

# Binding of hnRNP L to the Pre-mRNA Processing Enhancer of the Herpes Simplex Virus Thymidine Kinase Gene Enhances both Polyadenylation and Nucleocytoplasmic Export of Intronless mRNAs

Shouhong Guang, Alicia M. Felthouser, and Janet E. Mertz\*

McArdle Laboratory for Cancer Research, University of Wisconsin School of Medicine, Madison, Wisconsin 53706-1599

Received 26 January 2005/Returned for modification 21 February 2005/Accepted 12 May 2005

Liu and Mertz (*Genes Dev.* 9:1766–1780, 1995) previously identified a 119-nt pre-mRNA processing enhancer (PPE) element within the herpes simplex virus type 1 thymidine kinase gene that enables intron-independent gene expression in higher eukaryotes by binding heterogeneous nuclear ribonucleoprotein L (hnRNP L). Here, we identify a 49-nt subelement within this PPE that enhanced stability, polyadenylation, and cytoplasmic accumulation of transcripts synthesized in CV-1 cells from an intronless variant of the human  $\beta$ -globin gene when present in two or more tandem copies. This 2 $\times$ TK49 PPE also enhanced (i) the efficiency of polyadenylation of intronless  $\beta$ -globin RNA in a cell-free polyadenylation system and (ii) the kinetics of nucleocytoplasmic export of an intronless variant of adenovirus major late leader region RNA in *Xenopus* oocytes. This 2 $\times$ TK49 PPE bound only hnRNP L. Analysis of 2 $\times$ TK49 PPE mutants showed a strong positive correlation existed between binding hnRNP L and enhancement of intronless  $\beta$ -globin gene expression. hnRNP L was found to associate with both the mRNA export factor TAP and the exon-exon junction complex protein Aly/REF. Thus, we conclude that hnRNP L plays roles in enhancing stability, polyadenylation, and nucleocytoplasmic export; it does so, at least in part, by directly recruiting to intronless PPE-containing RNAs cofactors normally recruited to intron-containing RNAs.

Heterogeneous nuclear ribonucleoproteins (hnRNPs) were first described as proteins that bind to nascent transcripts synthesized by RNA polymerase II, packaging these heterogeneous nuclear RNAs into hnRNP particles (9, 25). Many hnRNPs are abundant nuclear proteins that shuttle between the nucleus and cytoplasm (44). hnRNPs interact via protein-protein contacts both with themselves and with other factors (24, 25). They bind to specific sequences in RNA, thereby influencing the utilization of specific 5' and 3' splice sites, the efficiency of 3' end formation, nucleocytoplasmic transport, localization within the cell, translation, and turnover of the RNA (25).

hnRNP L was first identified as an abundant nuclear protein present in lampbrush chromosomes (43). More recently, it has been shown (i) to affect translation by binding the 3' border of the internal ribosomal entry site of hepatitis C virus (14), (ii) to affect stability by binding the 3' untranslated regions of human vascular endothelial growth factor (48) and glucose transporter 1 (15) mRNAs, and (iii) to stimulate stability and splicing by binding a CA repeat region in endothelial cell nitric oxide synthase pre-mRNA (21, 22).

The pre-mRNA processing enhancer (PPE) is a 119-nt sequence element present within the transcribed region of the naturally intronless herpes simplex virus type 1 thymidine kinase (HSV-TK) gene that can enhance cytoplasmic accumulation of intronless  $\beta$ -globin mRNA in the absence of splicing

(32). It can also enhance cytoplasmic accumulation without splicing of an intron-containing mRNA, doing so through a Crm1-independent pathway (17, 41). Liu and Mertz (32) also reported that hnRNP L binds the TK119 PPE, with a positive correlation existing between hnRNP L binding and enhancement of intronless gene expression. However, other proteins were also noted to bind this TK119 PPE. Thus, it remained unclear whether hnRNP L binding to the PPE was sufficient to recruit to the RNA all of the factors needed to enhance stabilization, processing, and nucleocytoplasmic export. Also unknown was the mechanism(s) by which hnRNP L enhances these multiple steps in mRNA biogenesis.

We report here the identification of a smaller, 49-nt subelement of the TK119 PPE that still enables intronless  $\beta$ -globin gene expression when present in two or more tandem copies. This 2 $\times$ TK49 PPE is bound only by hnRNP L. Its presence enhances both polyadenylation and nucleocytoplasmic export of intronless RNAs. We also report that hnRNP L associates with the mRNA export factor TAP and the exon-exon junction complex protein Aly/REF. Thus, we conclude that the presence of this PPE facilitates processing of intronless transcripts by binding hnRNP L which, in turn, functions as an adaptor protein to enhance both polyadenylation and nucleocytoplasmic export by recruiting to the RNA cellular factors normally recruited by exon-exon junction complexes during recognition and excision of introns.

\* Corresponding author. Mailing address: McArdle Laboratory for Cancer Research, 1400 University Avenue, University of Wisconsin, Madison, WI 53706-1599. Phone: (608) 262-2383. Fax: (608) 262-2824. E-mail: mertz@oncology.wisc.edu.

## MATERIALS AND METHODS

**Cells and transfections.** The African green monkey kidney cell line CV-1PD was grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 100 U penicillin and streptomycin per ml and 5% fetal bovine serum as described

