading frame (ORF) (Fig. 1A).

Here, we extend these earlier reports by identifying several novel mRNAs encoded within this region of the KSHV genome and by directly demonstrating which KSHV pre-mRNAs can also function as KSHV primary miRNAs. Moreover, we demonstrate that KSHV miRNA expression requires readthrough of a viral polyadenylation site that appears to have evolved to be intrinsically inefficient.

#### MATERIALS AND METHODS

**Construction of molecular clones.** Plasmids pEGFP-N3, pTRE2hyg, and pTet-Off were purchased from BD/Clontech. pCMV-Luc has been previously described (3). KSHV promoter plasmids were generated by substitution of the cytomegalovirus promoter present in pEGFP-N3 with PCR-generated KSHV DNA fragments in either the sense or antisense orientation. The *gfp* gene in pEGFP-N3 was then replaced with the firefly luciferase indicator gene. The pLuc-ORF71p(A) and pLuc-HIV-LTRp(A) indicator constructs were generated by inserting ~300-bp PCR fragments containing the KSHV polyadenylation site at 122070, or containing the human immunodeficiency virus type 1 long terminal repeat polyadenylation site, into a unique XhoI site located in the 3' untranslated region of the luciferase gene present in pCMV-Luc. pTRE(123751–117169) and pTRE(127880–117169) were generated by inserting the relevant KSHV genomic DNA fragments into pTRE2hyg in place of the genomic  $\beta$ -globin polyadenylation site.

**Cell culture and transfection.** The latently KSHV-infected B-cell lines BC-1 and BCBL-1, the uninfected B-cell line BJAB, and 293T cells were maintained as previously described (4). BC-1 cells were incubated with tetradecanoyl phorbol acetate (TPA; Sigma, 25 ng/ml) for 48 hours to induce lytic KSHV replication. All 293T cell transfections were performed using Lipofectamine 2000 in 24-well plates. For luciferase assays, 400 ng of each firefly luciferase-based reporter plasmid was cotransfected with 1 ng of a *Renilla* luciferase internal control plasmid in each well. Induced luciferase activities were measured at 48 hours after transfection, as previously described (3).

**RNA isolation and analysis.** All RNA samples were extracted using TRIzol (Invitrogen). Reverse transcription-PCR assays were performed as previously described (3) using oligo(dT)-primed cDNA and the KSHV-specific primers F (5'-AGGCAGCTGCGCCACGAAG-3') and R (5'-CACCGCTCCTCGATACGTGTA-3'). RNase protection assays (RPAs) were used to determine the 5' and 3' ends of all KSHV transcripts analyzed in this study. The riboprobes used in the RPAs were generated by *in vitro* transcription of PCR products containing the T7 promoter and a 5' 30-bp nonspecific tag. The RPAs were performed using the Hybspeed RPA kit (Ambion). Band density was then quantified using a PhosphorImager.

After transfection of 293T cells with pTRE(123751–117169) or pTRE(127880–117169), in the presence or absence of pTet-Off, the cells were cultured in the presence or absence of 1  $\mu$ g/ml of doxycycline. At 48 hours, total RNA was harvested and Northern analysis of kaposin mRNA expression was performed as previously described (3, 4). The template for the K12-specific probe was amplified from KSHV genomic DNA using previously described primers (4). The primers used for KSHV miRNA detection by primer extension analysis have the following sequences: PE-K1, 5'-TGTCAGCTTACACCCAGTTCC-3'; PE-K3, 5'-ACTACGTCGCTGCCGTCCTCAGA-3'; PE-K4, 5'-CATTGCCCTAGAGTACTGCGGT-3'; and PE-K9, 5'-ACTCGCAGCGGGGTTTACGCAG-3'.

#### RESULTS

Previous work analyzing the transcripts generated from the latency-associated region of the KSHV genome has identified four promoter elements and two polyadenylation sites (Fig. 1A) (8, 16, 19, 23, 26, 29). The KSHV ORFs 73, encoding LANA; 72, encoding v-cyclin; and 71, encoding v-Flip, are all expressed from mRNAs that initiate at ~127886 and have been reported to be polyadenylated at 122070 (8, 26, 29). This ~5.8-kb initial transcript can be expressed in an unspliced form or can be alternatively spliced to give ~5.4-kb or ~1.8-kb mRNA species, neither of which contain any of the KSHV miRNAs. In the case of the 1.8-kb mRNA, splicing removes an intron extending from 127813 to 123776 and gives rise to an mRNA that expresses both v-cyclin and v-Flip (11). This promoter element is active in latently infected cells and is not significantly affected by induction of lytic KSHV replication by treatment with TPA.

A second KSHV promoter transcribes an unspliced mRNA that initiates transcription at 118758 and is polyadenylated at 117436, which encodes the KSHV kaposin proteins (23). The kaposins are translated from overlapping open reading frames that include the K12 sequence as well as two adjacent direct repeat elements termed DR1 and DR2 (Fig. 1A) (23). This mRNA, which also contains the viral miR-K10 miRNA, is almost undetectable in latently KSHV-infected cells but is dramatically induced when lytic replication is induced by TPA treatment.

The fifth transcript reported for the KSHV latency-associated region has been reported (16) to initiate at 123842 and polyadenylate at 117436 (Fig. 1A). This ~1.6-kb mRNA is spliced, to remove an intron extending from 123594 to 118799, and also encodes the kaposins (16). Moreover, this pre-mRNA contains all the viral miRNAs, with miR-K1 to miR-K9, as well as miR-K11, mapping to the intron, while miR-K10 is located in the K12 ORF.

