JOURNAL OF VIROLOGY, June 1997, p. 4199–4208 Vol. 71, No. 6
0022-538X/97/$04.00+0

The Herpes Simplex Virus Type 1 2.0-Kilobase Latency-Associated Transcript Is a Stable Intron Which Branches at a Guanosine

JANICE M. ZABOLOTNY, CLAUDE KRUMMENACHER, AND NIGEL W. FRASER*
The Wistar Institute, Philadelphia, Pennsylvania 19104-4268

Received 9 January 1997/Accepted 28 February 1997

We have used a minigene construct of the herpes simplex virus type 1 (HSV-1) latency-associated transcript (LAT) gene to analyze its transcripts in transient transfection assays. A 2.8-kb fragment of the approximately 8.5-kb LAT gene encompassing the 2.0-kb LAT was cloned into a eukaryotic expression vector downstream of the cytomegalovirus immediate-early gene promoter. Northern hybridization of RNA isolated from transfected COS-1 cells identified three LAT-specific transcripts, 3.4, 2.0, and 1.4 kb in size. Mapping of these transcripts by Northern hybridization indicated that the 1.4- and 2.0-kb RNAs are nonoverlapping, while the 3.4-kb RNA overlaps both smaller RNAs. Reverse transcription-PCR (RT-PCR) and partial sequencing of the 1.4-kb RNA revealed that this RNA is the spliced exons of the 3.4-kb primary transcript. The 2.0-kb LAT appears to be an intron accumulating after splicing of the minor LAT (mLAT) pre-mRNA. The splice donor and acceptor sites for the 2.0-kb LAT identified in transfected and HSV-1-infected cells are identical. Mapping of the branch point of this intron by RT-PCR in transfected and HSV-1-infected cells, as well as in latently infected murine trigemial ganglia, shows that it is a guanosine. This branch site does not bear homology to consensus mammalian branch site sequences. These data provide evidence that the 2.0-kb LAT is an intron of the mLAT pre-mRNA with a unique branch point.

Herpes simplex virus type 1 (HSV-1) is neurotropic virus capable of forming latent infections for the lifetime of an individual. During latency, transcription is restricted to a single diploid gene within the long repeat elements of the viral genome. Transcription of this gene generates a family of transcripts known as the latency-associated transcripts (LATs) (49, 54, 58, 65) (Fig. 1). An 8.5-kb LAT is postulated on the basis of in situ hybridization of infected tissues and the presence of a promoter element mapping to its 5′ end and of a polyadenylation signal near its 3′ end (41). This transcript is referred to as the minor LAT (mLAT) based on its abundance. Two transcripts, the 2.0-kb LAT and the 1.5-kb LAT, are routinely detectable by Northern hybridization (49, 54, 58, 65). These RNAs are partially collinear and are thought to evolve by differential splicing. The 1.5-kb LAT is observed only during latency in neurons, whereas the 2.0-kb LAT is detectable in productive infections in tissue culture and animals as well as during latency (54, 55, 62). During productive infections, the 2.0-kb LAT is produced with the kinetics of a late gene, but because it is synthesized in latency apparently without prior viral gene expression, it is considered to be a unique class of genes, known as the λ class (55).

The 2.0-kb LAT does not appear to be polyadenylated (19, 44, 54, 62) and also lacks a cap at its 5′ end (19) which maps to a splice donor sequence GT (33, 34, 56, 62). It has been proposed that the 2.0-kb LAT is a stable intron derived from the larger 8.5-kb mLAT (20). Consistent with this proposal, Farrell et al. (23) showed that the 2.0-kb LAT RNA could be spliced out of a β-galactosidase transcript containing the LAT sequences in transient transfections. Additionally, Wu et al. (67) have recently shown that the majority of the 2.0-kb LAT transcript is in a nonlinear structure, most likely a lariat. However, the spliced exons of the putative primary transcript mLAT have never been detected.

Two promoters involved in the generation of the 2.0-kb LAT have been identified. These are known as the latency active promoter 1 (LAP1) (6, 20, 69–71) and the latency active promoter 2 (LAP2) (26). The LAP1 promoter, mapping to the 5′ end of mLAT, is the promoter of this putative transcript. Speculation of a second promoter for the 2.0-kb LAT was prompted by the observation that deletion mutants of the mLAT promoter (LAP1) still produce 2.0-kb LAT during productive infections in tissue culture and animals but not during latency (43). Subsequently, a second promoter element, LAP2, was mapped at or near the 5′ end of the 2.0-kb LAT (26). This promoter, which lacks a TATA box but has a putative initiator element, may drive transcription of the 2.0-kb LAT during productive infections (11, 43). Studies with virus mutants have revealed that LAP1 operates primarily in latency (11, 20) and LAP2 is mainly active in productive infections (11, 43). Since the LAP2 promoter abuts the 5′ end of the 2.0-kb LAT, it has been proposed that the 2.0-kb LAT and the 8.5-kb LAT may be separate transcripts generated by transcription from these different promoters (26).

The functions of the LATs have not been clearly determined. Some LAT deletion mutants display a slow reactivation phenotype (57) or appear to establish latency with reduced efficiency (52). However, these phenotypes may be directly attributable not to the LATs but to other newly identified transcripts that overlap the LAT region (8, 35–37, 53, 68). During latency, the LATs are nuclear localized; however, in productive infections of tissue culture cells and SCID mice brainstems, the 2.0-kb LAT is present in the cytoplasm (44). We have previously shown that the majority of cytoplasmic 2.0-kb LAT is not associated with polysomes in productively infected tissue culture cells (44), suggesting that this transcript is not translated under these conditions. Nevertheless, the preservation of the stability of the LAT intron in all isolated

* Corresponding author. Mailing address: The Wistar Institute, 36th St. at Spruce, Philadelphia, PA 19104-4268. Phone: (215) 898-3847. Fax: (215) 898-3849. E-mail: fraser@wista.wistar.upenn.edu.

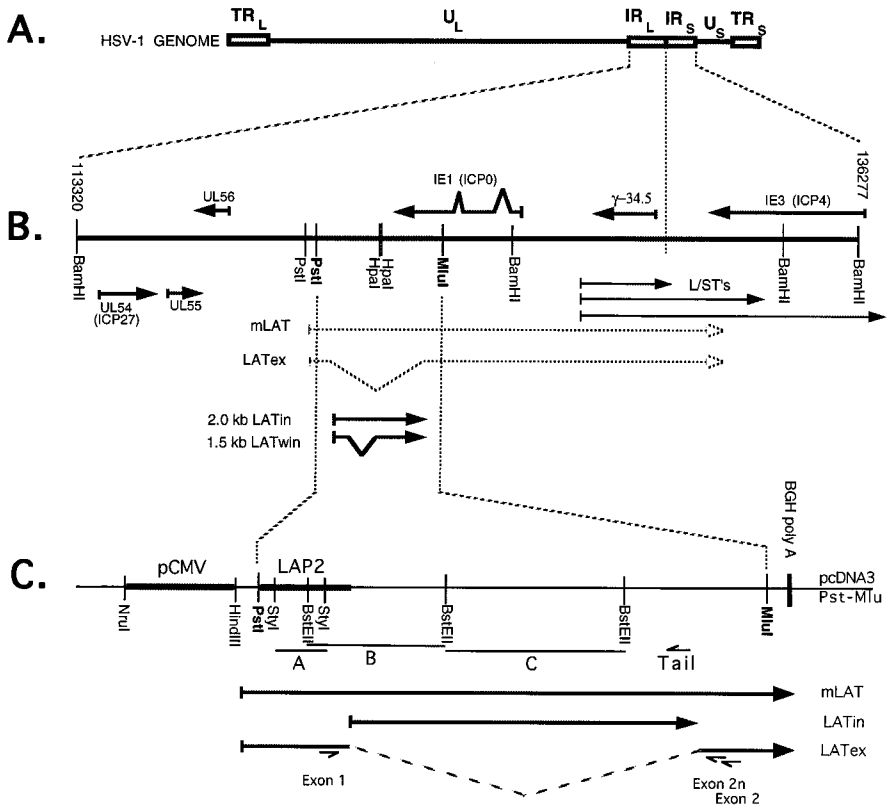

FIG. 1. The HSV-1 LATs. (A) Linear map of the HSV-1 genome. The 152-kb double-stranded DNA genome with its unique long region (U_L) and a unique short region (U_S) bounded by inverted repeat elements is represented. The terminal repeat long (TR_L) and the internal repeat long (IR_L) flank the unique long region, and the terminal repeat short (TR_S) and the internal repeat short (IR_S) flank the unique short region. (B) The LAT region of the HSV-1 genome. The LAT region of the genome is enlarged to show different transcripts that map to this region (LATs, L/STs, ICP0, γ34.5, ICP4, UL54, UL55, and UL56). The relevant restriction enzyme sites are shown. The putative 8.5-kb primary transcript mLAT and the postulated spliced exons of mLAT, LATex, are represented by dotted arrows. (C) The LAT minigene expression vector. A 2.8-kb fragment of the mLAT gene encompassing the 2.0-kb LAT region was cloned into a eukaryotic expression vector containing the CMV IE gene promoter to generate plasmid pcDNA3 PstI-MluI (pLAT). The CMV IE promoter and the postulated LAP2 promoter region are represented by boldface lines. Transcripts expected if the 2.0-kb LAT is an intron are depicted. The DNA fragments used as probes are shown: A, 0.4-kb *Sty*I-*Sty*I; B, 0.9-kb *Bst*EII-*Bst*EII; and C, 1.0-kb *Bst*EII-*Bst*EII. Positions of the oligonucleotide Tail and PCR primers specific for exon 1 and exon 2 of mLAT are indicated. BGH, bovine growth hormone.

HSV-1 and HSV-2 strains suggests that this RNA may have additional undiscovered functions.