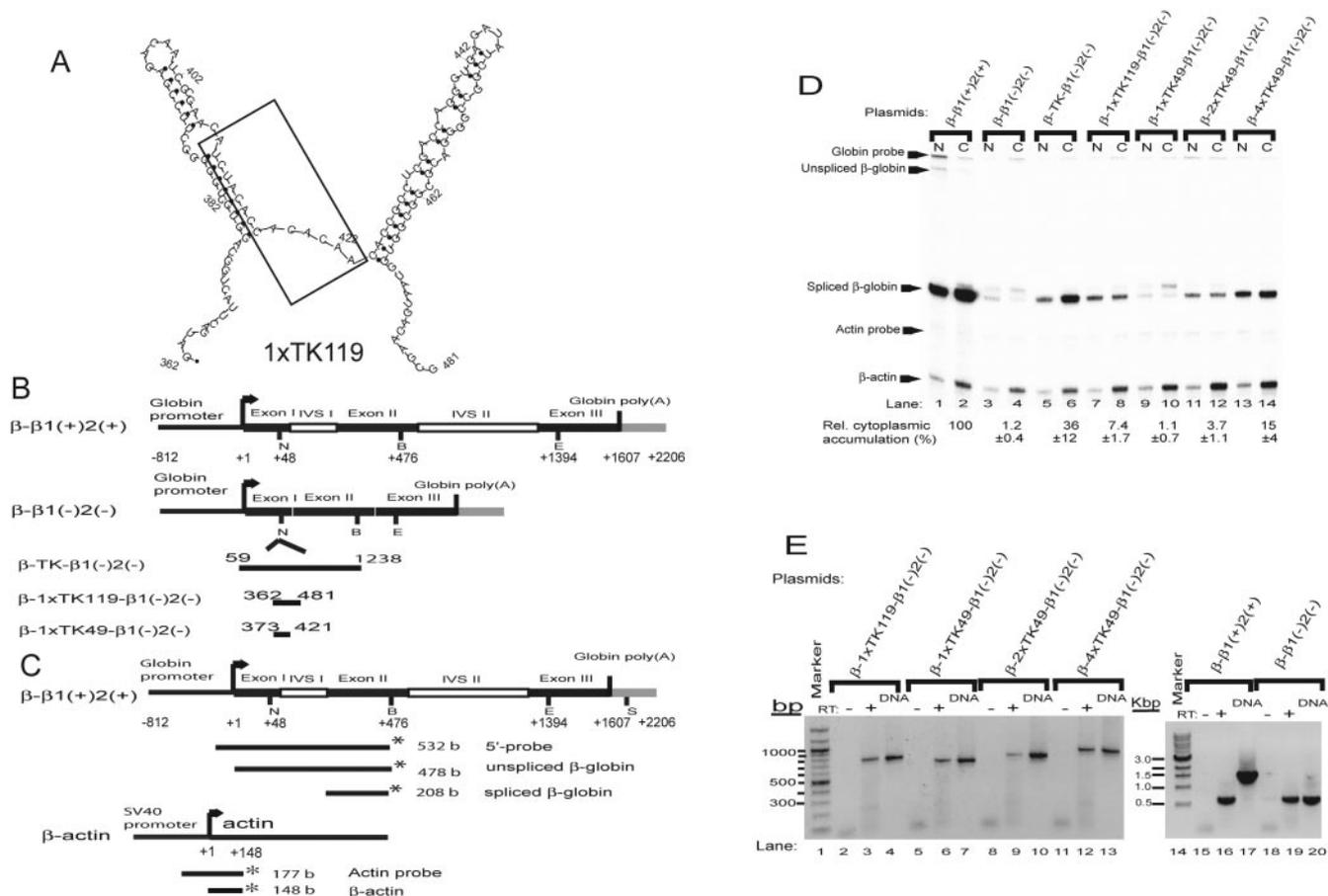


FIG. 1. A 49-nt sequence from the HSV-TK gene can enhance cytoplasmic accumulation of intronless  $\beta$ -globin-like transcripts when present in two or more tandem copies. (A) Computer-predicted secondary structure of 1xTK119 PPE. Secondary structure was determined using the mFOLD program (37, 52). The boxed region includes bases previously determined by cluster mutagenesis to be important for PPE function (32). (B) Structures of plasmids containing HSV-TK's full-length coding region, 1xTK119, and 1xTK49 inserted into the NcoI site in exon I of a cDNA variant of the human  $\beta$ -globin gene. All numbers are given relative to the site of transcription initiation from each gene. Black rectangles indicate exons. Open rectangles indicate introns. Gray rectangles indicate sequences 3' of the site of cleavage for polyadenylation. Rectangles are not drawn to scale. N, NcoI; B, BamHI; E, EcoRI. (C) Structure of the human  $\beta$ -globin and human  $\beta$ -actin probes used for S1 nuclease mapping analysis. Probes were amplified by PCR and end-labeled with  $^{32}$ P indicated by the \* as described in Materials and Methods. The sizes of the DNA fragments resulting from protection by hybridization with the corresponding RNAs are indicated. S, SspI. (D) Autoradiogram of quantitative S1 nuclease mapping analysis of the human  $\beta$ -globin-like RNAs accumulated in the nucleus and cytoplasm of CV-1PD cells transfected with the indicated plasmids. CV-1PD cells were cotransfected in parallel with 2  $\mu$ g of each of the indicated plasmids together with 1  $\mu$ g of the SV40 T antigen-encoding plasmid pRSV-Tori. Nuclear (N) and cytoplasmic (C) RNAs were harvested 48 h later and analyzed by concurrent S1 nuclease mapping with the 5'-end-labeled  $\beta$ -globin and  $\beta$ -actin probes shown in panel C. The amount of  $\beta$ -globin-like RNA accumulated in the cytoplasm was internally normalized to the amount of cellular  $\beta$ -actin RNA present in the same sample. The data were calculated as percentages relative to the amount of  $\beta$ -globin RNA accumulated in the cytoplasm of cells transfected in parallel with  $\beta$ - $\beta$ 1(+)(+) RNA after internal normalization as well to the relative amounts of DpnI-resistant globin-encoding plasmid DNA present in these cells. The numbers below the lanes are means  $\pm$  SEMs of data obtained from three independent experiments similar to the one shown here. (E)  $\beta$ -Globin-like RNAs accumulated in the cytoplasm are unspliced. Portions of the cytoplasmic RNA samples from the experiment shown in panel A were reverse transcribed and then amplified by PCR with the primers described in Materials and Methods (+RT, lanes 3, 6, 9, 12, 16, and 19). As controls, PCR amplification reactions were performed on the RNA samples without prior reverse transcription (-RT, lanes 2, 5, 8, 11, 15, and 18) and on the plasmid DNAs used in the transfections (DNA, lanes 4, 7, 10, 13, 17, and 20). Shown here is a photograph of an ethidium bromide-stained, 1% agarose gel in which the PCR products were electrophoresed. Lanes 1 and 14 contain size markers.

previously (11). Cotransfections were performed by the DEAE-dextran/chloroquine procedure as described previously with 3  $\mu$ g DNA total per 100-mm-diameter dish (30–32). 293T cells were grown in 100-mm-diameter dishes in DMEM supplemented with 10% fetal bovine serum. They were transfected in the presence of 300  $\mu$ l whole medium containing 6  $\mu$ g total plasmid DNA premixed with 18  $\mu$ l TransIT LT-1 reagent (Mirus Corp.), followed by incubation for 48 h.

**Recombinant plasmids.** Plasmids  $\beta$ - $\beta$ 1(+)(+),  $\beta$ - $\beta$ 1(-)(-),  $\beta$ -TK- $\beta$ 1(-)(-), and  $\beta$ -1xTK119- $\beta$ 1(-)(-) have been described previously (32, 33, 50). They contain the nt -812 to +2206 region of the human  $\beta$ -globin gene

relative to the transcription initiation site with (+) or without (-)  $\beta$ -globin's two introns in a pBR322-based cloning vector containing a simian virus 40 (SV40) origin of DNA replication (Fig. 1B). Synthetic oligonucleotides named TK49, TK49LS0, TK49LS0com, and TK49m (Fig. 2A) were annealed and inserted as two or four head-to-tail tandemly repeated copies into the BamHI and BglII sites of pSP72 (Promega) to generate plasmids pSP72/2xTK49 or pSP72/4xTK49, respectively, and the TK49 mutant variants thereof. Plasmids  $\beta$ -2xTK49wt- $\beta$ 1(-)(-),  $\beta$ -2xTK49LS0- $\beta$ 1(-)(-),  $\beta$ -2xTK49LS0com- $\beta$ 1(-)(-),  $\beta$ -2xTK49m- $\beta$ 1(-)(-), and  $\beta$ -4xTK49wt- $\beta$ 1(-)(-) were constructed by PCR amplification of the TK sequences in these pSP72/2xTK49 and pSP72/4xTK49

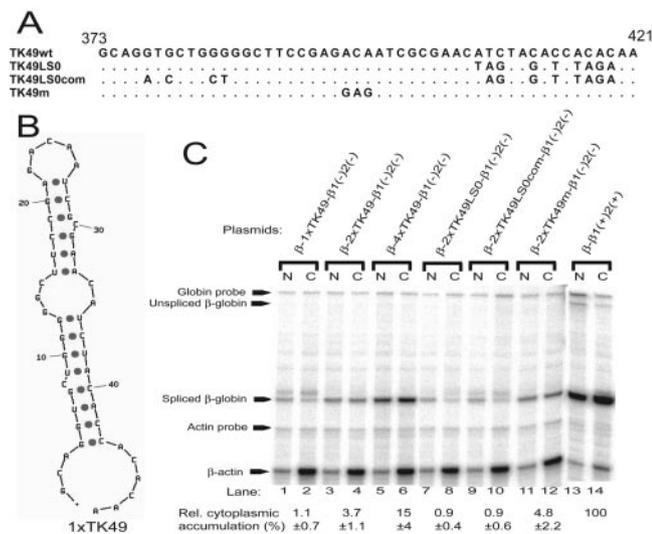


FIG. 2. Effects of mutations in 2xTK49 on enhancement of intron-independent expression of human  $\beta$ -globin gene. (A) Sequences of TK49 PPE and its cluster base substitution mutants. (B) Computer-predicted secondary structure of 1xTK49. (C) Autoradiogram of quantitative S1 nuclease mapping analysis of  $\beta$ -globin-like RNAs accumulated in the nucleus (N) and cytoplasm (C) by 48 h after cotransfection of CV-1PD cells with the indicated plasmids and pRSV-Tori. The S1 nuclease mapping probes were the 5'-end-labeled ones shown in Fig. 1C. The samples were analyzed as described in the legend to Fig. 1D. The numbers shown here are means  $\pm$  standard errors of the mean (SEMs) of data obtained from three independent experiments similar to the one shown here.

wild-type (wt) and mutant variant plasmids, cleavage with NcoI, and insertion at the NcoI site in exon I of p $\beta$ - $\beta$ 1(-)2(-) in the sense orientation (Fig. 1B). Plasmids pGEM5z/TK119 and pGEM5z/TK119LS0 have been described previously (32). Plasmids pSP72/2xTK49wt/ $\beta$ (A) and pSP72/2xTK49LS0/ $\beta$ (A) were constructed by insertion of part of the  $\beta$ -globin gene sequence, -100 to +416 nt relative to the cleavage site for polyadenylation, into the XbaI site of plasmid pSP72/2xTK49. Plasmid pRSV-Tori, encoding SV40 large T antigen transcribed from the Rous sarcoma virus promoter, has been described previously (32). Plasmid pAdML- $\Delta$ i, containing the first two exons of the adenovirus type 2 major late leader region, has been described previously (35). Plasmid p $\Delta$ i-2xTK49 was constructed by PCR amplification and insertion of 2xTK49wt into the KpnI site of pAdML- $\Delta$ i.