Importantly, while this promoter element appears to be active in latently infected cells, it has also been proposed to be TPA responsive (16). Although this pre-mRNA is therefore a strong candidate for the KSHV primary miRNA precursor, the lack of significant (>2-fold) induction of most viral miRNAs upon TPA treatment of latently infected cells (4, 21) either appears inconsistent with the hypothesis that

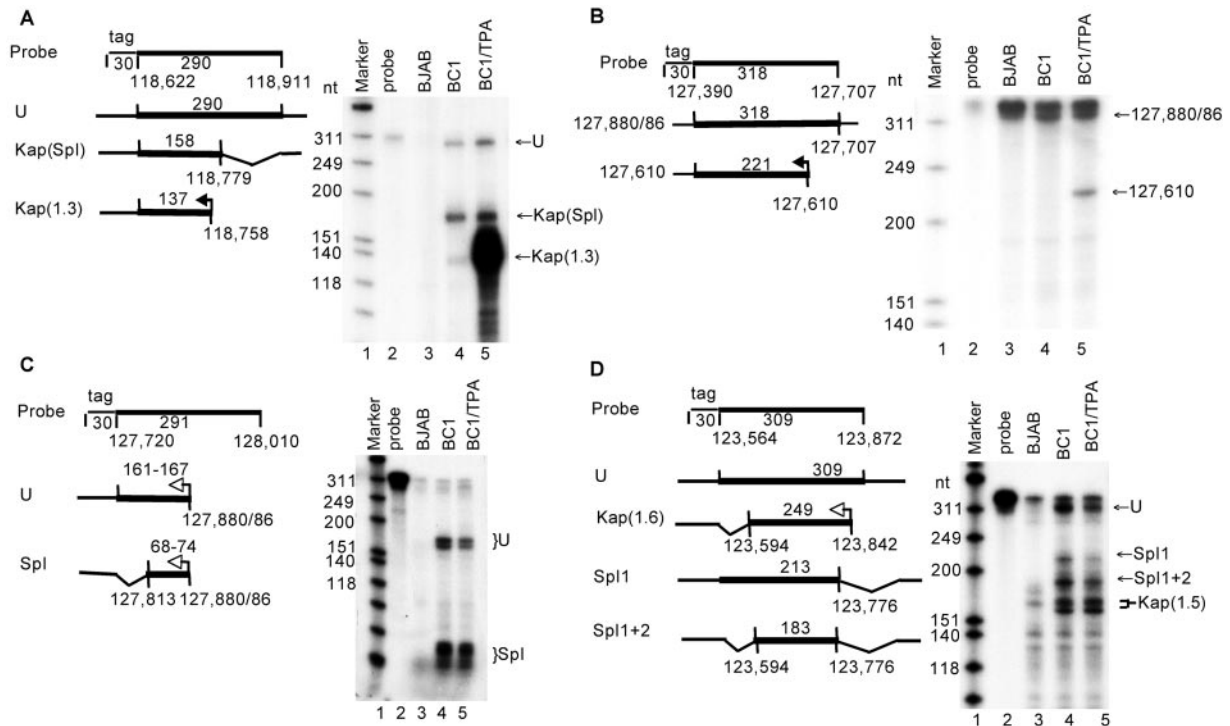


FIG. 2. Characterization of promoters located in the KSHV latency-associated region. (A) This RPA utilized a probe designed to detect RNAs that initiate transcription at position 118758 as well as spliced (Spl) and unspliced (U) KSHV RNAs that initiate 5' to position 118911. The RNA samples used were from the uninfected B-cell line BJAB (lane 3) or from the latently KSHV-infected cell line BC-1. BC-1 cells were cultured under conditions that allow only ~0.9% of cells to spontaneously enter lytic KSHV replication (lane 4) or were treated with TPA, which induces lytic KSHV replication in ~20% of the BC-1 cells (lane 5). (B) Similar to panel A, except that this 348-nucleotide RPA probe was designed to detect transcripts initiating at the 127610 promoter. Transcripts initiating at 127610 would rescue a 221-nucleotide probe fragment, while RNAs initiating at the 127880/86 promoter would rescue a 318-nucleotide probe fragment. (C) Similar to panel A, except that the RPA probe used detects transcripts that initiate at the 127880/86 promoter. Unspliced RNAs (U) and RNAs that utilize the 5' splice site at 127813 (Spl) were detected. (D) Similar to panel A except that this RPA probe was designed to detect transcripts adjacent to the proposed 123842 cap site. Transcripts initiating at the 127880/86 promoter should give rise to 309- and 213-nucleotide rescued fragments representing unspliced (U) and singly spliced (Spl) RNAs, respectively. A 183-nucleotide probe fragment is predicted if a transcript initiating at 127880/86 is doubly spliced (Spl1+2). No band at the predicted size of 249 nucleotides, reflecting transcription initiation at 123842, was detected. The provenance of the ~167- and ~158-nucleotide probe fragments is discussed in the text.

this promoter is TPA responsive or suggests that the viral miRNAs derive from a different primary miRNA.

Recently, a fourth promoter element, termed the  $LT_i$  promoter, has been reported in this region of the KSHV genome (19). This promoter initiates transcription at ~127610, i.e., ~270 bp 3' to the latent promoter located at ~127880 and also 3' to the 5' splice site located at 127813 (Fig. 1A). It has been proposed that transcripts initiating at the 127610 promoter are polyadenylated at the 122070 polyadenylation addition site (Fig. 1A) and function as unspliced mRNAs encoding the ORF73, and possibly also the ORF72 and ORF71, proteins (19). While the 127610 promoter has been reported to be inactive in latently infected cells, and so could not contribute to KSHV miRNA expression in these cells, this promoter is induced during lytic replication (19).

**Functional analysis of KSHV promoter elements.** As a first step towards characterizing the transcriptional origin of the KSHV miRNAs, we decided to confirm the location and TPA responsiveness of the four KSHV promoter elements listed in Fig. 1A by RNase protection analysis using RNA samples derived from the KSHV-infected primary effusion lymphoma (PEL) cell line BC-1 (25), either untreated or treated with

TPA, or derived from the uninfected human B-cell line BJAB. Because of the proximity of splice sites to many of the transcription initiation sites, this approach also allowed us to examine whether these transcripts are processed as predicted.