In this report, we present evidence that the 2.0-kb LAT is an intron (LATin) accumulating after splicing of the mLAT pre-mRNA both in cells transfected with a LAT minigene expression vector and in cells infected with HSV-1. Additionally, we report the detection of the spliced exons of mLAT (LATex) in transiently transfected cells by Northern hybridization and in infected cells by reverse transcription-PCR (RT-PCR). In transfected cells, the 2.0-kb LAT intron localizes to the cytoplasm, demonstrating that LAT localization is not dependent on viral gene expression. We have mapped the branch point of the 2.0-kb LAT intron expressed by the LAT minigene vector and HSV-1 by using a novel RT-PCR technique. We have identified the branch point as a guanosine, the first identified example of an intron using this nucleotide as a branch point in vivo. These data imply that the stability of the 2.0-kb LATin may relate to its unique branch point.

MATERIALS AND METHODS

Plasmids. The *Pst*I-*Mlu*I restriction fragment of HSV-1 strain F DNA encompassing the 2.0-kb LAT coding region was previously subcloned into the vector pGEM4Z (Promega) to yield pGEM4Z PstI-MluI (56). The HSV-1 DNA was removed from this plasmid by digestion with restriction enzymes *Eco*RI and *Hin*dIII and cloned into the eukaryotic expression vector pcDNA3 (Invitrogen) at these sites to generate plasmid pcDNA3 PstI-MluI (Fig. 1C). pcDNA3 expresses the neomycin resistance gene and contains a polylinker region flanked by the cytomegalovirus (CMV) immediate-early (IE) gene promoter and the bovine growth hormone gene polyadenylation sequence. Plasmid pcDNA3 PstI-MluI ΔCMV was generated by deletion of the *Nru*I-*Hin*dIII restriction fragment containing the CMV IE promoter from the parent plasmid pcDNA3 PstI-MluI (Fig. 1C).

Cell culture and HSV-1 preparation. CV-1 cells were propagated and maintained in Eagle's minimum essential medium supplemented with 5% calf serum at 37°C with 5% CO_2. Subconfluent CV-1 monolayers were infected with 1 PFU of HSV-1 strain F or strain 17+ per cell. Virus was concentrated from the medium as described previously (17, 18). Virus titers were assayed on CV-1 and BHK cells.

Cell culture and transfection. COS-1 cells were maintained and propagated in Iscove's media supplemented with 5% calf serum at 37°C with 5% CO_2. Cells (7.5×10^5 to 1.5×10^6) were seeded in 100-mm-diameter plates and grown overnight. Monolayers at approximately 70% confluence were transfected by the calcium phosphate precipitation method with 20 μg of plasmid DNA in 125 mM $CaCl_2$, 25 mM HEPES–140 mM NaCl–0.75 mM Na_2HPO_4 (pH 7.05) (4). After 16 h of incubation, cells were washed with phosphate-buffered saline (PBS), shocked with 15% glycerol in PBS for 1 to 2 min, washed with PBS, and refed with fresh medium. At 44 to 48 h posttransfection, cells were harvested for RNA isolation as described below.

Cell fractionation, RNA extraction, and Northern hybridization. Transfected cells to be fractionated were washed with PBS and harvested in reticulocyte standard buffer (RSB; 10 mM Tris-HCl [pH 7.4], 10 mM NaCl, 3 mM $MgCl_2$) at 4°C. After a 10-min incubation at 4°C, cells were lysed with the addition of 0.1 volume of 10% Nonidet P-40 and subsequent vortexing. Nuclei were isolated by centrifugation through a sucrose cushion (0.33 M sucrose in RSB). Nuclear RNA

was isolated from the pellet, and cytoplasmic RNA was isolated from the supernatant (44).

RNA was isolated by the guanidium thiocyanate extraction-cesium chloride centrifugation procedure of Chirgwin et al. (12) as modified by Spivack and Fraser (54). Total RNA and nuclear RNA were isolated by homogenization of cells or tissue, or nuclei, in guanidium thiocyanate solution (4 M guanidium thiocyanate, 0.5% sodium lauroylsarcosine, 100 mM β-mercaptoethanol, 25 mM sodium citrate [pH 7.0], 0.1% antifoam A [Sigma Chemical Co.]) for 20 s with a cell disrupter (Brinkmann Instruments Inc.). To isolate cytoplasmic RNA, 0.1 volume of extraction buffer (1 M β-mercaptoethanol, 0.25 M sodium citrate, 5% sodium lauryl sarcosinate) and guanidium thiocyanate (to a final concentration of 4 M) were added to cytoplasmic extracts in RSB. COS-1 or CV-1 cells were infected at 1 PFU per cell and harvested at 16 to 20 h postinfection for total RNA.

Northern analysis was performed as previously described (44, 54), with some modifications. Ten-microgram aliquots of glyoxalated RNA were electrophoresed on 1.2% agarose gels, vacuum blotted to a nylon membrane (GeneScreen Plus; NEN), and UV cross-linked. DNA probes were subfragments of the HSV-1 strain F *Bam*HI fragment B (46). *Bam*HI-B subfragments were generated by restriction digestion, isolated by gel electrophoresis, and purified. The positions of these fragments relative to the LAT locus are indicated in Fig. 1C. RNAs were hybridized to heat-denatured ^{32}P-labeled nick-translated DNA probes overnight (54). Membranes were washed twice each in 1×, 0.5×, and 0.1× SSPE (1× SSPE is 180 mM NaCl, 10 mM monobasic sodium phosphate [pH 7.7], and 1 mM EDTA) with 1% sodium dodecyl sulfate for 20 min at 65°C (44). The oligonucleotide Tail (5′ CAAGAGGAAACCTCCCTCGGCCCCCGCGCTGCTTCTG GGCCGCGGGGGCCGAGGAAGTGTG 3′) was also used as a probe for Northern hybridization. The position of this oligonucleotide relative to the 2.0-kb LAT RNA is shown in Fig. 1C. The ^{32}P end-labeled oligonucleotide was hybridized to RNAs as outlined above, and membranes were washed as indicated except that washes were carried out at 50°C. All membranes were exposed for autoradiography to Reflection film (Dupont) with an intensifying screen (DuPont).

RT-PCR and sequencing of the LAT splice junction. cDNA was synthesized from total RNA of transfected or infected cells with the Superscript preamplification system (Gibco-BRL) as instructed by the manufacturer. Primers Exon 1 (5′ GCTCCATCGCCTTTCCTGTTC 3′) and Exon 2 (5′ TGACGTCCTCGG CGGCCTC 3′) were used to amplify LATex cDNA from transiently transfected cells. PCR amplifications were done with 2.5 U of *Taq* 2000 polymerase (Stratagene) in 60 mM Tris-HCl–15 mM $(NH_4)_2SO_4$–1.5 mM $MgCl_2$ at pH 10.0 (Invitrogen) with 1 μM each primer and 1 mM deoxynucleotides. Amplifications consisted of 35 cycles of denaturation at 94°C for 1 min, annealing at 58°C for 1 min, and extension at 72°C for 1 min. PCR products were resolved by agarose gel electrophoresis and visualized by ethidium bromide staining. The PCR products were cloned into a TA cloning vector (Invitrogen) prior to sequencing by an ABI Prism cycle sequencer.

PCR with primers Exon 1 and Exon 2n (5′ TCCTCTGCCTCTTCCTCCTCG 3′) was used to amplify LATex cDNA from HSV-1-infected cells (due to mispriming with primer Exon 2 on viral cDNA). PCR amplifications were done with 2.5 U of *Taq* polymerase (Fisher) in 60 mM Tris-HCl–15 mM $(NH_4)_2SO_4$–2.0 mM $MgCl_2$ at pH 9.5 (Invitrogen) with 1 μM each primer and 1 mM deoxynucleotides. Amplifications consisted of 35 cycles of denaturation at 94°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for 1 min. PCR products were resolved by agarose gel electrophoresis and visualized by ethidium bromide staining. The 245-bp PCR product was cloned into a TA cloning vector (Invitrogen) prior to sequencing.

Mapping of the 2.0-kb LATin branch point by RT-PCR. DNase-treated total RNA isolated from transfected or infected cells or infected ganglia was transcribed into cDNA with the Superscript preamplification system (Gibco-BRL) as instructed by the manufacturer. cDNA was synthesized from transfected cell RNA with primer Branch (the sequences of primers for mapping the LAT branch point are shown in Table 1) at 50°C for amplifications with primers E and End and at 42°C for amplifications with primer E and primer A, B, C, or D. Primer Branch is complementary to the 5′ end of the 2.0-kb LAT RNA, with the addition of three degenerate nucleotides at the 3′ end of the primer. PCR amplifications with primers E and D, E and C, E and B, or E and A were done with 2.5 U of *Taq* polymerase (Fisher) in 60 mM Tris-HCl–15 mM $(NH_4)_2SO_4$–2.0 mM $MgCl_2$ at pH 10.0 (Invitrogen) with 1 μM each primer and 1 mM deoxynucleotides. Amplifications consisted of 35 cycles of denaturation at 94°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for 1 min. PCR amplifications for the LAT branch point were done by using primers E and End with 2.5 U of *Taq* polymerase (Fisher) in 60 mM Tris-HCl–15 mM $(NH_4)_2SO_4$–1.5 mM $MgCl_2$ at pH 9.0 (Invitrogen) with 1 μM each primer and 1 mM deoxynucleotides. Amplifications consisted of 35 cycles of denaturation at 94°C for 1 min, annealing at 71°C for 1 min, with the temperature dropping 1°C every 4 cycles and the final 15 cycles at 66°C, and extension at 72°C for 1 min. PCR products were resolved by agarose gel electrophoresis and visualized by ethidium bromide staining. The PCR products were cloned into a TA cloning vector (Invitrogen) and subsequently sequenced with an ABI Prism cycle sequencer.

cDNA was synthesized from infected cell RNA with primer Branch, BranchG, or End or random primers at 50°C. PCR amplification of the viral LAT branch point was done with primers E and End. Amplifications were done with 2.5 U of *Taq* polymerase (Fisher) in 60 mM Tris-HCl–15 mM $(NH_4)_2SO_4$–1.5 mM $MgCl_2$ at pH 9.0 (Invitrogen) with 1 μM each primer and 1 mM deoxynucleotides. Amplifications consisted of 35 cycles of denaturation at 94°C for 1 min, annealing at 71°C for 1 min, with the temperature dropping 1°C every 4 cycles and the final 15 cycles at 66°C, and extension at 72°C for 1 min. PCR products were resolved by agarose gel electrophoresis and visualized by ethidium bromide staining. The 300- and 250-bp PCR products were cloned into a TA cloning vector (Invitrogen) and subsequently sequenced with an ABI Prism cycle sequencer.