Plasmid pHCL3, containing a full-length cDNA encoding hnRNP L expressed from an SP6 promoter, has been described previously (32). A plasmid encoding a glutathione S-transferase (GST)-hnRNP L fusion protein was made by PCR amplification of the hnRNP L sequences of pHCL3 encoding amino acid residues 51 to 558 of hnRNP L and insertion between the BamHI and EcoRI sites of pGEX-2T (Amersham Bioscience). Deletion variants of GST-hnRNP L were constructed likewise. A plasmid expressing mouse Aly/REF2-II protein from a T7 promoter was kindly provided by E. Izaurralde (49). Plasmids pFLAG-hnRNP A1 (19), pFLAG-TAP (19), pFLAG-SRm160 (38), and pCS2/FLAG-Aly (4) have been described previously.

**RNA purification and S1 nuclease mapping analysis.** Cells were harvested 48 h after transfection. Nuclei and cytoplasm were separated by treatment with 0.5% NP-40 as described previously (30-32). RNA was purified using TRIzol reagent following the manufacturer's suggestions (Gibco). The relative amounts of globin-like RNA accumulated in the nuclear and cytoplasmic fractions were determined by quantitative S1 nuclease mapping as described previously (32, 33) and quantified with a PhosphorImager and ImageQuant software. Cellular  $\beta$ -actin RNA served as an internal control for recovery of the RNA samples (32, 33). The probes used in the S1 nuclease mapping assays are shown in Fig. 1C and Fig. 3A. They were prepared by PCR amplification followed by 5'-end-labeling with T4 DNA kinase or 3'-end-labeling with Klenow polymerase after cleavage with EcoRI endonuclease. The templates used for synthesis of the 5'-end-labeled  $\beta$ -globin and  $\beta$ -actin probes are indicated in Fig. 1C. The template for synthesis of the 3'-end-labeled  $\beta$ -globin probe was prepared by subcloning into the EcoRV

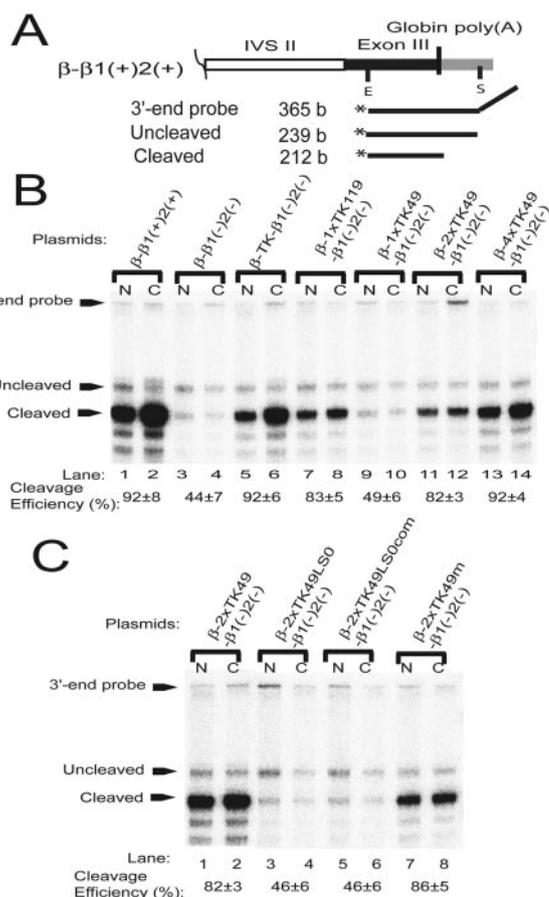


FIG. 3. PPEs also enhance 3' end processing of intronless  $\beta$ -globin-like RNAs in situ. (A) Structure of  $\beta$ -globin 3' end probe used in S1 mapping analysis. Probes were amplified by PCR and 3'-end-labeled with  $^{32}$ P. The sizes of the DNA fragments resulting from protection by hybridization with the corresponding RNAs are indicated. (B and C) Autoradiograms of quantitative S1 nuclease mapping analysis of the 3' ends of the  $\beta$ -globin-like RNAs accumulated in the nucleus and cytoplasm of CV-1PD cells. Portions of the RNA samples from the experiment shown in Fig. 1 and 2 were mapped with the 3'-end-labeled probe shown in panel A. The numbers below the lanes are means  $\pm$  SEMs of the cleaved RNA (N+C) divided by the cleaved plus uncleaved RNA (N+C) times 100% of data obtained from three independent experiments similar to the one shown here.

site of pGEM5z(+), (Promega) an SspI fragment (460 bp) of p $\beta$ - $\beta$ 1(+)-2(+), which contains part of intron II, exon III, and an additional 27 bp downstream of the site of cleavage for polyadenylation. Southern blot analysis of the relative amounts of DpnI-resistant,  $\beta$ -globin-encoding plasmid DNA present in the nuclear and cytoplasmic fractions prior to treatment with DNase I was performed as described previously (31, 47) to assay both relative transfection efficiencies and contamination of cytoplasmic RNA with nuclear nucleic acid. Contamination of nuclear RNA by cytoplasmic RNA was determined by Northern blot analysis of 18S rRNA. The occasional sample with greater than 20% cross-contamination of nuclear RNA with cytoplasmic RNA or vice versa was discarded.

**RT-PCR analysis.** Each RNA sample examined by reverse transcription (RT)-PCR analysis was first pretreated with RNase-free DNase I (Ambion) to degrade contaminating plasmid DNA. Each RT reaction contained 300 ng total RNA, 20 U avian myeloblastosis virus reverse transcriptase (Roche), and 15 pmol  $\beta$ -globin reverse primer (5'-CCAGATGCTCAAGGCC-3'), corresponding to the 3' end of the  $\beta$ -globin gene, in a total volume of 20  $\mu$ l. After incubation at 42°C for 1 h, the reaction mixture was incubated at 99°C for 5 min. Afterwards, the PCR was performed as described previously (32) with Taq DNA polymerase (Promega) after addition of the  $\beta$ -globin forward primer (5'-ACATTTGCTTCTGACACA ACTG), corresponding to the 5' end of  $\beta$ -globin mRNA. A thermal cycler

(Techne) was used, with denaturation at 94°C for 30 s, annealing at 50°C for 30 s, and extension at 72°C for 2 min for 40 cycles (32).

**Cell-free polyadenylation assay.** Plasmids pSP72/2×TK49wt/β(A) and pSP72/2×TK49LS0/β(A) were PCR amplified with primers 5'-TAATACGACTCACT ATAGG-3' and 5'-TTTTTTTTTTTTTTTGGCAATGAAAATAAATGTTTTTTT ATTAGG-3' to generate templates containing a T7 promoter, a 2×TK49 PPE, the 3'-terminal 100 nt of β-globin's exon III, and 15 (A:T) bp directly adjacent to the cleavage site for polyadenylation (Fig. 4A). <sup>32</sup>P-labeled, m<sup>7</sup>G-capped RNAs were synthesized from these templates with a RiboProbe kit (Promega) in the presence of [<sup>32</sup>P]UTP and the cap analogue m<sup>7</sup>G(5')ppp(5')G and gel purified. HeLa cell nuclear extract was prepared with a modification of the Dignam method (1, 6). Cell-free polyadenylation assays were performed essentially as described previously (39, 40) with 48% nuclear extract (6 to 7 μg protein per μl reaction volume) in 40 mM KCl, 1 mM ATP, 20 mM phosphocreatine, 0.2 mM dithiothreitol, 0.35 mM EDTA, 1.5 mM MgCl<sub>2</sub>, 2.8% polyvinyl alcohol, and 10 U RNasin. Reactions were incubated at 30°C for the times indicated. The resulting RNAs were purified by phenol extraction and ethanol precipitation and separated by electrophoresis in 8 M urea-6% polyacrylamide gels.