To analyze the KSHV promoter that initiates at 118758 (Fig. 1A), we used an RPA probe that extends from 118622 to 118911 (Fig. 2A). Like all other probes used in this paper, this 320-nucleotide RPA probe contains a 5' 30-nucleotide sequence of heterologous origin that acts as a "tag" to allow us to discriminate the input probe from any rescued probe fragment. Based on the RPA probe design used, we would predict that we might recover a 290-nucleotide probe fragment that would result from unspliced KSHV RNAs that initiated 5' to 118911. Any RNAs that initiate 5' to 118911 that utilize the 3' splice acceptor at 118779 should give rise to a 158-nucleotide probe fragment. Finally, mRNAs that initiate transcription at the 118758 promoter would rescue a 137-nucleotide probe fragment.

As shown in Fig. 2A, all three of these probe fragments were recovered when RNA from uninduced BC-1 cells was analyzed (Fig. 2A, lane 4), while no probe fragments were rescued by BJAB RNA (Fig. 2A, lane 3). When the BC-1 cells were



treated with TPA to induce lytic gene expression, the level of probe fragment rescued by the spliced or unspliced KSHV RNAs that initiate 5' to 118911 was largely unaffected. In contrast, transcription initiating at 118758 was dramatically induced (Fig. 2A, lane 5). These data confirm the existence of the 118758 KSHV promoter and also confirm that it is induced by lytic replication (23). Our data do not address whether this promoter is active in latently infected cells, as the very weak signal seen in Fig. 2A, lane 4, could be due to the ~0.9% of BC-1 cells that spontaneously enter lytic replication under our culture conditions (4). These data also confirm the existence of the 3' splice site at 118779 (16) and argue that KSHV RNAs that use this site, i.e., presumably mRNAs that initiate at 123842, are at most weakly induced during lytic replication.

Analysis of the KSHV promoter located at 127886 also gave the predicted result (Fig. 2C). This transcription initiation site has actually been reported to be imprecise, in that it extends from 127880 to 127886 (8). Consistent with this earlier work, we observed probe bands of 161 and 167 nucleotides, representing fragments rescued by unspliced RNAs initiating at this promoter, and fragments extending from 68 to 74 nucleotides, as predicted if the probe annealed to KSHV RNAs that initiated at 127880 or 127886 and then utilized the 5' splice site at 127813 (Fig. 1A). Importantly, and as previously reported (8, 26, 29), this promoter is not significantly induced by TPA treatment (Fig. 2C, lane 5).

In Fig. 2B, we analyze transcription initiating from the LT<sub>1</sub> promoter located at ~127610 (19). As predicted from earlier work, this promoter proved to be inactive in latently KSHV-infected cells but was induced by treatment with TPA, as demonstrated by rescue of an ~221-nucleotide probe fragment by RNA derived from TPA-treated (Fig. 2B, lane 5) but not uninduced (Fig. 2B, lane 4) BC-1 cells. Even after induction, this promoter appeared to be relatively weak, as the level of the ~318-nucleotide probe fragment rescued by transcripts initiating 5' to the 127610 promoter, i.e., presumably unspliced mRNAs initiating at the 127880/86 latent promoter, remained somewhat higher even in the presence of TPA (Fig. 2B). However, it should be recalled that TPA treatment induces lytic KSHV replication in only ~20% of BC-1 cells, so that this result undoubtedly underestimates the level of transcription from the 127610 promoter during lytic replication. Nevertheless, compared to the lytic promoter located at 118911 (Fig. 2A), the promoter located at 127610 is clearly fairly weak.

While the three KSHV promoters located at 118758, 127610, and 127880/86 therefore behaved as predicted, analysis of the proposed KSHV promoter located at 123842 (16) provided several surprises (Fig. 2D). This promoter is located adjacent to a 5' splice site at 123594, so an RNA that initiates at 123842 and splices at 123594 should rescue a 249-nucleotide probe fragment. However, no such probe fragment was detected (Fig. 2D, lanes 4 and 5). Instead, we detected probe fragments of 213 and 183 nucleotides in length as well as a doublet at ~167 and ~158 nucleotides.

The 213-nucleotide rescued fragment is expected, as this is the probe fragment that should be rescued by the 1.8-kb mRNA that initiates at 127886 and utilizes the 3' splice site at 123776 (indicated as Sp11 in Fig. 2D). The 183-nucleotide rescued probe fragment is the size that would be predicted if both the 3' splice site at 123776 and the 5' splice site at 123594

were used by a KSHV RNA initiating at 127886. As discussed in more detail below, this doubly spliced RNA could exist only if some of the KSHV RNAs that initiate at 127886 read through the polyadenylation site at 122070 and then use the 3' splice site at 118799 (Fig. 1B).

Finally, we believe that the probe doublet of ~167 and ~158 nucleotides actually represents transcription initiation at 123760 and 123751, i.e., we hypothesize that the KSHV promoter proposed to initiate at 123842 has been incorrectly mapped and that the actual cap site is 82 to 91 nucleotides 3' to this proposed cap site. In fact, in the original paper that reported this viral promoter element (16), the authors obtained discordant data when they mapped the transcription start site by S1 nuclear protection analysis, which gave the 123842 transcription initiation site, and 5' random amplification of cDNA ends, which indicated a cap site 90 nucleotides 3' to 123842, i.e., at the same location mapped in this analysis.

To address this issue in more detail, we constructed indicator plasmids that link three of the proposed KSHV promoters listed in Fig. 1A to the firefly luciferase indicator gene. The promoter fragments used extended from 119208 to 118708 (-450 to +50 relative to the 118758 cap site), 124292 to 123820 (-450 to +22 relative to the proposed cap site at 123842), 124292 to 123692 (-532 to +59/+68 relative to the alternative cap sites at 123751 and 123760), and finally from 128336 to 127836 (-450 to +50 relative to the cap site at 127886). These indicator constructs were then transfected into 293T cells and luciferase activity assayed 48 hours after transfection. An inverted form of the 128336/127836 promoter was used as a negative control.