TABLE 1. HSV-specific oligonucleotides for intron branch point mapping

Primer[a]	Sequence[b]	Location[b]
A	CTGGGAGGGAGACAAGAGGAAA	121418–121397
B	CGAGGAAGTGTGCCCGGAAGAC	121357–121336
C	AGACGCGCCACGCGGAGACTTC	121339–121318
D	GGAGACTTCCGGGGCCGTCC	121326–131307
E	GGGGCATCACTGTGTTACCC	121058–121077
Branch	AGAAGCAGGTGTCTAACCTACNNN	119484–119464
End	AGAAGCAGGTGTCTAACCTAC	119484–119464
BranchG	AGAAGCAGGTGTCTAACCTACCCG	119484–119464 and 121344–121342

[a] As indicated in Fig. 5.
[b] Based on published sequences (40).

RESULTS

The 2.0-kb LAT is an intron spliced from mLAT in transient transfection assays and viral infection. To examine the expression of the HSV-1 2.0-kb LAT in transient transfection assays, we constructed a LAT minigene expression vector. A detailed map of the LAT locus and the expression construct is depicted in Fig. 1. A 2.8-kb fragment of the mLAT gene encompassing the 2.0-kb LAT RNA was cloned into the polylinker site of vector pcDNA3 (Invitrogen), between the CMV IE gene promoter and the bovine growth hormone gene polyadenylation signal. The vector was transfected into COS-1 cells, and at 44 to 48 h posttransfection, total RNA, nuclear RNA, or cytoplasmic RNA was isolated and analyzed by Northern hybridization. Additionally, total RNA was isolated from COS-1 cells and from COS-1 cells infected for 16 h with HSV-1 strain 17+ at a multiplicity of infection of 1. The same Northern blot was hybridized to three different probes (Fig. 2), mapping 5′ to the 2.0-kb LAT (probe A), overlapping the 2.0-kb LAT as well as the region upstream (probe B), and within the 2.0-kb LAT (probe C) (Fig. 1C). Northern blots were also routinely probed with plasmid pBSS (27), containing the human rRNA 5′ external transcribed spacer, to verify that cytoplasmic RNA was not contaminated with nuclear RNA (data not shown) (44).

Probe B detects the expression of three transcripts, 3.4, 2.0, and 1.4 kb in size (Fig. 2, lanes 3 to 5). The 3.4-kb RNA is mostly nuclear, the 1.4-kb RNA is mainly cytoplasmic, and the 2.0-kb RNA is similarly abundant in equal quantities of nuclear and cytoplasmic RNA (Fig. 2; compare lanes 4 and 5). The size and cellular distribution of these transcripts suggested that the 3.4-kb RNA may be the truncated mLAT primary transcript and that the 1.4-kb RNA may be the spliced exons of this transcript.

To clarify the relationship between these three transcripts, the RNAs were mapped by reprobing this blot with two other probes, one specific for sequences within the 2.0-kb LAT (probe C) and one specific for sequences upstream of the 2.0-kb LAT (probe A). Two transcripts, the 3.4- and 2.0-kb RNAs, were detected with probe C (Fig. 2, lanes 8 to 10). In virally infected cells, this probe detects the 2.0-kb LAT as well as the antisense ICP0 transcript (Fig. 2, lane 7). Probe A also

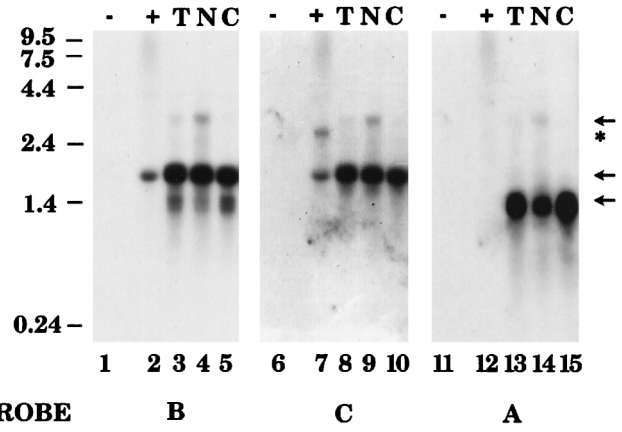

FIG. 2. Expression of truncated mLAT, LATin, and LATex RNAs from a minigene expression plasmid. COS-1 cells were transfected by calcium phosphate with a LAT minigene expression plasmid (pLAT), and total RNA was prepared from cells (T) or from nuclear (N) and cytoplasmic (C) fractions at 48 h post-transfection. Total RNA was also isolated from mock-infected (−) or HSV-1 strain 17+-infected (+) COS-1 cells at 16 h postinfection. RNAs were analyzed by Northern hybridization. The same Northern blot was probed with three different probes, A, B, and C (Fig. 1). The positions of RNA size markers (in kilobases) are indicated. The three LAT RNAs, 3.4, 2.0, and 1.4 kb, are indicated by arrows; the asterisk indicates the ICP0 mRNA.

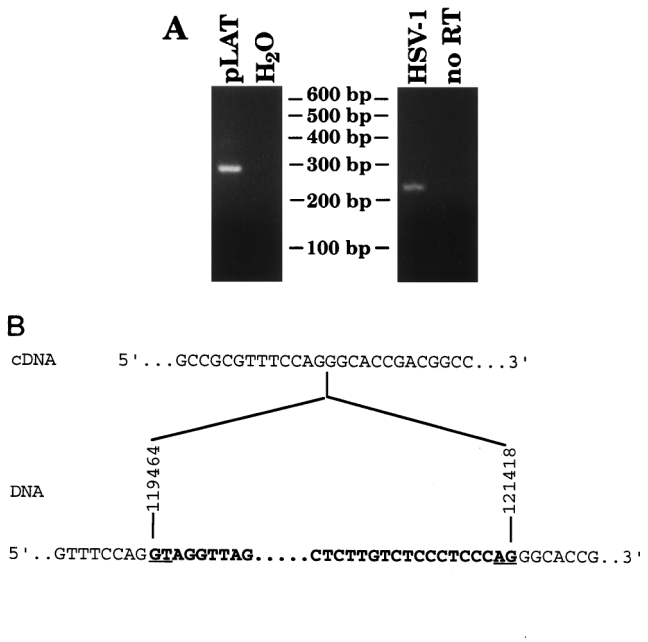

FIG. 3. Identification of the LATin splice donor and acceptor sites. Total RNA of COS-1 cells transfected with a LAT minigene expression plasmid or CV-1 cells infected with HSV-1 was reverse transcribed into cDNA. PCR amplification of transfected cell cDNA (pLAT) or H_2O was done with primers Exon 1 and Exon 2 (Fig. 1C). PCR amplification of HSV-1 cDNA or RNA from infected cells (no RT) was done with primers Exon 1 and Exon 2n (Fig. 1C). The PCR products were electrophoretically separated on a 2% agarose gel (A). The sequence at the splice junction of the LATex cDNA is shown in panel B, with the corresponding sequences of the mLAT DNA. Intron sequences are in boldface, and the splice donor and acceptor sites are underlined. Consensus mammalian sequences for splicing are also shown.

detects only two transcripts, the 3.4- and 1.4-kb RNAs (Fig. 2, lanes 13 to 15). These results are consistent with the 3.4-kb RNA being the primary transcript from which the 2.0-kb LATin and the 1.4-kb LATex are spliced.

To verify this hypothesis, we demonstrated by RT-PCR and sequencing that these exons are spliced together and identified the splice donor and acceptor sites of the 2.0-kb LAT. Total RNA from COS-1 cells transiently transfected with pcDNA3 PstI-MluI was amplified by RT-PCR with primers Exon 1 and Exon 2, specific for the putative mLAT exons (Fig. 1C). PCR products were separated by agarose gel electrophoresis (Fig. 3A). The major 275-bp PCR product was cloned into a TA cloning vector, and several clones were sequenced. To determine if splicing of mLAT was also occurring in virally infected cells and if the same splice donor and acceptor sites of mLAT were recognized, total RNA from CV-1 cells infected overnight with HSV-1 strain F was also amplified by RT-PCR with primers specific for the putative mLAT exons. In this case, primers Exon 1 and Exon 2n were used for PCR due to mispriming of primer Exon 2 on viral cDNA. PCR products were separated by agarose gel electrophoresis (Fig. 3A), and the 245-bp PCR product was cloned into a TA cloning vector prior to sequencing. The splice junction of the LAT exons identified in transfected and infected cells was the same. The sequence at the splice junction of the LATex is shown in Fig. 3B, with the corresponding sequences of the mLAT DNA. The GT of the splice donor site at the 5′ end of the 2.0-kb LATin begins at nucleotide 119464, and the AG of the splice acceptor site of the intron ends at nucleotide 121418 of the HSV-1 genome (40). This is in agreement with the previously published splice donor and acceptor sites of the 2.0-kb LATin (23). A minor 400-bp PCR product was also occasionally amplified by RT-PCR from cells transfected with the LAT expression vector. The splice donor and acceptor sites of the 2.0-kb LAT at nucleotides 119464 and 121418 were used, but this product contained an additional exon corresponding to nucleotides 119611 to 119739 of the LAT gene. The sequence at 119739 is the splice donor site of the inner intron of the 2.0-kb LAT, which was thought to be neuron specific (56). We attribute these splicing events to overexpression of the LAT in transfected cells relative to virally infected cells (Fig. 2) and note that these products are not amplified from virally infected cells (Fig. 3A).