**Xenopus oocyte nucleocytoplasmic export assay.** Plasmids pAdML-Δi and pΔi-2×TK49 were linearized at their BamHI site (Fig. 5A). RNAs were synthesized with a RiboProbe kit (Promega) in the presence of [<sup>32</sup>P]UTP and the cap analogue m<sup>7</sup>G(5')ppp(5')G. The U3 and U1 RNAs were synthesized as described previously (42). Coinjection of the <sup>32</sup>P-labeled RNAs into the nuclei of *Xenopus* oocytes was performed as described previously (42). After incubation at 18°C for the times indicated, the oocytes were manually dissected to separate the nuclei from the cytoplasm. The RNAs were purified by phenol extraction and ethanol precipitation and separated by electrophoresis in 8 M urea-6% polyacrylamide gels.

**UV-cross-linking assay.** All RNA substrates used for cell-free studies were synthesized with a RiboProbe kit (Promega) in the presence of [<sup>32</sup>P]UTP and the cap analogue m<sup>7</sup>G(5')ppp(5')G. Plasmids pGEM5z/TK119 and pGEM5z/TK119LS0 were linearized at their NcoI site. Plasmids pSP72/1×TK49, pSP72/2×TK49wt, pSP72/2×TK49LS0, pSP72/2×TK49LS0com, pSP72/2×TK49m, and pSP72/4×TK49wt were linearized at their BamHI site. UV-cross-linking and immunoprecipitation assays were performed essentially as described previously (32). For UV-cross-linking assays, 10<sup>5</sup> cpm of <sup>32</sup>P-labeled RNA was incubated with 20 μg of HeLa cell nuclear extract at 30°C for 10 min. Afterward, the reaction mixture was irradiated with a UV-Stratalinker (Stratagene) for 10 min on the automatic setting. RNases A and T1 were added, and the mixtures were further incubated at 37°C for 10 min. The proteins in each sample were separated by electrophoresis in a sodium dodecyl sulfate (SDS)-12% polyacrylamide gel and quantified with a PhosphorImager. Immunoprecipitation was performed after RNase treatment by addition of the hnRNP L-specific monoclonal antibody 4D11 (generous gift of G. Dreyfuss) (43) and precipitation with protein G plus/protein A agarose beads (Oncogene Research Products). After extensive washing of the beads, the bound proteins were eluted by incubation with SDS loading buffer and separated by SDS-12% PAGE.

**GST pull-down assay.** [<sup>35</sup>S]methionine-labeled proteins were synthesized with a rabbit reticulocyte lysate-TNT system (Promega). Chimeric GST-hnRNP L fusion proteins were expressed in *Escherichia coli* cells. Their synthesis was induced by incubation at 27°C for 6 h with 0.4 mM IPTG (isopropyl-β-D-thiogalactopyranoside). The bacteria were harvested and lysed by sonication in the presence of 3% Sarkosyl. After pelleting of the debris, the supernatants were collected, and Triton X-100 was added to 1%. GST fusion proteins were bound to glutathione Sepharose-4B (Amersham Pharmacia). Pull-down analysis was performed as described previously (19, 36). Where indicated, the TNT-synthesized proteins were pretreated with RNase A (125 μg/ml) and RNase T1 (100 U/ml) at 21°C for 30 min before addition of the GST-hnRNP L proteins. The bound proteins were separated by SDS-12% PAGE.

**Coimmunoprecipitation assay.** Cells were harvested 48 h after transfection with the indicated expression plasmids essentially as described previously (19, 36). Coimmunoprecipitations were performed in the presence of 200 μg/ml of RNase A with anti-FLAG M2 affinity gel (Sigma).

## RESULTS

**The 49-nt subelement of TK119 PPE enhances intron-independent gene expression when present in two or more tandem copies.** Liu and Mertz (32) previously identified a 119-nt sequence element within the HSV-TK gene, called a PPE, that (i) enhances cytoplasmic accumulation of intronless β-globin-

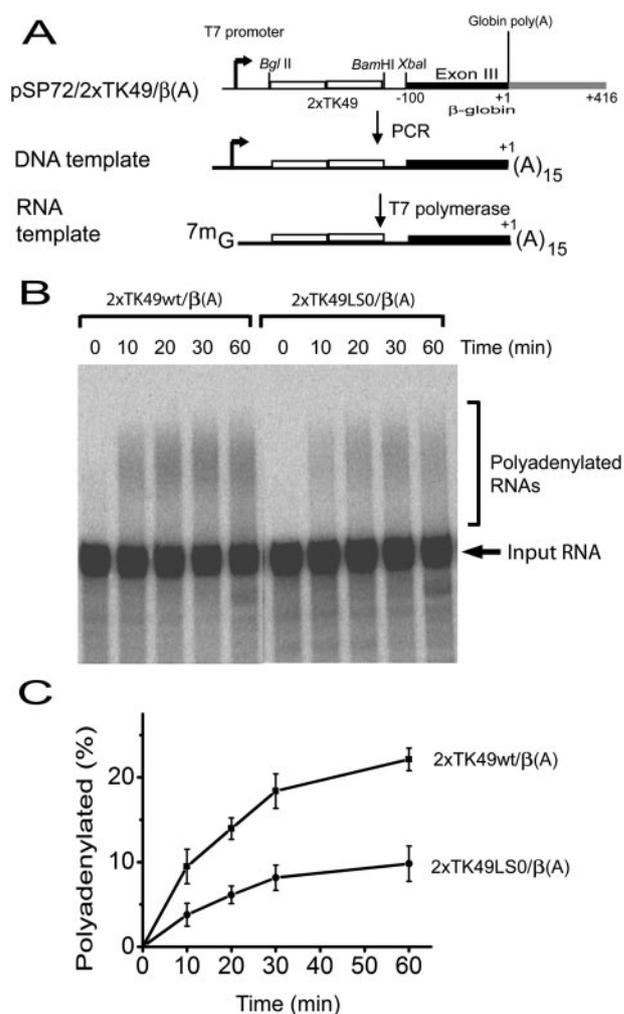


FIG. 4. Presence of 2×TK49 enhances polyadenylation in a cell-free system. (A) Structures of the plasmids employed in synthesis of the RNAs used in the cell-free polyadenylation assay. Plasmids pSP72/2×TK49wt/β(A) and pSP72/2×TK49LS0/β(A) were PCR amplified to generate double-stranded linear DNAs containing a T7 promoter, wild-type or LS0 mutant 2×TK49 PPE sequence, the 3' half of β-globin's exon III, and 15 thymidine nucleotides. Radiolabeled, 5'-m<sup>7</sup>G-capped RNAs were synthesized from these linear DNA templates with T7 RNA polymerase. The arrow indicates the site of transcription initiation. The open rectangles indicate the TK sequences inserted between the BglII and BamHI sites in pSP72. The black rectangle indicates the β-globin exon III sequences relative to β-globin's natural cleavage site for polyadenylation. The gray rectangle indicates the β-globin sequences downstream of its cleavage site for polyadenylation. (B) Autoradiogram showing enhancement of polyadenylation of PPE-containing RNA. The 5'-m<sup>7</sup>G-capped, <sup>32</sup>P-labeled RNA templates prepared as outlined in panel A were incubated in parallel with HeLa cell nuclear extract prepared as described in Materials and Methods in the presence of 1 mM ATP at 30°C for the indicated times. The resulting RNAs were phenol extracted, ethanol precipitated, and sized by electrophoresis in an 8 M urea-6% polyacrylamide gel. (C) Data averaged from four experiments similar to the one shown here indicating the means ± SEMs of the percentage of polyadenylated RNA relative to total input RNA as a function of time incubated at 30°C.