As shown in Fig. 3, the indicator plasmids containing KSHV DNA sequences flanking the 118758, 123751/60, and 127880/86 cap sites all gave rise to readily detectable levels of luciferase activity in transfected cells. In contrast, the indicator construct that contained genomic KSHV sequences extending 22 bp 5' and 450 bp 3' to the proposed cap site at 123842 did not give any detectable luciferase activity. Moreover, RPA analysis of RNA recovered from 293T cells transfected with the indicator construct containing the 124292 to 123692 KSHV DNA segment mapped exactly the same two transcription initiation sites seen in the KSHV-infected BC-1 cells (Fig. 2D and data not shown). Based on these data, we therefore conclude that the transcription initiation site for this KSHV promoter is actually at 123751/60 and not at 123842, as previously proposed (16). Our data indicate that this promoter is not activated by induction of KSHV lytic replication by TPA (Fig. 2D, lane 5), a result which is consistent with the similar lack of TPA induction of KSHV mRNAs that utilize the 3' splice site at 118799 (Fig. 2A, lane 5).

In conclusion, the RPA analysis performed in Fig. 2 has confirmed the existence of four KSHV promoters within the viral latency-associated region. Two of these promoters, one initiating transcription at 127880/86 and the second now remapped to 123751/60, are latent promoters that are not induced by TPA treatment. The other two promoters, which initiate transcription at 118758 and 127610, respectively, are induced by TPA treatment and appear not to be active during latent infection. In addition, the RPA analysis presented in Fig. 2 confirms the existence of the two 5' KSHV splice sites pre-

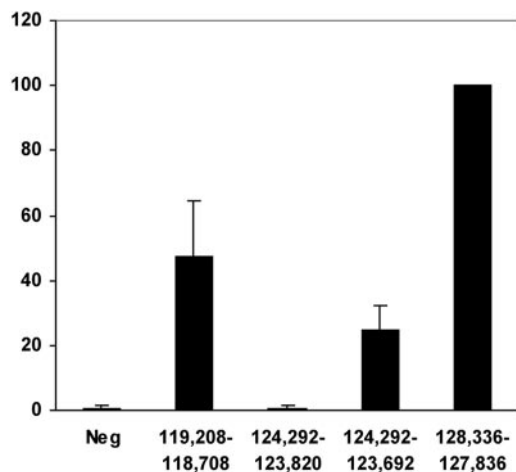


FIG. 3. Biological activity of candidate KSHV promoter elements. Plasmids containing the indicated KSHV DNA fragments linked to the luciferase indicator gene were transfected into 293T cells and induced luciferase activities determined. A *Renilla* luciferase expression vector was cotransfected as an internal control. The data presented were normalized to the internal control and are given relative to the pTRE(128336-127836)Luc indicator plasmid, which was set at 100. The negative (Neg) control contains the antisense version of the 128336/127836 KSHV sequence linked to the luciferase gene. The average for three experiments with standard deviations is indicated.

viously mapped to 123594 and 127813 as well as the two viral 3' splice sites located at 118779 and 123776 (8, 16, 26, 29).

**A KSHV polyadenylation site is subject to frequent read-through.** An interesting observation reported in Fig. 2D is the evidence suggesting the existence of KSHV mRNAs that utilize both the 3' splice site at 123776 and the 5' splice site at 123594. Each of these splice sites has been previously reported (the 3' splice site at 123776 in the ~1.8-kb mRNA encoding ORF72 and ORF71 and the 5' splice site at 123594 in the ~1.6-kb mRNA encoding the kaposin proteins) (Fig. 1A) (8, 16, 26, 29) but no KSHV mRNA that uses both these splice sites has been previously identified. Based on the known KSHV promoters (Fig. 1), the only possible source of this novel mRNA is the latent promoter located at 127886. Consistent with this hypothesis, the probe fragment derived from this doubly spliced KSHV mRNA is not induced by treatment of BC-1 cells with TPA (Fig. 2D, lane 5).

To confirm the existence of a latent KSHV mRNA that initiates at 127886 and is polyadenylated at 117436, we performed a reverse transcription-PCR analysis on RNA samples derived from the latently KSHV-infected PEL cell lines BC-1 and BCBL-1. As shown in Fig. 4, both cell lines gave rise to the 332-bp PCR product, labeled L, that would be predicted if this doubly spliced KSHV mRNA indeed existed. Moreover, a shorter 149-bp PCR product, labeled S, was also identified in the BC-1 cells, but not in the BCBL-1 cells, that derives from a KSHV mRNA that splices directly from 127813 to 188779, thereby omitting the 183-nucleotide noncoding exon that extends from 123776 to 123594. The predicted identity of both the L and the S reverse transcription-PCR products identified in Fig. 4 was confirmed by DNA sequencing (data not shown). As expected, neither the L nor the S reverse transcription-PCR