The 2.0-kb LAT as a primary transcript expressed from the LAP2 promoter could not be detected. The sequences for the LAP2 promoter element lie within the LAT sequences present in our minigene vector. To assess the contribution of this promoter element to the synthesis of the 2.0-kb LAT in this system, we deleted the CMV IE promoter element from this vector, generating plasmid pcDNA3 PstI-MluI ΔCMV (Fig. 1C). Total RNA was isolated from COS-1 cells transfected with the wild-type expression vector pcDNA3 PstI-MluI or the CMV promoter deletion mutant pcDNA3 PstI-MluI ΔCMV and analyzed by Northern hybridization with probe C (Fig. 4). The amount of 2.0-kb LAT detected in cells transfected with plasmid pcDNA3 PstI-MluI ΔCMV was low (Fig. 4, lane 1) compared to the level detected in cells transfected with the wild-type plasmid (Fig. 4, lane 3). By PhosphorImager (Molecular Dynamics) analysis, the ratio of LAT expression of wild-type expression plasmid to this mutant was approximately 500:1 in this experiment and was consistently greater than 100:1 in different experiments. Thus, LAP2 promoter sequences do not contribute significantly to the levels of 2.0-kb LAT observed in this system.

To determine whether the 2.0-kb LAT synthesized from the

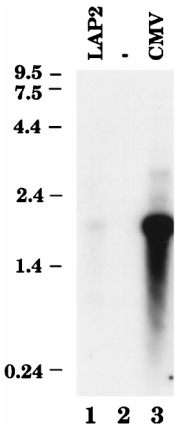

FIG. 4. Expression of the 2.0-kb LAT from LAP2. COS-1 cells were transfected with the wild-type LAT minigene expression plasmid (CMV) or the LAT expression plasmid deleted for the CMV IE promoter but still retaining the LAP2 promoter sequences (LAP2). Total RNA was prepared from mock-transfected cells (−) and cells 44 h posttransfection. RNAs were analyzed by Northern hybridization with probe C (Fig. 1). The positions of RNA size markers (in kilobases) are indicated.

LAP2 promoter is a primary transcript, we assayed for the presence and abundance of the LATex by semiquantitative RT-PCR using primers Exon 1 and Exon 2. Semiquantitative PCR was done as previously described (16), using serial dilutions of cDNA and selecting a number of cycles within the linear range of template amplification. LATex cDNA was amplified from cells transfected with the wild-type expression vector as well as the vector deleted for the CMV IE promoter element. The ratio of LATex expressed by wild-type expression plasmid to that expressed by the mutant was approximately 200:1 (data not shown). The difference in expression of LATex assayed by RT-PCR corresponds to the difference in expression of LATin assayed by Northern hybridization. This finding suggests that LAP2 promoter-driven transcription of the 2.0-kb LAT may not initiate at the splice donor site in our LAT minigene vector.

The branch point of the 2.0-kb LAT intron is a guanosine. It was recently reported that a portion of the 2.0-kb LAT has a nonlinear configuration, most likely a lariat (67). Because the 2.0-kb LAT is an intron, some of the molecules may be in a lariat conformation. As such, the 2.0-kb LAT should be branched. We used a novel RT-PCR technique to map the branch point of this intron. DNase-treated total RNA isolated from COS-1 cells transfected with plasmid pcDNA3 PstI-MluI was the substrate for cDNA synthesis. Primer Branch used for cDNA synthesis is complementary to the 5′ end of the 2.0-kb LATin, with three degenerate nucleotides at the 3′ end to extend the area of hybridization past the putative branched nucleotide (Fig. 5; Table 1). cDNA was synthesized at 50°C for identification of the LAT intron branch point. Due to limited mispriming of this degenerate primer during cDNA synthesis at 42°C, we were able to amplify a series of PCR products from the 3′ end of the 2.0-kb LATin by using primers E and D, E and C, E and B, or E and A. Primer D begins 91 bases upstream of the 3′ end of the 2.0-kb LATin, and primer A begins at the splice acceptor site at the 3′ end of the 2.0-kb LATin (Table 1). PCR products were amplified from the 3′ end of the LATin with these primer sets for size comparison to branch point PCR products (see below). Most intron branch points lie within 100 bases upstream of the splice acceptor site (42), and so we anticipated that the branch point PCR product

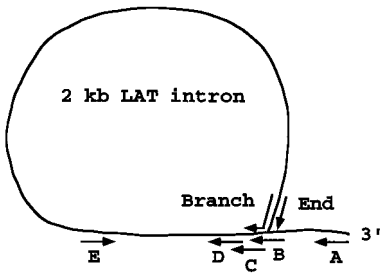

FIG. 5. Mapping the 2.0-kb LATin branch point. A schematic diagram of the 2.0-kb LATin in a lariat configuration is shown with the primers used for cDNA synthesis and PCR amplification of the branch point sequence. Sequences of the primers and their locations in the HSV-1 genome are presented in Table 1.

would probably be larger than the PCR product amplified by primers E and D.

PCR amplification of the branch point was accomplished with a primer complementary to the 5′ end of the 2.0-kb LAT (End) and an upstream primer (E) (Fig. 5; Table 1). The PCR products were separated by agarose gel electrophoresis (Fig. 6A). Two PCR products of 250 and 300 bp are visible. These PCR products were cloned into a TA cloning vector, and several clones were sequenced. The nucleotide 5′ to the splice donor site of the intron in the 300-bp PCR product was the G at position 121344 of the HSV-1 genome (40). In the 250-bp PCR product, the nucleotide 5′ to the splice donor site of the 2.0-kb LAT was the G at position 121283. These sequences are shown in Fig. 7B and C, respectively, hybridized to the primer Branch. Sequences 3′ to the G at position 121344 show no complementarity to primer Branch; however, the four nucleotides 3′ to the G at position 121283 are complementary to the four nucleotides in primer Branch 5′ to the 3 degenerate bases at its 3′ end (Fig. 7C). It is likely that the 250-bp PCR product results from a misprime at the level of cDNA synthesis, al-

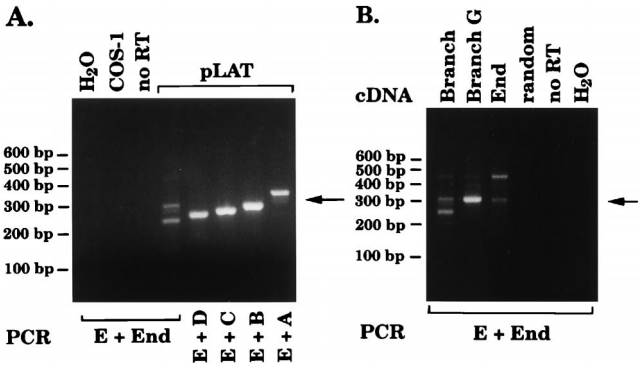

FIG. 6. RT-PCR mapping of the 2.0-kb LATin branch point. (A) Amplification of the 2.0-kb LATin branch point from transiently transfected cells. cDNA was synthesized from total RNA of untransfected COS-1 cells (COS-1) or cells transfected with plasmid pcDNA3 PstI-MluI (pLAT) by using primer Branch. PCR amplification of cDNAs, RNA from transfected cells (no RT), and H_2O was done with the indicated primer sets. PCR products were electrophoretically separated on a 2% agarose gel and stained with ethidium bromide. Positions of DNA molecular weight markers are shown. The arrow indicates the PCR product containing the branch point of the 2.0-kb LATin. (B) Amplification of the 2.0-kb LATin branch point from virally infected cells. cDNA was synthesized from total RNA of CV-1 cells infected with HSV-1 strain F by using primer Branch, BranchG, or End or random primers. PCR amplification of cDNAs, RNA from infected cells (no RT), and H_2O was done with primers E and End. PCR products were resolved on a 2% agarose gel and visualized with ethidium bromide staining. Positions of molecular weight markers are shown. The arrow indicates the PCR product containing the LATin branch point.

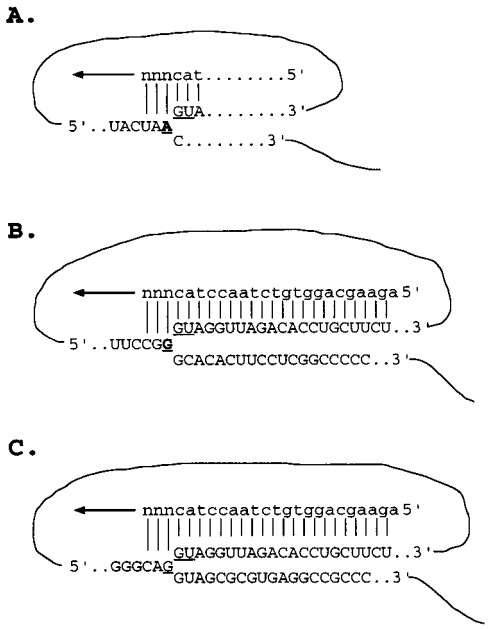

FIG. 7. RT-PCR identification of branch points. (A) Identification of the branch point of a yeast consensus branch site sequence (UACUAAC). A schematic diagram of a primer hybridized to a yeast consensus branch site (UACUAAC) of an intron is shown. The intron is depicted in a lariat conformation. The branched nucleotide A is in boldface and underlined. The underlined splice donor site, GU, at the 5′ end of the intron is linked to the branch point A by a 2′-5′ phosphate linkage. The primer used for cDNA synthesis is shown in lowercase hybridized to the 5′ end of the intron containing the splice donor sequence. The three degenerate nucleotides (n) of the primer for cDNA synthesis are lacking in the primer for PCR. The arrow indicates the direction of synthesis. (B and C) Identification of the branch point of the HSV-1 LATin. LATin is depicted in a lariat conformation. The primer Branch used for cDNA synthesis is shown in lowercase hybridized to the 5′ end of LATin containing the splice donor sequence. The underlined splice donor site, GU, is linked by a 2′-5′ phosphate linkage to the underlined branch point nucleotide. The arrow indicates the direction of cDNA synthesis. The three degenerate nucleotides (n) at the 3′ end of primer Branch for cDNA synthesis are lacking in primer End used for PCR. The branch point G at position 121344 (40) is in boldface and underlined (B). The G at position 121283 was also identified as a potential branch point sequence with this technique and is underlined in (C). The sequence identified in panel C may be a misprime at the level of cDNA synthesis due to complementarity of the surrounding sequences to primer Branch.

though we cannot rule out that the G at position 121283 is another LATin branch point with this technique.