like mRNA when inserted into exon I of a cDNA version of the human β-globin gene and (ii) binds hnRNP L. However, since other proteins were also noted to bind this TK119 PPE (32), it remained unclear whether hnRNP L binding to the PPE was

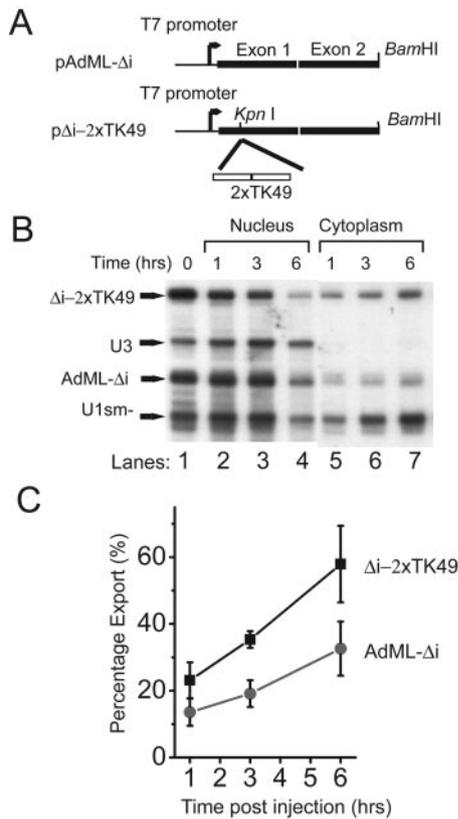


FIG. 5. 2 $\times$ TK49 PPE enhances kinetics and efficiency of nucleocytoplasmic export of intronless RNA. (A) Structures of plasmids employed in synthesis of RNAs used in *Xenopus* oocyte export experiment. The 2 $\times$ TK49 PPE (open rectangles) was inserted into pAdML- $\Delta$ i, a plasmid containing a cDNA version of the first two exons of the adenovirus major late leader region (solid rectangles) present downstream of a T7 promoter. Template DNAs were linearized with BamHI prior to transcription with T7 RNA polymerase in the presence of [<sup>32</sup>P]UTP and the cap analogue m<sup>7</sup>G(5')ppp(5')G. (B) Autoradiogram of RNA present in the nuclei versus cytoplasm of *Xenopus* oocytes at the indicated times after nuclear coinjection of a mixture of the indicated <sup>32</sup>P-labeled RNAs. The oocytes were manually dissected at the indicated times postinjection as described previously (42). The RNAs were purified and sized by electrophoresis in an 8 M urea-6% polyacrylamide gel. (C) Summary of data (means  $\pm$  SEMs) obtained from three pools of oocytes analyzed in parallel as described in panel B. Percentage export was calculated as the amount of the indicated RNA in the cytoplasm relative to the total amount of that RNA in the nucleus plus cytoplasm at the indicated time.

sufficient to recruit to the RNA all of the factors needed to enhance stabilization, 3' end processing, and nucleocytoplasmic export. A computer-generated secondary structure analysis of this 119-nt PPE predicted the possible existence of two stem-loops (Fig. 1A). The boxed region indicates bases previously shown by base substitution mutational analysis to include ones important for this PPE both (i) to enhance intron-independent gene expression (32) and (ii) to bind hnRNP L (30, 32). This region contains a CA repeat. CA repeat regions have recently been shown by others to preferentially bind hnRNP L (14, 15, 21, 22, 48). We speculated that the left stem-loop together with the CA repeat region might be sufficient to provide PPE function via binding hnRNP L.

One, two, and four copies of a 49-nt sequence encompassing

this subregion of TK119 PPE were inserted into exon I of the cDNA variant of the human  $\beta$ -globin gene (Fig. 1B). Each of these plasmids, along with controls, was transiently cotransfected in parallel into CV-1PD cells together with pRSV-Tori, a plasmid encoding the SV40 large T antigen. The presence of the latter plasmid results in replication of the test plasmid to high copy number. Therefore, the resulting  $\beta$ -globin-like RNAs are synthesized at high levels and can be readily analyzed by quantitative S1 nuclease mapping with the probes shown in Fig. 1C and 3A.

Precise deletion of the human  $\beta$ -globin gene's two introns does not significantly affect the gene's rate of transcription as determined by nuclear run-on assays (5, 46). Nevertheless, transcripts synthesized from this cDNA variant are highly defective in stabilization, cleavage for polyadenylation, and cytoplasmic accumulation compared to ones synthesized from the wild-type parental gene (3, 5, 32, 46, 50) (Fig. 1D, lane 4 versus lane 2; see also Fig. 3B, lanes 3 and 4 versus lanes 1 and 2). Insertion of a nearly full-length copy of the coding region of the HSV-TK gene, containing several PPEs (3, 41), enhances approximately 30-fold cytoplasmic accumulation of  $\beta$ -globin-like RNA synthesized from the intronless gene (32) (Fig. 1D, lane 6 versus lane 4). Insertion of a single copy of TK119 PPE enhances cytoplasmic accumulation sixfold (32) (Fig. 1D, lane 8 versus lane 4). However, insertion of a single copy of TK49 at the same site failed to enhance cytoplasmic accumulation of RNA synthesized from the intronless gene (Fig. 1D, lane 10 versus lane 4).

A likely explanation for these findings is that efficient enhancement requires the presence of multiple PPE-like elements. Consistent with this hypothesis was the finding that insertion of two copies of TK49 in tandem enhanced cytoplasmic accumulation of globin-like RNA approximately threefold (Fig. 1D, lane 12). Insertion of four copies of TK49 in tandem enhanced accumulation even further, up to approximately 12-fold (Fig. 1D, lane 14). Similar results were observed when these plasmids were transfected into COS-M6 cells, a derivative of CV-1 cells transformed by SV40 large T antigen (data not shown). Thus, this 49-nt subregion of the TK119 PPE can indeed function fairly well to enhance expression of an intronless gene, but only when present in multiple copies. TK119 PPE also functions better when present in two copies rather than one (32).

To rule out the possibility that the RNA had accumulated in the cytoplasm due to cryptic splicing, we performed RT-PCR analysis with primers corresponding to sequences located near the very 5' and 3' ends of the  $\beta$ -globin mRNA. As expected, the globin-like RNAs accumulated in the cytoplasm of cells transfected with these chimeric constructs were similar in size to their corresponding template DNAs (Fig. 1E). No smaller bands corresponding to cryptically spliced products were present at detectable levels. Primer extension analysis did not indicate any utilization of alternative, upstream 5' ends (data not shown). Cryptic polyadenylation followed by splicing to sequences situated downstream of the natural cleavage site for polyadenylation also could not have accounted for a significant portion of the accumulated RNA, since the relative amounts of globin-like RNA present in the cytoplasm as quantified with this exon I/II 5'-end-labeled probe were similar to the relative amounts as quantified with an exon III 3'-end-labeled probe

(Fig. 3). Thus, at most a small percentage of the cytoplasmic RNA quantified in these experiments could have arisen as a consequence of cryptic splicing.

Thus, we conclude that two copies of TK49 are sufficient to enhance expression in mammalian cells of an intronless variant of this highly intron-dependent gene.

**2×TK49 PPE enhances stabilization of pre-mRNA.** Insertion of the PPE also enhanced nuclear accumulation of the intronless  $\beta$ -globin pre-mRNA (Fig. 1D, lanes 7, 11, and 13 versus lane 3). Precise deletion of the  $\beta$ -globin gene's two introns does not significantly affect either the rate of transcription of this gene (5, 46) or the half-life of the resulting mRNA (17, 34). Thus, stabilization of the primary transcript is likely one of several steps necessary for mRNA biogenesis that was enhanced by the PPE's presence. Consistent with this hypothesis is the finding that the PPE functions well when located near the 5' end of the primary transcript but poorly, if at all, when inserted in  $\beta$ -globin's exon III near the site of polyadenylation (13a, 32; D. Willard and J. E. Mertz, unpublished).