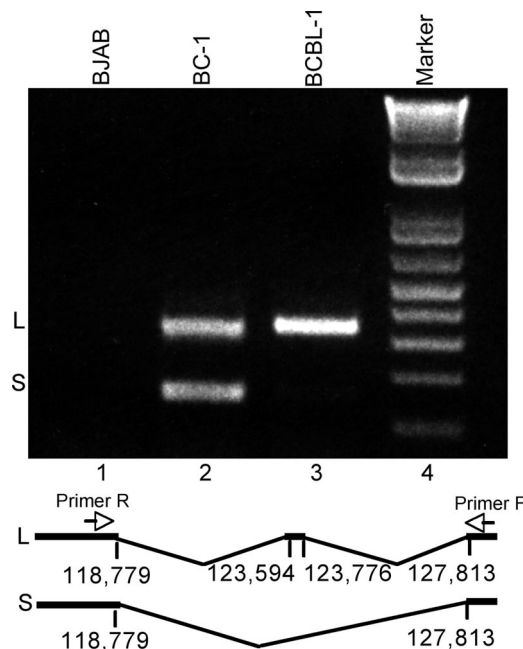


FIG. 4. Reverse transcription-PCR analysis of spliced kaposin mRNAs derived from the 127880/86 promoter. Total RNA derived from BJAB, BC-1, or BCBL-1 cells was subjected to reverse transcription using oligo(dT) primers and then PCR using primers extending from 118098 to 118718 (R) and from 127861 to 127879 (F) in the KSHV genome. This analysis yielded a large (L) fragment of 332 bp and a short (S) fragment of 149 bp. DNA sequencing demonstrated that these two fragments were derived from KSHV RNAs that had undergone alternative splicing as indicated.

product was obtained when RNA from uninfected BJAB cells was analyzed (Fig. 4).

If KSHV indeed expresses kaposin mRNAs in latently infected cells that initiate at 127886 and are polyadenylated at 117436, then the signals that trigger polyadenylation at 122070 must be ignored at least some of the time. Indeed, the previously reported (16) kaposin mRNA that was proposed to initiate at 123842 (now revised to 123751/60) must also read through this same polyadenylation site. We therefore asked whether we could detect high levels of readthrough of the KSHV polyadenylation site at 122070, compared to the polyadenylation site at 117436, using RPA.

As shown in Fig. 5A, 42% of the KSHV transcripts that traverse the 122070 polyadenylation site were found to be processed at this site at steady state, while 58% were not processed. In contrast, 86% of KSHV transcripts that traverse the polyadenylation site at 117436 were found to be processed at this site, while 14% were not processed (Fig. 5B). While these data strongly suggest that the 122070 polyadenylation site is indeed less efficiently utilized than the viral 117436 polyadenylation site, we actually felt that this result was likely to overstate the efficiency of polyadenylation at the 122070 site. In particular, if polyadenylation at this site did occur, it would give rise to a stable mRNA encoding KSHV ORF73, ORF72, and/or ORF71 (Fig. 1A). In contrast, if polyadenylation did not occur, then the polyadenylation site would be expected to form part of an intron that would be removed in the nucleus

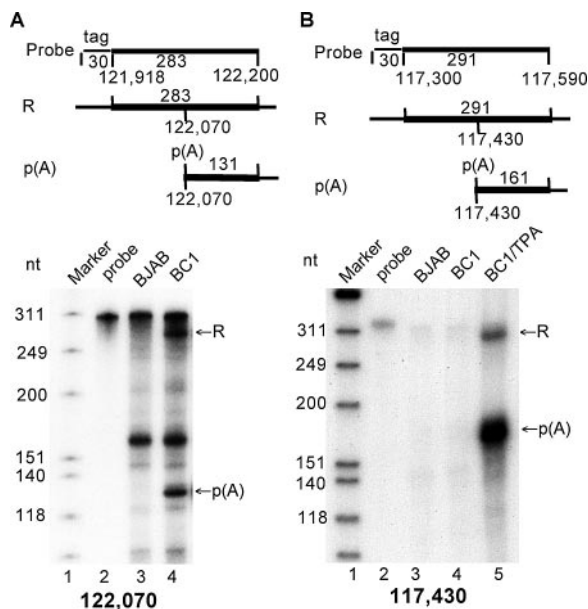


FIG. 5. Relative polyadenylation efficiency at the KSHV 117430 and 122070 polyadenylation sites in BC-1 cells. This RPA analysis uses probes that detect KSHV transcripts that read through (R) the 117430 or 122070 polyadenylation site or are processed [p(A)] at these sites. The RNA samples used were from control BJAB cells or from latently KSHV-infected BC-1 cells. The probe design and the predicted rescued probe fragments are indicated at the top.

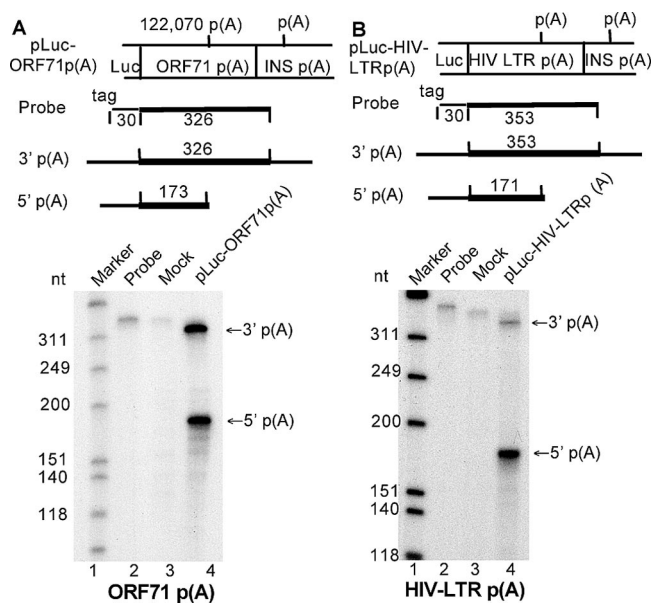


FIG. 6. Analysis of processing efficiency at the KSHV 122070 poly(A) addition site. This analysis used indicator constructs in which ~300-bp DNA fragments that include the entire KSHV 122070 polyadenylation site (A) or the human immunodeficiency virus type 1 long terminal repeat (HIV-1 LTR) polyadenylation site [p(A)] (B) were introduced between a luciferase indicator gene and a highly active genomic rat preproinsulin II (INS) polyadenylation site. The probes utilized distinguish between RNAs processed at the introduced 5' polyadenylation site and those processed at the 3' insulin polyadenylation site or that remain unprocessed.

and then degraded either by the cellular debranching enzyme or as a result of cleavage by the Droscha/DGCR8 dimer.