As a control for determining which nucleotide within the branch point region was the nucleotide which had formed a 2′-5′ phosphodiester bond with the 5′ splice donor site, RT-PCR was also done on introns by using a yeast consensus branch point sequence, 5′ UACUAAC 3′. The underlined A is utilized as the branch site in this sequence (42). Sequenced RT-PCR products revealed that the nucleotide 5′ to the splice donor site was the branched nucleotide (Fig. 7A). By analogy, the G at position 121344, the nucleotide 5′ to the splice donor site in the LAT branch point PCR product, is the branched nucleotide of this intron (Fig. 7B).

To identify the LATin branch point from virally infected cells, cDNA was synthesized from total RNA of CV-1 cells infected with HSV-1 strain F, using primer Branch, BranchG, or End or random primers. In primer BranchG, the three degenerate nucleotides at the 3′ end of the primer are replaced with the bases complementary to the branch point sequence of LATin at position 121344 identified from the LAT expression plasmid (Table 1; Fig. 7B). The LATin branch point was amplified by PCR with primers E and End. The PCR products were separated by agarose gel electrophoresis (Fig. 6B). Again,

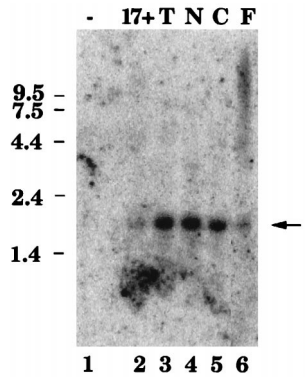

FIG. 8. Detection of the tail of the 2.0-kb LAT. The Northern blot in Fig. 2 is shown hybridized with probe Tail. COS-1 cells were transfected by calcium phosphate with a LAT minigene expression plasmid, and total RNA was prepared from cells (T) or from nuclear (N) and cytoplasmic (C) fractions at 48 h posttransfection. Total RNA was also isolated from mock-infected (−) or HSV-1 strain 17+-infected (17+) COS-1 cells at 16 h postinfection. An additional lane of this blot, of total RNA isolated from CV-1 cells infected with HSV-1 strain F at 16 h postinfection (F), is shown. The positions of the RNA size markers (in kilobases) are indicated. The arrow indicates the position of the 2.0-kb LAT.

two major PCR products are visible at 250 and 300 bp. The PCR products were cloned into a TA cloning vector (Invitrogen), and a number of clones were sequenced. The sequences of these RT-PCR products were the same as those amplified from cells transiently transfected with the LAT expression plasmid. Therefore, the G at position 121344 is also the branch point of the LATin in virally infected cells.

We have also amplified the LATin branch point from RNA isolated from the trigeminal ganglia of latently infected mice. BALB/c mice were infected as previously described (54) with 10^4 PFU of HSV-1 strain 17+ per eye. After several months, mice were sacrificed and trigeminal ganglia were removed for extraction of RNA. The RT-PCR products obtained (data not shown) were identical to those amplified from transiently transfected cells and virally infected tissue culture cells (Fig. 6). The sequences of several branch point PCR clones were identical to those amplified from transfected and infected tissue culture cells (Fig. 7B and C). Thus, the LATin branch point in latently infected neurons is also the guanosine at position 121344.

The tail of the 2.0-kb LAT is detectable in transfected and infected cells. To determine if the tail of the 2.0-kb LAT downstream of the branch point at nucleotide 121344 was present on the 2.0-kb LAT, we reprobed the blot shown in Fig. 2 with probe Tail (Fig. 1C). An additional lane of this blot, of total RNA isolated from CV-1 cells infected with HSV-1 strain F at 16 h postinfection, is shown. The results are shown in Fig. 8. Probe Tail detects the 2.0-kb LAT in total, nuclear, and cytoplasmic fractions of cells transfected with the LAT minigene expression plasmid (lanes 3 to 5). Probe Tail also detects the 2.0-kb LAT from COS-1 cells infected with HSV-1 strain 17+ or CV-1 cells infected with HSV-1 strain F (lanes 2 and 6). These data indicate that the tail of the 2.0-kb LAT is present on at least a proportion of the 2.0-kb LAT molecules.

DISCUSSION

The data presented in this report provide evidence supporting the hypothesis that the 2.0-kb LAT is an intron spliced from the pre-mRNA mLAT. Previously, Farrell et al. (23) presented strong evidence in support of this hypothesis by demonstrating that the 2.0-kb LAT could be spliced out of a β-galactosidase transcript containing LAT sequences in tran-

sient transfections. This is the first report demonstrating the splicing of the 2.0-kb intron (LATin) from within the context of mLAT and the observation of the spliced exons of mLAT (LATex).

Detection of the LATex during transfection and HSV-1 infection confirms that the 2.0-kb LAT is an intron. Transient transfection of our LAT expression plasmid generated three LAT-specific transcripts 3.4, 2.0, and 1.4 kb in size. We did not expect that the 1.4-kb transcript was the 1.5-kb LAT (Fig. 1B), because this transcript is observed in latency only in small amounts (54, 55, 62). In addition, the 1.4-kb RNA hybridizes with probe A, which does not overlap the 1.5-kb RNA observed during HSV-1 latency. The 1.5-kb LAT appears to arise by splicing of a 500-base intron from the 2.0-kb LATin (56) and therefore may be a twintron (LATwin) (13).

The nuclear localization of the 3.4-kb RNA and the mostly cytoplasmic localization of the 1.4-kb RNA, as well as the sizes of these transcripts, suggested that the 3.4-kb RNA is a truncated mLAT pre-mRNA and the 1.4-kb RNA is the LATex. Mapping of the transcripts by Northern hybridization (Fig. 2) and RT-PCR and partial sequencing of the 1.4-kb LATex (Fig. 3) confirmed this hypothesis. The splice donor and splice acceptor sites of the 2.0-kb LATin, as determined by sequencing of the splice junction of LATex cDNA, are the expected consensus sequences CAG/GT and NYAG/G, respectively. Immediately upstream of the splice acceptor AG is a stretch of 19 pyrimidines out of 20 nucleotides comprising the polypyrimidine tract associated with the splice acceptor site. These sites are the same as those identified by the Farrell et al. (23), used in splicing the 2.0-kb LATin from a β-galactosidase transcript in which the 2.0-kb LAT had been inserted. Significantly, although the viral mLAT pre-mRNA and LATex transcripts are not detectable by Northern hybridization, the same splice donor and acceptor sites of mLAT are recognized in productive viral infections (Fig. 3).

The mLAT pre-mRNA encodes a number of open reading frames (36). The splicing of the mLAT to yield the LATex mRNA generates a new open reading frame, which has previously been referred to as open reading frame B (ORFB) (36). ORFB is not well conserved between HSV-1 strains F and 17+ (36, 40, 66). The protein encoded by ORFB has not been detected in productively infected cells (36), although it may be expressed during HSV-1 latency or reactivation.

LATin localization to the cytoplasm is not dependent on expression of other viral genes. The cellular localization of the 2.0-kb LATin is unusual. In transiently transfected COS-1 cells, a significant amount of the 2.0-kb LATin is cytoplasmic (Fig. 2). Nevertheless, this corresponds to what is observed in HSV-1-infected CV-1 cells and murine brainstem cells (44). Additionally, the localization of the 2.0-kb LATin in transfected cells demonstrates that no viral gene products are required for cytoplasmic LATin localization. In latently infected neurons, the 2.0-kb LATin and the smaller 1.5-kb LATwin are localized in the nucleus (54, 58, 59). It is unclear what cellular mechanism mediates the localization of 2.0-kb LATin, but it is likely related to the dynamic state of the cell. The murine T-cell receptor β RNA introns $IVS1_{C\beta1}$ and $IVS1_{C\beta2}$ and the human T-cell receptor β RNA intron $IVS1_{C\beta1}$ have been described as relatively stable (47). However, unlike the LATin, the murine $IVS1_{C\beta1}$ intron is localized in the nucleus (47). The HSV-1 2.0-kb LATin is the only example now identified of a naturally occurring intron localizing to the cytoplasm in vivo.

The mechanism by which mRNAs and snRNAs are exported from the nucleus appears to involve RNA binding proteins (21). From transport studies of retroviral RNAs, it has been hypothesized that pre-mRNAs are retained in the nucleus due to bound splicing factors (9, 31, 38). Unspliced retroviral RNAs code for additional viral proteins and are able to overcome blocked nuclear export. For HIV RNAs, this mechanism involves binding to Rev (15). Like pre-mRNAs, introns are not normally exported from the nucleus. Introns may be retained in the nucleus by binding the same proteins involved in the nuclear retention of pre-mRNAs. Because no viral proteins are required for LATin nuclear export, it is unclear how the LATin would overcome this potential block. Alternatively, lack of intron nuclear export could be due to rapid debranching and degradation of introns. The LATin appears to be resistant to debranching enzymes and remarkably stable (50). As a result, LATin may be available for nuclear export by binding the same proteins involved in nuclear export of mRNA.