**Regions of 2×TK49 PPE necessary for function.** To identify regions of this 2×TK49 PPE necessary to enhance intronless gene expression, we constructed several cluster base substitution mutant variants of it (Fig. 2A). Figure 2B shows a computer-predicted secondary structure of 1×TK49. TK119LS0 is a previously described cluster base substitution mutant variant of TK119 known to be defective in both *in situ* functioning and hnRNP L binding (32). The base substitution mutations present in TK119LS0 were placed into 2×TK49 to generate 2×TK49LS0. These mutations grossly perturbed the computer-predicted structure (data not shown). We also constructed 2×TK49LS0com, a variant of 2×TK49LS0 containing four additional base substitution mutations which, in theory, restore the original stem-loop structure but retain the mutations in the adjacent CA repeats. Last, 2×TK49m was constructed. This variant of 2×TK49 contains three base substitution mutations in the loop region while maintaining the stem-loop structure and CA repeat region.

Each of these variants of 2×TK49 was inserted into plasmid p $\beta$ - $\beta$ 1(-)2(-) at the NcoI site in exon I. These plasmids were cotransfected in parallel into CV-1PD cells together with pRSV-Tori and analyzed as described above. The 2×TK49 variants containing the LS0 and LS0com mutations failed to enhance cytoplasmic accumulation of  $\beta$ -globin RNA (Fig. 2C, lanes 8 and 10). However, the 2×TK49m variant enhanced intronless gene expression as well as the parental 2×TK49wt did (Fig. 2C, lane 12 versus lane 4). Similar results were observed when these plasmids were transfected into COS-M6 cells (data not shown). Thus, we conclude that specific sequences within 2×TK49, including the CA repeat region, and not its secondary structure per se are important for enhancing intronless gene expression.

**2×TK49 PPE enhances proper 3' end formation.** Previous reports indicated that transcripts synthesized from intronless variants of the  $\beta$ -globin gene are also defective in 3' end processing (3, 5, 32, 46, 50). Insertion of TK119 PPE enhances accumulation of intronless  $\beta$ -globin-like RNA in part by enhancing the efficiency of proper 3' end formation (17, 32). Thus, 2×TK49 PPE probably also enhances proper 3' end processing of intronless  $\beta$ -globin transcripts.

To test this prediction, we analyzed the  $\beta$ -globin-like RNAs

from the above experiment by quantitative S1 nuclease mapping with an exon III 3'-end-labeled probe (Fig. 3A). Most of the  $\beta$ -globin RNA accumulated in cells transfected with p $\beta$ - $\beta$ 1(+)-2(+), the intron-containing genomic version of the human  $\beta$ -globin gene, was properly processed at the  $\beta$ -globin polyadenylation site (Fig. 3B, lanes 1 and 2). As expected, cells transfected with p $\beta$ - $\beta$ 1(-)-2(-), the intronless cDNA variant, contained as much uncleaved primary transcript as did cells transfected with p $\beta$ - $\beta$ 1(+)-2(+), but little RNA that had been cleaved at the  $\beta$ -globin polyadenylation site (Fig. 3B, lanes 3 and 4). Insertion of the full-length TK coding region, 1×TK119, 2×TK49, 2×TK49m, or 4×TK49 into exon I of p $\beta$ - $\beta$ 1(-)-2(-) led to a dramatic increase in the accumulation of  $\beta$ -globin-like RNA properly cleaved at the  $\beta$ -globin polyadenylation site (Fig. 3B, lanes 5 to 8 and 11 to 14, and C, lanes 7 and 8). On the other hand, cells transfected with p $\beta$ -1×TK49- $\beta$ 1(-)-2(-) or either of the stem/CA repeat region substitution mutants p $\beta$ -2×TK49LS0- $\beta$ 1(-)-2(-) and p $\beta$ -2×TK49LS0com- $\beta$ 1(-)-2(-) failed to accumulate RNA properly cleaved for polyadenylation above the level observed in cells transfected with the parental p $\beta$ - $\beta$ 1(-)-2(-) (Fig. 3B, lanes 9 and 10, and C, lanes 3 to 6, versus Fig. 3B, lanes 3 and 4). The relative amounts of properly 3'-end-cleaved  $\beta$ -globin RNA measured with this probe were similar to the relative amounts of  $\beta$ -globin RNA measured with the 5'-end-labeled probe (Fig. 1 and 2 versus Fig. 3). Thus, this 3'-end-specific probe was not failing to detect significant quantities of  $\beta$ -globin-like RNA that, in theory, might have been cleaved elsewhere for polyadenylation, e.g., within adjacent vector sequences, and then spliced. The small amounts of uncleaved RNA observed in the cytoplasm here as well as by others (34) were likely due to a combination of (i) cross-contamination of cytoplasm with nuclear material, measured as described in Materials and Methods, (ii) low levels of cryptic polyadenylation, and (iii) unprocessed RNA being inefficiently exported to the cytoplasm.

Also noteworthy is the finding that the total amounts of uncleaved RNA accumulated in the cells were similar for all of the samples regardless of the presence of introns or PPE elements (Fig. 3). Thus, the apparent defect in cleavage efficiency observed with the PPE-less, intronless transcripts could be due, at least in part, to degradation of cleaved RNAs that failed to be polyadenylated (see below) rather than to their not being cleaved with normal efficiency. Regardless, we conclude that insertion of two or four copies of the TK49 PPE can significantly rescue *in situ* the defect in 3' end processing of intronless  $\beta$ -globin transcripts.

**2×TK49 PPE enhances polyadenylation.** To examine directly whether the presence of the 2×TK49 PPE enhances polyadenylation of intronless transcripts, we measured the kinetics of polyadenylation of 2×TK49 versus 2×TK49LS0-containing RNA in a cell-free polyadenylation system. Plasmid pSP72/2×TK49wt/ $\beta$ (A) contains an insertion of 2×TK49 between a T7 promoter and the 3' end of the human  $\beta$ -globin gene including the site of cleavage for polyadenylation (Fig. 4A). Plasmid pSP72/2×TK49LS0/ $\beta$ (A) is a variant of it containing the LS0 cluster base substitution mutations. These two plasmid DNAs were used as templates to generate by PCR amplification double-stranded linear DNAs containing the T7 promoter, the wild-type or LS0 2×TK49 PPE sequence, the 3'

half of  $\beta$ -globin's exon III, and 15 thymidine nucleotides directly adjacent to  $\beta$ -globin's natural site for polyadenylation (Fig. 4A). Radiolabeled, 5'-m<sup>7</sup>G-capped RNAs were synthesized from the latter template DNAs with T7 RNA polymerase. These RNAs were incubated in parallel with a HeLa cell nuclear extract under conditions for addition of poly(A)<sub>n</sub> to the 3' end of the RNA and sized by electrophoresis in a denaturing polyacrylamide gel. Poly(A)<sub>n</sub> was added to the 2 $\times$ TK49wt-containing RNA with approximately twofold-faster kinetics and greater efficiency than it was to the 2 $\times$ TK49LS0-containing RNA (Fig. 4B and C). 2 $\times$ TK49wt-containing template RNA with a 3' end at +12 relative to the natural site of polyadenylation rather than 15 adenosine nucleotides was also polyadenylated with approximately twofold-faster kinetics and greater efficiency than its corresponding 2 $\times$ TK49LS0-containing RNA (data not shown) (13). Thus, the presence of 2 $\times$ TK49 does indeed enhance the kinetics and efficiency of addition of poly(A)<sub>n</sub> to the 3' end of intronless  $\beta$ -globin-like RNAs.