As a more rigorous, albeit also more artificial, test of polyadenylation site efficiency, we therefore used a competitive assay for polyadenylation efficiency that has been described previously (2). In this assay, a 296-bp KSHV DNA fragment, extending 141 bp 5' and 149 bp 3' to the AATAAA polyadenylation signal for the 122070 polyadenylation site, was inserted into an expression plasmid between the luciferase indicator gene and a highly active genomic polyadenylation site derived from the rat preproinsulin II gene. Similarly, a 302-bp DNA fragment extending 168 bp 5' and 149 bp 3' to the AATAAA polyadenylation signal located in the human immunodeficiency virus type 1 long terminal repeat promoter element was also inserted in the same location. When the luciferase gene in these indicator plasmids is transcribed, the resultant mRNAs have a choice of being polyadenylated at the 5' KSHV or human immunodeficiency virus type 1 polyadenylation site or at the 3' preproinsulin polyadenylation site. If the inserted 5' polyadenylation site is effective, then it will be favored over the 3' polyadenylation site because it is transcribed first (2). However, if the inserted polyadenylation site is weak, then the 3' preproinsulin polyadenylation site will be used preferentially.

As can be observed in Fig. 6, the KSHV 122070 polyadenylation site was used 39% of the time, while 61% of the luciferase mRNAs were either polyadenylated at the preproinsulin polyadenylation site or left unprocessed (Fig. 6A). (Because the preproinsulin polyadenylation site is highly active, we have previously observed very little nonpolyadenylated RNA when using comparable indicator plasmids [2]). In contrast, when the

human immunodeficiency virus type 1 long terminal repeat polyadenylation site was introduced, ~90% of the luciferase mRNAs were polyadenylated at the human immunodeficiency virus long terminal repeat polyadenylation site, while only ~10% were polyadenylated at the preproinsulin polyadenylation site or left unprocessed (Fig. 6B).

Together, these data, looking at the efficiency of polyadenylation at the KSHV 122070 polyadenylation site either in infected cells (Fig. 5) or in transfected cells (Fig. 6), indicate that a substantial percentage of mRNA transcripts that traverse this site are not polyadenylated at this site. These data are therefore consistent with the evidence, reported in Fig. 2 and 4 and previously by Li et al. (16), arguing for the existence of KSHV mRNAs that result from readthrough of the 122070 polyadenylation site.

**Two KSHV pre-mRNAs can function as viral primary miRNAs.** The data presented so far identify three pre-mRNAs that could serve as primary miRNAs for all the KSHV miRNAs expressed in latently infected cells (Fig. 1B). These include a pre-mRNA that initiates at 123751/60 and is polyadenylated at 117436 and contains a single intron extending from 123594 to 118799. This pre-mRNA is closely similar to a pre-mRNA proposed by Li et al. (16) (Fig. 1A) except that we have defined a slightly different transcription start site. In addition, we have identified two novel KSHV pre-mRNAs that extend from 127880/86 to 117436 and that differ in terms of the inclusion or exclusion of a short noncoding exon (Fig. 1B). Of note, all three of these pre-mRNAs are predicted to give rise





thus strongly suggesting that latently KSHV-infected cells indeed express this novel ~1.7-kb v-Flip/v-cyclin mRNA.

## DISCUSSION

The primary aim of this study was to define the transcriptional origin of the 11 KSHV miRNAs expressed in latently infected human cells. As all these miRNAs are encoded within the latency-associated region of KSHV in the same transcriptional orientation (4, 21, 24), it seemed possible that they might all be processed out of a single primary miRNA precursor. Indeed, a candidate KSHV primary miRNA had been previously reported (16) in the form of a kaposin pre-miRNA that was reported to initiate at position 123842 in the KSHV genome and to be polyadenylated at 117436 (Fig. 1A). Ten of the 11 KSHV miRNAs map to an ~4.8-kb intron present in this pre-miRNA, while the eleventh miRNA is located within the K12 ORF that is used for translation of one of the kaposin proteins (Fig. 1A). However, this kaposin pre-mRNA had been reported to be induced when latently KSHV-infected cells are triggered to enter lytic replication by treatment with TPA (16). In contrast, expression of the KSHV miRNAs is largely unaffected by TPA treatment (4, 21). The KSHV latency-associated region has also been reported to express a pre-mRNA that initiates at 127880/60 and is polyadenylated at 122070 that gives rise to mRNAs encoding the ORF 71, 72, and 73 proteins (Fig. 1A). The promoter driving these mRNAs had been reported to be largely unaffected by treatment of latently infected cells with TPA.

Using a series of RPA probes specific for the regions flanking each reported promoter in the latency-associated region, we have largely confirmed these previous reports (Fig. 2). One exception is that our analysis suggests that the proposed 123842 transcription initiation site (16) is inaccurate and that this KSHV promoter actually initiates transcription at 123751/60 (Fig. 1B). Moreover, our RPA data argue that this promoter element is not affected by TPA treatment, i.e., it is a true latency-specific KSHV promoter element.

While the 123751/60 promoter is therefore clearly predicted to be able to act as a primary miRNA precursor for the KSHV miRNAs during latent infection, our results also identify two novel kaposin pre-mRNAs that initiate at the latent 127880/86 promoter element and are then polyadenylated at the more distal 117436 polyadenylation site rather than at the 122070 polyadenylation site (Fig. 1B and 4). The majority of these ~10,500-nucleotide pre-mRNA transcripts were then spliced to remove two previously defined introns (8, 16, 26, 29) extending from 127813 to 123776 and from 123594 to 118799, although a minority were spliced directly from 127813 to 118799 (Fig. 4). Importantly, these two alternatively spliced kaposin pre-mRNAs should also have the potential to encode the various KSHV miRNAs. Indeed, direct expression of these two transcripts, and of the shorter 123751 to 117436 KSHV pre-mRNA, under the control of a tetracycline-regulatable promoter in transfected 293T cells directly demonstrated the expression of not only the predicted kaposin mRNAs (Fig. 7A) but also KSHV miRNAs (Fig. 7B). Because this expression was observed only when the TRE promoter was activated by coexpression of the Tet-Off activator in the absence of doxycycline (Fig. 7), this result strongly argues that these are the authentic

KSHV primary miRNAs and that no other KSHV promoter contributes substantially to KSHV miRNA expression in latently infected cells.