Generation of the 2.0-kb LAT from LAP2: intron or exon? Although the LAP2 sequence elements are present in our LAT expression vector pcDNA3 PstI-MluI, the promoter is not very active in transfected COS-1 cells (Fig. 4). The levels of 2.0-kb LAT produced by the pcDNA3 PstI-MluI ΔCMV vector were consistently at least 100-fold less than that observed for the wild-type expression plasmid pcDNA3 PstI-MluI (Fig. 4). If the 2.0-kb LAT is produced as a primary transcript from LAP2, this does not contribute significantly to the amount of LAT observed in this system. Thus, the majority of the 2.0-kb LAT produced by this vector is an intron accumulating after splicing.

The small amount of 2.0-kb LAT produced by the LAP2 promoter may also be an intron. RT-PCR amplification of LATex with primers Exon 1 and Exon 2 from cells transfected with pcDNA3 PstI-MluI ΔCMV suggests that transcription of the 2.0-kb LAT from LAP2 may initiate upstream of the splice donor site of the intron (data not shown). Primer extension analysis by Chen et al. (11) demonstrated that the 5′ ends of some transcripts initiating at LAP2 mapped upstream of the splice donor site of the 2.0-kb LATin. This is consistent with the LAP2 2.0-kb LAT also being an intron.

The 2.0-kb LAT intron branches at a guanosine. Notably absent near the 3′ end of the 2.0-kb LATin is the mammalian branch site consensus sequence, YNYURAC, with the branch point at the adenosine (42). The sequence UUCUAAC, a consensus branch site, is present at position +690 relative to the 5′ end of the 2.0-kb LATin. This sequence lies in the smaller 500-base intron within the 2.0-kb LATin and is presumably utilized as the branch point in the splicing of this intron to generate the 1.5-kb LATwin observed in latently infected neurons (56). The lack of a consensus branch point sequence and the unusual stability of the 2.0-kb LATin led us to postulate that the branch site of this intron would be unique.

Using a novel RT-PCR technique, we have identified the branch point of the 2.0-kb LATin, expressed in both transfected and infected tissue culture cells as well as latently infected neurons, as a guanosine. This is the first known example of an intron utilizing this nucleotide as a branchpoint in vivo. Additionally, the branch site of LATin, UUCCGGG, bears little homology to the consensus mammalian branch site sequence YNYURAC. By Northern hybridization, the tail of the 2.0-kb LAT downstream of the branch point is detectable on at least proportion of the 2.0-kb LAT in transfected and infected cells (Fig. 8). Because exonucleolytic degradation of RNA appears to follow debranching of lariat intermediates (10), it is likely that at least some of these molecules are in a lariat conformation.

The HSV-1 2.0-kb LAT and HSV-2 2.3-kb LAT, which overall have little homology (50%), are highly homologous at their 3′ ends (80%) (34, 39) due to overlap with the ICP0 mRNA coding region on the opposite strand (Fig. 1B). The 3′ ends of

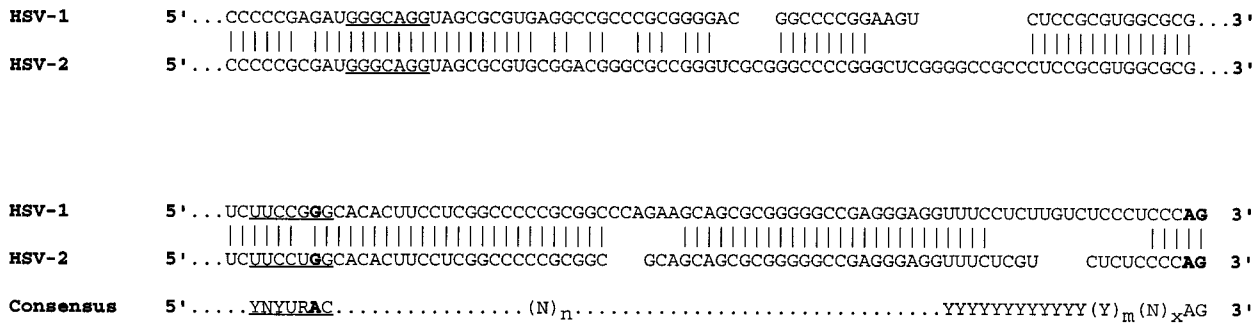

FIG. 9. Sequence homology between the 3′ ends of the HSV-1 and HSV-2 LAT introns. The sequences of the 3′ ends of the HSV-1 2.0-kb LAT introns and the HSV-2 2.3-kb LAT introns are shown with the consensus sequences for the mammalian branch point, polypyrimidine tract, and splice acceptor (AG) site. Identity is indicated by lines between the sequences. The potential HSV-1 2.0-kb LAT intron branch point sequences identified by RT-PCR are underlined in the HSV-1 and HSV-2 LAT intron sequences. The HSV-1 LAT intron branched nucleotide at position 121344 is in boldface. Spaces were introduced into the HSV-1 or HSV-2 sequences to preserve identity. n, m, and x indicate approximately 100, 12 to 20, and 0 to several nucleotides, respectively.

the HSV-1 and HSV-2 LATs are shown in Fig. 9. Interestingly, there is a base change between the two viruses within the branch site sequence of the HSV-1 LATin. The corresponding sequence in HSV-2, UUCuGGG, has less homology to the consensus mammalian branch point sequence YNYURAC than the HSV-1 branch site UUCCGGG. Thus, it is unclear whether this sequence in HSV-2 would be recognized as the branch point of this intron.

The branch point G of the HSV-1 LATin at position 121344 is complementary to the last base of the ICP0 codon 621, GCC, coding for alanine (39). In HSV-2, the corresponding G is complementary to the last base of the HSV-2 ICP0 codon 663, GCC for alanine (39). Changing the HSV-1 branch point guanosine to the usual branch point nucleotide adenosine would alter these codons to GCU, which still encodes alanine. Thus, there appears to be no selective pressure maintaining this nucleotide as a G. Indeed, the base preceding this G is altered in HSV-2, yet the corresponding ICP0 codons in both HSV-1 (CGG) and HSV-2 (AGG) encode the amino acid arginine.

Splicing of the 2.0-kb LAT is unusual during the limited splicing of HSV-1 infection. Of the more than 70 genes encoded by HSV-1, only 6 generate transcripts that are spliced to any great extent (14, 25, 45, 63, 64; reviewed in reference 28), although nominal splicing of a few other viral transcripts has been reported (7). These six genes include the IE genes ICP0, ICP22, and ICP47, the early gene UL15, the late gene glycoprotein C, and LAT. The IE protein ICP27 disrupts splicing in infected cells at early times after infection, presumably to favor viral gene expression (51). Unspliced cellular and viral transcripts accumulate in the nucleus at late times postinfection (28). Consistent with this finding, only unspliced glycoprotein C mRNA codes for protein (25). Splicing of LAT is unusual; there is no accumulation of unspliced mLAT pre-mRNA detectable by Northern analysis in productively infected cells at late times postinfection. It is possible that the splicing of LAT is independent of some elements of the splicing machinery required by other transcripts, or that LAT requires other components for splicing. Recently, a new class of rare genes containing AT-AC introns has been described. These introns are spliced by several newly identified snRNPs which form an alternate spliceosome (60, 61). By sequence analysis, the splicing of LAT appears to require the normal spliceosome, but the branch site of the LATin bears little homology to consensus mammalian branch point sequences and may therefore not be recognized by the U2 snRNP.

Is LAT intron stability mediated through a guanosine branchpoint? Some introns use nonadenosine branch points in vivo (reviewed in reference 48). The second intron of the human immunodeficiency virus type 1 (HIV-1) *tat*/*rev* pre-mRNA branches mainly at a uracil (22). Selection of this branch point is suboptimal, allowing some messages to remain unspliced, which results in the generation of additional viral proteins. Two cellular genes, the human growth hormone gene and the calcitonin/calcitonin gene-related peptide I gene, also have introns with unusual branch points. The first intron of the human growth hormone pre-mRNA branches at a cytosine (29). A uracil is the branch point of the third intron of the calcitonin/calcitonin gene-related peptide I (1, 2). It has been postulated that the use of uracil and cytosine rather than adenosine as branch points may be involved in regulation of splicing of these pre-mRNAs, because selection of cytosine or uracil branch points occurs rather inefficiently in vitro, while the exon joining remains efficient (48).

The choice of this LAT branch point may be facilitated by the secondary structure of the 3′ end of the intron. The secondary structure of an RNA between a potential branch point sequence and the 3′ splice site may regulate splicing (5). A schematic diagram of the potential secondary structure in this region is shown in Fig. 10. There is a potential hairpin structure between the polypyrimidine tract at the 3′ splice site and the guanosine branch point of this intron. The ΔG of this hairpin is calculated to be −39.7 kcal/mol, using version 7.0 of the GCG package software (Genetic Computer Group, Inc., Madison Wis.), which suggests that this is a stable stem-loop structure. This hairpin may be a critical element influencing the selection of the LATin guanosine branchpoint. Additionally, the LATin 3′ end, which is detectable on at least a proportion of the transcripts (Fig. 8), may also stabilize the intron. This potential hairpin may augment the stability of this intron by binding to proteins and/or blocking the branch point from debranching enzymes.