**2 $\times$ TK49 PPE enhances nucleocytoplasmic export.** To examine whether the presence of 2 $\times$ TK49 can also enhance nucleocytoplasmic export of an intronless mRNA, we inserted 2 $\times$ TK49 near the 5' end of the transcribed region of pAdML- $\Delta$ i. Plasmid pAdML- $\Delta$ i contains a cDNA version of the first two exons of the adenovirus major late leader region adjacent to a T7 promoter (35, 51). 5'-m<sup>7</sup>G-capped, <sup>32</sup>P-labeled, intronless RNA was synthesized in vitro from BamHI-cleaved pAdML- $\Delta$ i DNA and the PPE-containing variant p $\Delta$ i-2 $\times$ TK49 (Fig. 5A). These RNAs were coinjected into the nuclei of *Xenopus* oocytes along with U1<sub>sm-</sub> and U3 RNAs as positive and negative controls for export, respectively (42). Nuclei and cytoplasm were manually isolated after incubation at 18°C for the indicated times. The RNAs were purified and separated by electrophoresis in a denaturing polyacrylamide gel. As expected (42), the U1<sub>sm-</sub> RNA was efficiently exported to the cytoplasm, while the U3 RNA was retained in the nucleus (Fig. 5B). Also as expected (35), the intronless adenovirus major late leader region RNA was slowly and inefficiently exported to the cytoplasm (Fig. 5B). The presence of the 2 $\times$ TK49 PPE enhanced approximately twofold both the kinetics and efficiency of export of the intronless adenovirus RNA to the cytoplasm (Fig. 5B; summarized in Fig. 5C). Thus, we conclude that the presence of a 2 $\times$ TK49 PPE can also enhance nucleocytoplasmic export of intronless RNA.

**2 $\times$ TK49 PPE specifically binds only hnRNP L.** To determine what protein(s) binds the 2 $\times$ TK49 PPE, we incubated radiolabeled, 5'-m<sup>7</sup>G-capped 2 $\times$ TK49 wild-type and mutant variant RNAs along with 1 $\times$ TK119 wild-type and mutant variant RNAs with a HeLa cell nuclear extract. Afterward, the bound proteins were UV-cross-linked to the RNAs, and the unprotected RNAs were degraded by digestion with RNases A and T1. The RNA-protein complexes were sized by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). As previously reported (32), several proteins were found bound to 1 $\times$ TK119 RNA, including the 64-kDa protein shown by several criteria to be hnRNP L (32) (Fig. 6B, lane 2). The PPE-defective mutant 1 $\times$ TK119LS0 failed to bind the 64-kDa protein, while still binding the one other abundant protein, approximately 24-kDa in size, that bound 1 $\times$ TK119 (Fig. 6B, lane 3 versus lane 2). Thus, this 24-kDa protein likely does not contribute to

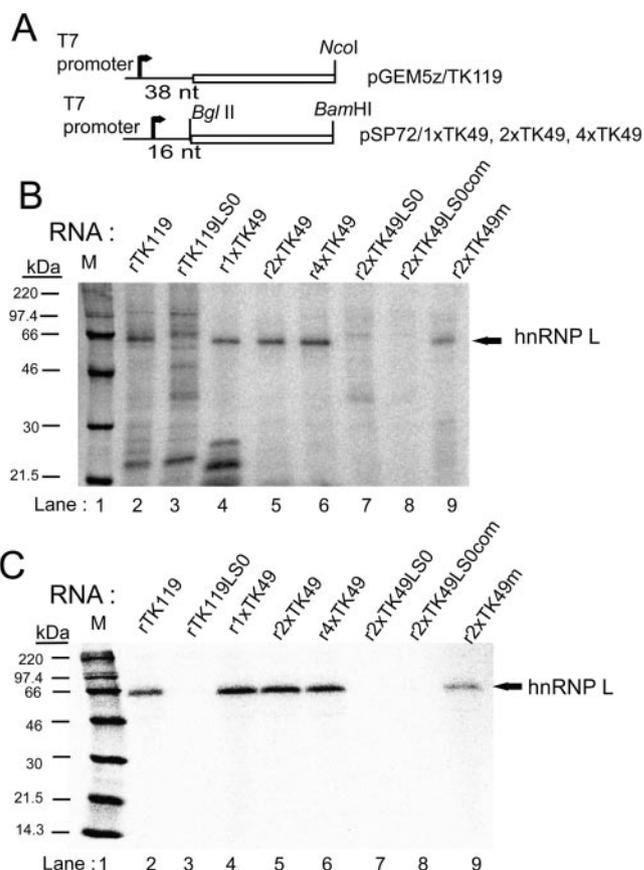


FIG. 6. 2 $\times$ TK49 is bound only by hnRNP L. (A) Structures of the plasmids used to synthesize the RNA probes employed in UV-cross-linking analysis. Symbols are the same as described in the legend to Fig. 4A. (B) Autoradiogram of SDS-PAGE analysis of the proteins that UV-cross-linked with the indicated variants of the HSV-TK PPE. 5'-m<sup>7</sup>G-capped, <sup>32</sup>P-labeled RNAs were synthesized in parallel with T7 RNA polymerase, incubated with 20  $\mu$ g of HeLa cell nuclear extract, exposed to UV, and incubated with RNases A and T1. Afterward, the proteins were sized by SDS-12% PAGE. Lane 1 contains <sup>14</sup>C-labeled Rainbow markers (Amersham Pharmacia). (C) Autoradiogram of the UV-cross-linked proteins from panel B sized by SDS-12% PAGE after digestion with RNases A and T1 and immunoprecipitation with the hnRNP L-specific monoclonal antibody 4D11.

1 $\times$ TK119's functional activities. As expected, the 1 $\times$ TK49, 2 $\times$ TK49, 4 $\times$ TK49, and 2 $\times$ TK49m RNAs, but not the 2 $\times$ TK49LS0 and 2 $\times$ TK49LS0com mutant RNAs, also bound the 64-kDa protein (Fig. 6B, lanes 4 to 9). Moreover, there were no other proteins bound to the 2 $\times$ TK49 and 4 $\times$ TK49 RNAs at detectable levels (Fig. 6B, lanes 5 and 6). Immunoprecipitation of the UV-cross-linked protein with the hnRNP L-specific antibody 4D11 followed by SDS-PAGE confirmed that the 64-kDa protein was, indeed, hnRNP L (32) (Fig. 6C). Immunoprecipitation with an hnRNP K-specific antibody failed to pull down this radiolabeled, 64-kDa band (32). Taken together with the findings above, we conclude that binding of hnRNP L is both necessary and sufficient to bring to PPE-containing RNAs the cellular factors involved in PPE-mediated enhancement of stability, proper 3' end formation, and cytoplasmic accumulation of intronless transcripts.