An important point about the KSHV primary miRNAs defined in Fig. 1B is that their expression is totally dependent on readthrough of the polyadenylation signal located at 122070 (Fig. 1). Indeed, we have directly observed that this viral polyadenylation addition site is utilized only inefficiently when analyzed in latently KSHV-infected cells (Fig. 5) or when analyzed in the context of artificial indicator constructs (Fig. 6). Partial utilization of this site is however essential in order to allow the expression of latent mRNAs transcribed from the 127880/86 promoter that encodes the ORF 71, 72, and 73 proteins (Fig. 1). Moreover, our data suggest that some transcripts that initiate at the 123751/60 promoter are also polyadenylated at the 122070 site, thus giving rise to a novel KSHV mRNA that has the potential to encode the ORF72 and ORF71 proteins (Fig. 6C and 1B).

The KSHV transcripts that initiate at 127880/86 can give rise, through alternative splicing and alternative polyadenylation site usage, to latent mRNAs that encode not only kaposin but also the ORF 71, 72, and 73 proteins, as well as functioning as KSHV primary miRNAs for all 11 viral miRNAs. It is therefore not entirely clear why KSHV has also evolved a second latent promoter at 123751/60 that serves essentially the same function, with the exception that this promoter cannot generate mRNAs encoding the ORF73 protein (Fig. 1B). Based on the RPA shown in Fig. 2D, it appears that these two promoters may be equivalently active, at least in latently infected BC-1 cells. Possibly, these promoters are however differentially regulated in other latently KSHV-infected cells *in vivo*.

While expression of the KSHV miRNAs is only modestly induced by activation of lytic replication by treatment of latently infected cells with TPA, a modest ( $\leq 2$ -fold) increase in expression, particularly of miR-K10, has been observed. Because miR-K10 is unique in being located within the kaposin mRNA expressed by the lytic 118758 promoter (Fig. 1A), enhanced miR-K10 expression would be predicted upon TPA treatment. Moreover, the 127610 promoter is also induced by TPA treatment (Fig. 2B) (19) and, if RNAs transcribed from this promoter can also read through the polyadenylation site at 122070, then a modest increase in the expression of the other KSHV miRNAs might also be expected after TPA treatment. Whether this increased expression is functionally relevant is currently unclear.

In conclusion, we have performed an extensive analysis of the transcripts that arise from the KSHV latency-associated region in latently infected cells. This analysis identified three novel KSHV latent transcripts and redefined the precise location and TPA responsiveness of one KSHV promoter (Fig. 1B). Moreover, these data demonstrate that utilization or readthrough of an inefficient viral polyadenylation site, located at 122070, is a key determinant of the fate of viral transcripts that initiate transcription at the latent promoters located at 127880/86 and at 123751/60. In fact, readthrough of this polyadenylation site appears to be critical for the expression of all of the KSHV miRNAs, and possibly also of the kaposin proteins, in latently KSHV-infected cells.

These data confirm the hypothesis (4, 21, 24) that 10 of the



KSHV miRNAs are processed out of a viral intron, while the 11th, miR-K10, is processed out of the K12 open reading frame. As the presence of the miR-K10 precursor does not appear to substantially inhibit the production of kaposin mRNA (Fig. 7A), we can only conclude that processing of the miR-K10 RNA stem-loop by the Drosha/DGCR8 dimer is inefficient. We note that our data also suggest that expression of the KSHV miRNAs, and of the viral kaposin, ORF71, and ORF72 proteins is coordinately regulated in latently infected cells. Whether this coexpression is, in fact, functionally important must await the definition of the mRNA targets for the KSHV miRNAs.

After submission of this manuscript, Pearce et al. (20) reported somewhat similar results. In particular, this group also showed that the latent KSHV promoter previously proposed to initiate transcription at position 123842 had been incorrectly mapped and that the transcription start site was both heterogeneous and close to position 123757, i.e., very close to the positions (123751/60) proposed in this report (20). This group also proposed the existence of a short, unspliced transcript initiating at this promoter that would encode ORFs 71 and 72 and proposed, but did not experimentally test, the hypothesis that the polyadenylation site located at 122070 was ignored by a significant proportion of the transcripts initiating at the 123751/60 promoter. Therefore, this work and that published recently by Pearce et al. (20) are in close agreement.

#### ACKNOWLEDGMENTS

We thank Dirk Dittmer for helpful discussions.

This work was sponsored by National Institute of Health grant GM071408.