Selection of guanosine branch points within a consensus branch site sequence in vitro is inefficient and inhibits subsequent exon joining (24, 30, 48). The selection efficiency of the 2.0-kb LATin guanosine branch point was not determined. Nevertheless, branch points at other nucleotides were not detected, although using this RT-PCR technique, we have amplified branch point sequences of introns which do not accumulate (Fig. 7A and data not shown). Additionally, accumulation of unspliced mLAT pre-mRNA or splicing intermediates of mLAT is not observed during HSV-1 infection

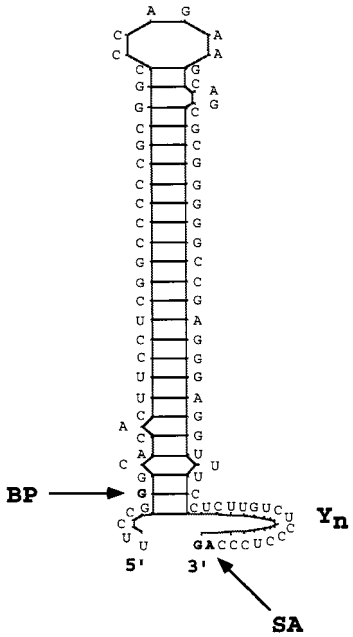

FIG. 10. Secondary structure at the end of the 2.0-kb LAT. A schematic diagram shows the potential secondary structure at the 3′ end of the 2.0-kb LAT. The branch point G (BP) and splice acceptor (SA) site nucleotides are in boldface and indicated by arrows. The polypyrimidine tract (Y_n) is shown upstream of the splice acceptor site. The ΔG of this hairpin is calculated to be −39.7 kcal/mol, using version 7.0 of the GCG package software (Genetic Computer Group, Inc.).

or transient transfections with our LAT expression vector. This may be due to the degradation of these intermediates. Alternatively, selection of the LATin branch point may not be inefficient or inhibitory to exon joining due to the surrounding sequences which bear little homology to the consensus mammalian branch site sequence or to a novel branch point selection mechanism.

Most introns are rapidly degraded within seconds of release from spliceosomes in vivo (reviewed in reference 42). The degradation of introns is thought to involve debranching of the intron followed by exonucleolytic cleavage (10). Guanosine branch points are poor substrates for mammalian debranching activity (3). The debranching of guanosine branch points occurs at approximately 50% the rate of adenosine branch points in vitro. Rodahl and Haarr (50) found that the 2.0-kb LAT lariat RNA was a poor substrate for debranching activity in vitro. In yeast cells, an intron containing the mutant branch point sequence UACUACC, with the underlined cytosine identified as the branch point, accumulated to high levels in a lariat configuration in vivo and was not efficiently debranched in vitro (32). Consistent with this finding, Wu et al. (67) have shown that the majority of the 2.0-kb LATin is in a nonlinear conformation, most likely a lariat. Together, these data imply that the stability of the 2.0-kb LATin is related to its unique branch point.

In conclusion, we have demonstrated that the 2.0-kb LAT is spliced out of the mLAT pre-mRNA and that the resulting stable intron has an unusual branch point—a guanosine. We hypothesis that the selection of a guanosine branch point plays a role in the extraordinary stability of the 2.0-kb LATin.

ACKNOWLEDGMENTS

We thank Giovanni Rovera for his interest and support of this work and the members of the Fraser and Rovera laboratories for their critique of the experiments. We also thank Miles F. Wilkinson (University of Texas M. D. Anderson Cancer Center) for many helpful discussions. We appreciate the excellent technical assistance provided by Andrew Dold and the members of the Wistar Institute DNA Synthesis and Sequencing facilities.

This work was supported by Public Health Service grant NS33768 from the National Institutes of Health. J.M.Z. was supported by National Institutes of Health training grant AI07324. C.K. was supported by the Swiss National Science Foundation.

REFERENCES

1. **Adema, G. J., R. A. L. Bovenberg, H. S. Jansz, and P. C. Baas.** 1988. Unusual branch point selection involved in splicing of the alternatively processed calcitonin/CGRP-1 pre-mRNA. Nucleic Acids Res. **16:**9513–9526.
2. **Adema, G. J., K. L. v. Hulst, and P. D. Baas.** 1990. Uridine branch acceptor is a cis-acting element involved in regulation of the alternative processing of calcitonin/CGRP-1 pre-mRNA. Nucleic Acids Res. **18:**5365–5373.
3. **Arenas, J., and J. Hurwitz.** 1987. Purification of a RNA debranching activity from HeLa cells. J. Biol. Chem. **262:**4274–4279.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York, N.Y.
5. **Balvay, L., D. Libri, and M. Y. Miszman.** 1993. Pre-mRNA secondary structure and the regulation of splicing. Bioessays **15:**165–169.
6. **Batchelor, A. H., and P. O'Hare.** 1990. Regulation and cell-type-specific activity of a promoter located upstream of the latency-associated transcript of herpes simplex virus type 1. J. Virol. **64:**3269–3279.
7. **Bludau, H., and U. K. Freese.** 1991. Analysis of the HSV-1 strain 17 DNA polymerase gene reveals the expression of four different classes of pol transcripts. Virology **183:**505–518.
8. **Bohenzky, R. A., M. Lagunoff, B. Roizman, E. K. Wagner, and S. Silverstein.** 1995. Two overlapping transcription units which extend across the L-S junction of herpes simplex virus type 1. J. Virol. **69:**2889–2897.
9. **Chang, D. D., and P. A. Sharp.** 1989. Regulation by HIV Rev depends upon recognition of splice sites. Cell **59:**789–795.
10. **Chapman, K. B., and J. D. Boeke.** 1991. Isolation and characterization of the gene encoding yeast debranching enzyme. Cell **65:**483–492.
11. **Chen, X., M. C. Schmidt, W. F. Goins, and J. C. Glorioso.** 1995. Two herpes simplex virus type 1 latency-active promoters differ in their contributions to latency-associated transcript expression during lytic and latent infections. J. Virol. **69:**7899–7909.
12. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Orientation of herpes simplex virus type-1 immediate early mRNAs. Biochemistry **18:**5294–5299.
13. **Copertino, D. W., and R. B. Hallick.** 1991. Group II twintron: an intron within an intron in a chloroplast cytochrome b-559 gene. EMBO J. **10:**433–442.
14. **Costa, R. H., K. G. Draper, T. J. Kelly, and E. K. Wagner.** 1985. An unusual spliced herpes simplex virus type 1 transcript with sequence homology to Epstein-Barr virus. J. Virol. **54:**317–340.
15. **Cullen, B. R., and M. H. Malim.** 1991. The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional activator. Trends Biochem. Sci. **16:**346–350.
16. **Dallman, M. J., and A. C. G. Porter.** 1991. Semi-quantitative PCR for the analysis of gene expression, p. 215–224. *In* M. J. McPherson, P. Quirke, and G. R. Taylor (ed.), PCR: a practical approach, vol. 1. Oxford University Press, Oxford, England.
17. **Deatly, A. M., J. G. Spivack, E. Lavi, and N. W. Fraser.** 1987. RNA from an immediate early region of the HSV-1 genome is present in the trigeminal ganglia of latently infected mice. Proc. Natl. Acad. Sci. USA **84:**3204–3208.
18. **Deatly, A. M., J. G. Spivack, E. Lavi, D. R. O'Boyle II, and N. W. Fraser.** 1988. Latent herpes simplex virus type 1 transcripts in peripheral and central nervous system tissues of mice map to similar regions of the viral genome. J. Virol. **62:**749–756.
19. **Devi-Rao, G. B., S. A. Goodart, L. M. Hecht, R. Rochford, M. A. Rice, and E. K. Wagner.** 1991. Relationship between polyadenylated and nonpolyadenylated herpes simplex virus type 1 latency-associated transcripts. J. Virol. **65:**2179–2190.
20. **Dobson, A. T., F. Sederati, G. Devi-Rao, J. Flanagan, M. J. Farrell, J. G. Stevens, E. K. Wagner, and L. T. Feldman.** 1989. Identification of the latency-associated transcript promoter by expression of rabbit beta-globin mRNA in mouse sensory nerve ganglia latently infected with a recombinant herpes simplex virus. J. Virol. **63:**3844–3851.
21. **Dreyfuss, G., M. Hentze, and A. I. Lamond.** 1996. From transcript to protein. Cell **85:**963–972.
22. **Dyhr-Mikkelsen, H., and J. Kjems.** 1995. Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro. J. Biol. Chem. **270:**24060–24066.
23. **Farrell, M. J., A. T. Dobson, and L. T. Feldman.** 1991. Herpes simplex virus latency-associated transcript is a stable intron. Proc. Natl. Acad. Sci. USA **88:**790–794.