#### REFERENCES

- Bartel, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**:281–297.
- Brown, P. H., L. S. Tiley, and B. R. Cullen. 1991. Efficient polyadenylation within the human immunodeficiency virus type 1 long terminal repeat requires flanking U3-specific sequences. *J. Virol.* **65**:3340–3343.
- Cai, X., C. H. Hagedorn, and B. R. Cullen. 2004. Human miRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**:1957–1966.
- Cai, X., S. Lu, Z. Zhang, C. M. Gonzalez, B. Damania, and B. R. Cullen. 2005. Kaposi's sarcoma-associated herpesvirus expresses an array of viral miRNAs in latently infected cells. *Proc. Natl. Acad. Sci. USA* **102**:5570–5575.
- Chendrimada, T. P., R. I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, K. Nishikura, and R. Shiekhattar. 2005. TRBP recruits the Dicer complex to Ago2 for miRNA processing and gene silencing. *Nature* **436**:740–744.
- Cullen, B. R. 2004. Transcription and processing of human miRNA precursors. *Mol. Cell* **16**:861–865.
- Denli, A. M., B. B. J. Tops, R. H. A. Plasterk, R. F. Ketting, and G. J. Hannon. 2004. Processing of primary miRNAs by the Microprocessor complex. *Nature* **432**:231–235.
- Dittmer, D., M. Lagunoff, R. Renne, K. Staskus, A. Haase, and D. Ganem. 1998. A cluster of latently expressed genes in Kaposi's sarcoma-associated herpesvirus. *J. Virol.* **72**:8309–8315.
- Doench, J. G., C. P. Petersen, and P. A. Sharp. 2003. siRNAs can function as miRNAs. *Genes Dev.* **17**:438–442.
- Gregory, R. I., K.-P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, and R. Shiekhattar. 2004. The microprocessor complex mediates the genesis of miRNAs. *Nature* **432**:235–240.
- Grundhoff, A., and D. Ganem. 2001. Mechanisms governing expression of the v-FLIP gene of Kaposi's sarcoma-associated herpesvirus. *J. Virol.* **75**:1857–1863.
- Hammond, S. M., E. Bernstein, D. Beach, and G. J. Hannon. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**:293–296.
- Han, J., Y. Lee, K.-H. Yeom, Y.-K. Kim, H. Jin, and V. N. Kim. 2004. The Drosha-DGCR8 complex in primary miRNA processing. *Genes Dev.* **18**:3016–3027.
- Jenner, R. G., M. Mar Albà, C. Boshoff, and P. Kellam. 2001. Kaposi's sarcoma-associated herpesvirus latent and lytic gene expression as revealed by DNA arrays. *J. Virol.* **75**:891–902.
- Lee, Y., M. Kim, J. Han, K.-H. Yeom, S. Lee, S. H. Baek, and V. N. Kim. 2004. MiRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**:4051–4060.
- Li, H., T. Komatsu, B. J. Dezube, and K. M. Kaye. 2002. The Kaposi's sarcoma-associated herpesvirus K12 transcript from a primary effusion lymphoma contains complex repeat elements, is spliced, and initiates from a novel promoter. *J. Virol.* **76**:11880–11888.
- Lund, E., S. Güttinger, A. Calado, J. E. Dahlberg, and U. Kutay. 2004. Nuclear export of miRNA precursors. *Science* **303**:95–98.
- Martinez, J., A. Patkaniowska, H. Urlaub, R. Lührmann, and T. Tuschl. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**:563–574.
- Matsumura, S., Y. Fujita, E. Gomez, N. Tanese, and A. C. Wilson. 2005. Activation of the Kaposi's sarcoma-associated herpesvirus major latency locus by the lytic switch protein RTA (ORF50). *J. Virol.* **79**:8493–8505.
- Pearce, M., S. Matsumura, and A. C. Wilson. 2005. Transcripts encoding K12, v-FLIP, v-cyclin, and the miRNA cluster of Kaposi's sarcoma-associated herpesvirus originate from a common promoter. *J. Virol.* **79**:14457–14464.
- Pfeffer, S., A. Sewer, M. Lagos-Quintana, R. Sheridan, C. Sander, F. A. Grässer, L. F. van Dyk, C. K. Ho, S. Shuman, M. Chien, J. J. Russo, J. Ju, G. Randall, B. D. Lindenbach, C. M. Rice, V. Simon, D. D. Ho, M. Zavolan, and T. Tuschl. 2005. Identification of miRNAs of the herpesvirus family. *Nat. Methods* **2**:269–276.
- Pfeffer, S., M. Zavolan, F. A. Grässer, M. Chien, J. J. Russo, J. Ju, B. John, A. J. Enright, D. Marks, C. Sander, and T. Tuschl. 2004. Identification of virus-encoded miRNAs. *Science* **304**:734–736.
- Sadler, R., L. Wu, B. Forghani, R. Renne, W. Zhong, B. Herndier, and D. Ganem. 1999. A complex translational program generates multiple novel proteins from the latently expressed kaposin (K12) locus of Kaposi's sarcoma-associated herpesvirus. *J. Virol.* **73**:5722–5730.
- Samols, M. A., J. Hu, R. L. Skalsky, and R. Renne. 2005. Cloning and identification of a miRNA cluster within the latency-associated region of Kaposi's sarcoma-associated herpesvirus. *J. Virol.* **79**:9301–9305.
- Sarid, R., O. Flore, R. A. Bohenzky, Y. Chang, and P. S. Moore. 1998. Transcription mapping of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) genome in a body cavity-based lymphoma cell line (BC-1). *J. Virol.* **72**:1005–1012.
- Sarid, R., J. S. Wieszorek, P. S. Moore, and Y. Chang. 1999. Characterization and cell cycle regulation of the major Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) latent genes and their promoter. *J. Virol.* **73**:1438–1446.
- Schwarz, D. S., G. Hutvagner, B. Haley, and P. D. Zamore. 2002. Evidence that siRNAs function as guides, not primers, in the *Drosophila* and human RNAi pathways. *Mol. Cell* **10**:537–548.
- Sullivan, C. S., A. T. Grundhoff, S. Tevethia, J. M. Pipas, and D. Ganem. 2005. SV40-encoded miRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* **435**:682–686.
- Talbot, S. J., R. A. Weiss, P. Kellam, and C. Boshoff. 1999. Transcriptional analysis of human herpesvirus-8 open reading frames 71, 72, 73, K14, and 74 in a primary effusion lymphoma cell line. *Virology* **257**:84–94.
- Yi, R., Y. Qin, I. G. Macara, and B. R. Cullen. 2003. Exportin-5 mediates the nuclear export of pre-miRNAs and short hairpin RNAs. *Genes Dev.* **17**:3011–3016.
- Zeng, Y., R. Yi, and B. R. Cullen. 2003. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl. Acad. Sci. USA* **100**:9779–9784.