24. **Freyer, G. A., J. Arenas, K. K. Perkins, H. M. Furneaux, L. Pick, B. Young, R. J. Roberts, and J. Hurwitz.** 1987. In vitro formation of the lariat structure containing a G2′-5′G linkage. J. Biol. Chem. **262:**4267–4273.
25. **Frink, R. J., R. J. Eisenberg, G. H. Cohen, and E. K. Wagner.** 1983. Detailed analysis of the portion of the herpes simplex virus type 1 genome encoding glycoprotein C. J. Virol. **45:**634–647.
26. **Goins, W. F., L. R. Sternberg, K. D. Croen, P. R. Krause, R. L. Hendricks, D. J. Fink, S. E. Straus, M. Levine, and J. C. Glorioso.** 1994. A novel latency-active promoter is contained within the herpes simplex virus type 1 UL flanking repeats. J. Virol. **68:**2239–2252.
27. **Gonzalez, A.-M., M. Buscaglia, M. Ong, and A. Baird.** 1990. Distribution of basic fibroblast growth factor in the 18-day rat fetus: localization in the basement membranes of diverse tissues. J. Cell Biol. **110:**753–765.
28. **Hardy, W. R., and R. M. Sandri-Goldin.** 1994. Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. J. Virol. **68:**7790–7799.
29. **Harmuth, K., and A. Barta.** 1988. Unusual branch point selection in processing of human growth hormone pre-mRNA. Mol. Cell. Biol. **8:**2011–2020.
30. **Hornig, H., M. Aebi, and C. Weissmann.** 1986. Effect of mutations at the lariat branch acceptor site on the beta-globin pre-mRNA splicing in vitro. Nature **324:**589–591.
31. **Izaurradle, E., and I. W. Mattaj.** 1995. RNA export. Cell **81:**153–159.
32. **Jacquier, A., and M. Rosbash.** 1986. RNA splicing and intron turnover are greatly diminished by a mutant yeast branch point. Proc. Natl. Acad. Sci. USA **83:**5835–5839.
33. **Krause, P. R., K. D. Croen, J. M. Ostrove, and S. E. Straus.** 1990. Structural and kinetic analysis of herpes simplex virus type 1 latency-associated transcripts in human trigeminal ganglia and in culture. J. Clin. Invest. **86:**5619–5623.
34. **Krause, P. R., J. M. Ostrove, and S. E. Straus.** 1991. The nucleotide sequence, 5′ end, promoter domain, and kinetics of expression of the gene encoding the herpes simples virus type 2 latency-associated transcript. J. Virol. **65:**5619–5623.
35. **Lagunoff, M., G. Randall, and B. Roizman.** 1996. Phenotypic properties of herpes simplex virus 1 containing a derepressed open reading frame P gene. J. Virol. **70:**1810–1817.
36. **Lagunoff, M., and B. Roizman.** 1994. Expression of a herpes simplex virus 1 open reading frame antisense to the gamma 34.5 gene and transcribed by an RNA 3′ coterminal with the unspliced latency-associated transcript. J. Virol. **68:**6021–6028.
37. **Lagunoff, M., and B. Roizman.** 1995. The regulation of synthesis and properties of the protein product of open reading frame P of the herpes simplex virus 1 genome. J. Virol. **69:**3615–3623.
38. **Legrain, P., and M. Rosbash.** 1989. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. Cell **57:**573–583.
39. **McGeoch, D. J., C. Cunningham, G. McIntyre, and A. Dolan.** 1991. Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2. J. Gen. Virol. **69:**1531–1574.
40. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69:**1531–1574.
41. **Mitchell, W. J., R. P. Lirette, and N. W. Fraser.** 1990. Mapping of low abundance latency associated RNA in the trigeminal ganglia of mice latently infected with herpes simplex virus type 1. J. Gen. Virol. **71:**125–132.
42. **Moore, M. J., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursors to messenger RNAs by the spliceosome, p. 1–30. *In* R. F. Gesteland and J. F. Atkins (ed.), The RNA world: the nature of modern RNA suggests a prebiotic RNA world. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
43. **Nicosia, M., S. L. Deshmane, J. M. Zabolotny, T. Valyi-Nagy, and N. W. Fraser.** 1993. Herpes simplex virus type 1 latency-associated transcript (LAT) promoter deletion mutants can express a 2-kilobase transcript mapping to the LAT region. J. Virol. **67:**7276–7283.
44. **Nicosia, M., J. M. Zabolotny, R. P. Lirette, and N. W. Fraser.** 1994. The HSV-1 2 kb latency-associated transcript is found in the cytoplasm comigrating with ribosomal subunits during productive infection. Virology **204:**717–728.
45. **Perry, L. J., F. J. Rixon, R. D. Everett, M. C. Frame, and D. J. McGeogh.** 1986. The IE110 gene of herpes simplex virus type 1: characterization by mRNA mapping, DNA sequence, oligopeptide antiserum and mutational analysis. J. Gen. Virol. **67:**2365–2374.
46. **Post, L. E., A. J. Conley, E. S. Mocarski, and B. Roizman.** 1980. Cloning of reiterated and nonreiterated herpes simplex virus 1 sequences as BamHI fragments. Proc. Natl. Acad. Sci. USA **77:**4201–4205.
47. **Qian, L., M. N. Vu, M. Carter, and M. F. Wilkinson.** 1992. A spliced intron accumulates as a lariat in the nucleus of T cells. Nucleic Acids Res. **20:**5345–5350.
48. **Query, C. C., S. A. Strobel, and P. A. Sharp.** 1996. Three recognition events at the branch-site adenine. EMBO J. **15:**1392–1402.
49. **Rock, D. L., A. B. Nesburn, H. Ghiasi, J. Ong, T. L. Lewis, J. R. Lokensgard, and S. M. Wechsler.** 1987. Detection of latency-related viral RNAs in trigeminal ganglia of rabbits latently infected with herpes simplex virus type 1. J. Virol. **61:**3820–3826.
50. **Rodahl, E., and L. Haarr.** 1997. Analysis of the 2-kilobase latency-associated transcript expressed in PC12 cells productively infected with herpes simplex virus type 1: evidence for a stable, nonlinear structure. J. Virol. **71:**1703–1707.
51. **Sacks, W. R., C. C. Greene, D. P. Ashman, and P. A. Schaffer.** 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. **55:**796–805.
52. **Sawtell, N. M., and R. L. Thompson.** 1992. Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. J. Virol. **66:**2157–2169.
53. **Singh, J., and E. K. Wagner.** 1993. Transcriptional analysis of the herpes simplex virus type 1 region containing the TRL/UL junction. Virology **196:**220–231.
54. **Spivack, J. G., and N. W. Fraser.** 1987. Detection of herpes simplex virus type 1 transcripts during latent infection in mice. J. Virol. **61:**3841–3847.
55. **Spivack, J. G., and N. W. Fraser.** 1988. Expression of herpes simplex virus type 1 (HSV-1) latency-associated transcripts and transcripts affected by the detection in the avirulent mutant HFEM: evidence for a new class of HSV-1 genes. J. Virol. **62:**3281–3287.
56. **Spivack, J. G., G. M. Woods, and N. W. Fraser.** 1991. Identification of a novel latency-specific splice donor signal within the herpes simplex virus type 1 2.0-kb latency-associated transcript (LAT): translation inhibition of LAT open reading frames by the intron within the 2.0-kb LAT. J. Virol. **65:**6800–6810.
57. **Steiner, I., J. G. Spivack, R. P. Lirette, S. M. Brown, A. R. MacLean, J. Subak-Sharpe, and N. W. Fraser.** 1989. Herpes simplex virus type 1 latency-associated transcripts are evidently not essential for latent infection. EMBO J. **8:**505–511.
58. **Stevens, J. G., E. K. Wagner, G. B. Devi-Rao, M. L. Cook, and L. T. Feldman.** 1987. RNA complementary to a herpes virus gene mRNA is prominent in latently infected neurons. Science **235:**1056–1059.
59. **Stroop, W. G., D. L. Rock, and N. W. Fraser.** 1984. Localization of herpes simplex virus in the trigeminal and olfactory systems of the mouse central nervous system during acute and latent infections by in situ hybridization. Lab. Invest. **51:**27–38.
60. **Tarn, W.-Y., and J. A. Steitz.** 1996. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science **273:**1824–1832.
61. **Tarn, W.-Y., and J. A. Steitz.** 1996. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) of introns in vitro. Cell **84:**801–811.
62. **Wagner, E. K., G. Devi-Rao, L. T. Feldman, A. T. Dobson, Y. F. Zhang, J. M. Hill, W. M. Flanagan, and J. G. Stevens.** 1988. Physical characterization of the herpes simplex virus latency-associated transcript in neurons. J. Virol. **62:**1194–1202.
63. **Wagner, E. K., W. M. Flanagan, G. Devi-Rao, Y. F. Zhang, J. M. Hill, K. P. Anderson, and J. G. Stevens.** 1988. The herpes simplex virus latency-associated transcript is spliced during the latent phase of infection. J. Virol. **62:**4577–4585.
64. **Watson, R. J., M. Sullivan, and G. F. V. Woude.** 1981. Structure of two spliced herpes simplex virus type 1 immediate-early mRNAs. J. Virol. **37:**431–444.
65. **Wechsler, S. L., A. B. Nesburn, R. Watson, S. M. Slanina, and and H. Ghiasi.** 1988. Fine mapping of the latency-related gene of herpes simplex virus type 1: alternate splicing produces distinct latency-related RNAs containing open reading frames. J. Virol. **70:**2163–2169.
66. **Wechsler, S. L., J. Nesburn, N. Zaagstra, and H. Ghiasi.** 1989. Sequence of the latency related gene of herpes simplex virus type 1. Virology **168:**168–172.
67. **Wu, T.-T., Y.-H. Su, T. M. Block, and J. M. Taylor.** 1996. Evidence that two latency-associated transcripts of herpes simplex virus are nonlinear. J. Virol. **70:**5962–5967.
68. **Yeh, L., and P. A. Schaffer.** 1993. A novel class of transcripts expressed with late kinetics in the absence of ICP4 spans the junction between the long and short segments of the herpes simplex virus type 1 genome. J. Virol. **67:**7373–7382.
69. **Zwaagstra, J., H. Ghiasi, A. B. Nesburn, and S. L. Wechsler.** 1991. Identification of a major regulatory sequence in the latency associated transcript (LAT) promoter of herpes simplex virus type 1 (HSV-1). Virology **182:**287–297.
70. **Zwaagstra, J., H. Ghiasi, A. B. Nesburn, and S. L. Wechsler.** 1989. In vitro promoter activity associated with the latency-associated transcript gene of herpes simplex virus type 1. J. Gen. Virol. **70:**2163–2169.
71. **Zwaagstra, J., H. Ghiasi, S. M. Slanina, A. B. Nesburn, S. C. Wheatley, K. Lillycrop, J. Wood, D. S. Latchman, K. Patel, and S. L. Wechsler.** 1990. Activity of herpes simplex virus type 1 latency-associated transcript (LAT) promoter in neuron-derived cells: evidence for neuron specificity and for a large LAT transcript. J. Virol. **64:**5019–5028.