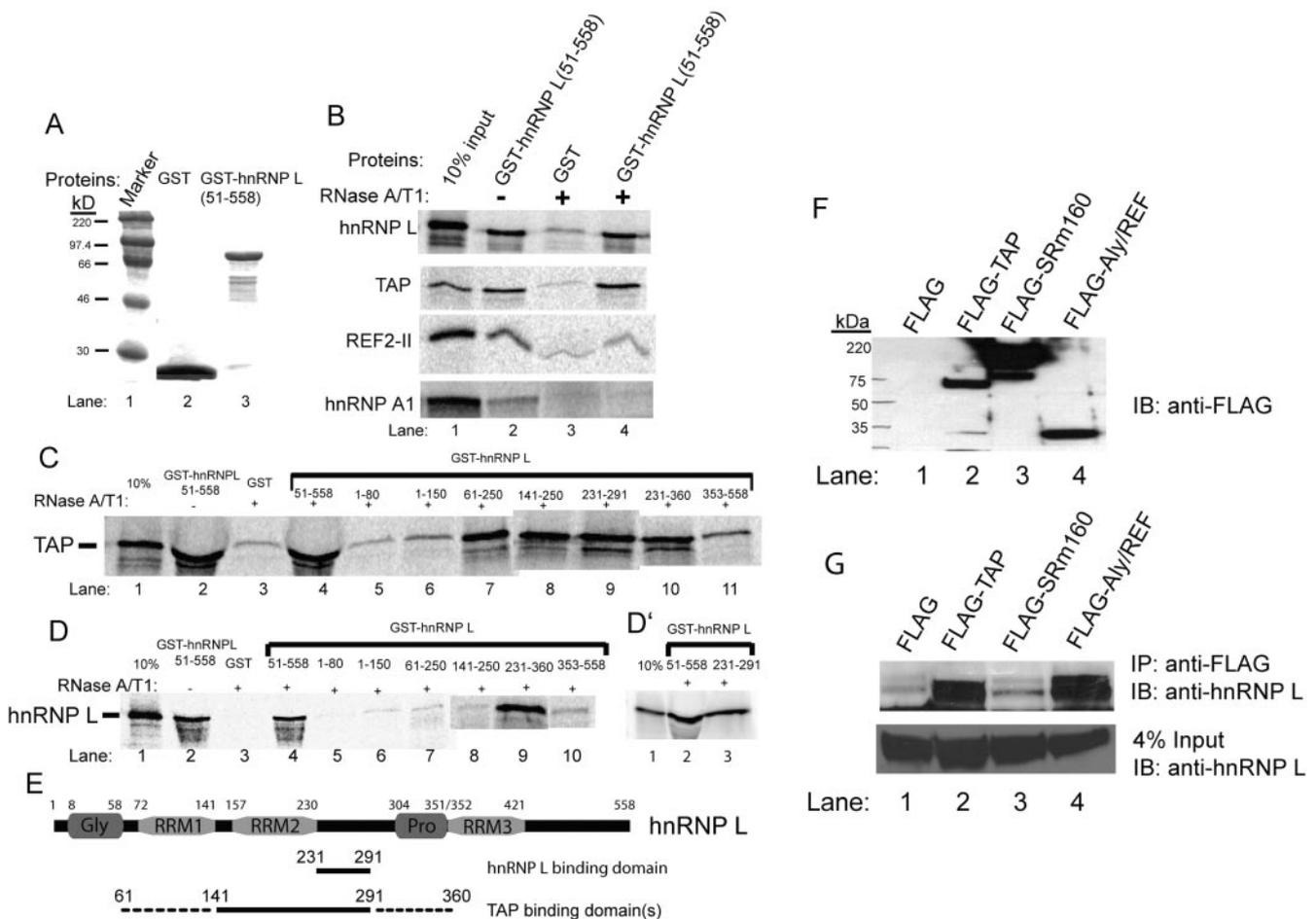


FIG. 7. hnRNP L binds itself, TAP, and Aly/REF. (A) Coomassie brilliant blue staining of an SDS-PAGE analysis of GST and GST-hnRNP L fusion protein expressed in *E. coli* and purified by glutathione Sepharose-4B. (B) Autoradiogram of GST pull-down analysis. [<sup>35</sup>S]methionine-labeled, nearly full-length hnRNP L, TAP, REF2-II, and hnRNP A1 were synthesized with a rabbit reticulocyte lysate system. Lysates were pretreated with RNases A and T1 (+) or not pretreated (–) before addition of the indicated GST fusion protein. The proteins that bound were sized by SDS-12% PAGE and exposed to a PhosphorImager. (C and D) Domains of hnRNP L involved in binding TAP and itself. GST pull-down assays were performed with equimolar amounts of each of the indicated deletion variants of hnRNP L synthesized in *E. coli* and <sup>35</sup>S-labeled TAP (C) or full-length hnRNP L (D). (E) Schematic showing known domains of hnRNP L. The numbers indicate positions of amino acid residues. RRM, RNA recognition motif; Gly, glycine-rich region; Pro, proline-rich region. Below the schematic is indicated the tentative mapping of the hnRNP L- and TAP-binding domains as determined from the data presented here. (F) FLAG-tagged proteins present in human 293T cells 48 h after transient transfection with expression plasmids encoding the indicated FLAG-tagged proteins. Whole-cell extracts were electrophoresed in an SDS-12% polyacrylamide gel and probed by immunoblotting with the anti-FLAG M2 antibody. (G) Immunoblot showing coimmunoprecipitation of endogenous hnRNP L with FLAG-tagged TAP and Aly/REF. The whole-cell extracts from panel F were incubated with RNase A followed by immunoprecipitation of the FLAG-tagged protein-containing complexes with anti-FLAG M2 affinity gel, electrophoresis in an SDS-12% polyacrylamide gel, and immunoblotting with the anti-hnRNP L monoclonal antibody 4D11. The bottom panel represents 4% of the amount of hnRNP L present in each whole-cell extract prior to immunoprecipitation.

**hnRNP L binds TAP and Aly/REF in vitro.** One hypothesis is that hnRNP L facilitates the nucleocytoplasmic export of PPE-containing RNAs via a novel pathway. Alternatively, hnRNP L could act as an adaptor protein to recruit to intronless PPE-containing mRNAs export factors normally recruited to intron-containing mRNAs during exon definition and intron excision. If the latter hypothesis is valid, hnRNP L might associate with TAP or Aly/REF, cellular factors involved in mRNA export that are recruited to mRNAs during splicing (27, 28). To test this possibility, we performed GST pull-down assays with these proteins and recombinant hnRNP L. GST-hnRNP L fusion protein was synthesized in *E. coli* cells and purified with glutathione Sepharose-4B beads (Fig. 7A). Ra-

diolabeled, full-length hnRNP L, TAP, and REF2-II were synthesized with a rabbit reticulocyte lysate system in the presence of [<sup>35</sup>S]methionine. Where indicated, the lysates were pretreated with RNases A and T1 before incubation with the GST-hnRNP L fusion protein to exclude possible artifacts due to these RNA-binding proteins simply binding the same RNA molecules rather than each other.

hnRNP L was found to bind both itself and TAP fairly well (Fig. 7B, lane 2), associations that were insensitive to treatment with RNases A and T1 (Fig. 7B, lane 4). hnRNP L also bound Aly/REF (Fig. 7B, lane 2). This association was slightly reduced by pretreatment with RNases A and T1, but not down to the level observed with the GST-only control (Fig. 7B, lane

4 versus 2). On the other hand, the association between hnRNP L and hnRNP A1, another RNA-binding protein, was completely disrupted by a similar pretreatment with RNases A and T1 (Fig. 7B, lane 4 versus 2). Thus, hnRNP L likely binds TAP, Aly/REF, and itself but not hnRNP A1.

To begin to determine the specificity of these protein-protein interactions, a series of deletion variants of GST-hnRNP L were constructed. These proteins were synthesized in *E. coli* and used in pull-down assays as described above. The domains of hnRNP L binding TAP localized to the amino acid 141-to-291 region of hnRNP L (Fig. 7C, lanes 8 and 9). Importantly, they were partially nonoverlapping with the domain that bound hnRNP L, amino acid residues 231 to 291 (Fig. 7D and D'; summarized in Fig. 7E). For example, the amino acid 141-to-250 region of hnRNP L bound TAP well yet failed to bind hnRNP L. Given this noncoincidence of binding to hnRNP L, TAP was not simply binding a sticky region of hnRNP L. Rather, these data indicate that either the TAP-binding domain is located within the amino acid 231-to-250 region of hnRNP L or there exists more than one TAP-binding domain on hnRNP L. Regardless, these findings confirm that hnRNP L specifically binds TAP, suggesting a mechanism by which hnRNP L's binding to PPEs facilitates nucleocytoplasmic export of PPE-containing intronless RNAs. Thus, we conclude that hnRNP L acts as an adaptor protein to recruit TAP and, likely, Aly/REF to PPE-containing mRNAs, factors normally involved in the export of intron-containing mRNAs.

**hnRNP L coimmunoprecipitates with TAP and Aly/REF.** To begin to determine the physiological relevance of hnRNP L binding TAP and Aly/REF, we examined whether hnRNP L coimmunoprecipitates with these proteins from extracts of mammalian cells. FLAG-tagged TAP and Aly/REF were transiently expressed in human 293T cells. FLAG and FLAG-tagged SRm160, another RNA-binding protein frequently found in association with exon-exon junction complexes (26), were also expressed in 293T cells as negative controls. Whole-cell extracts were prepared (Fig. 7F). The FLAG-tagged protein-containing complexes were immunoprecipitated with anti-FLAG antibodies in the presence of RNase A and separated by SDS-PAGE. Immunoblot analysis with the anti-hnRNP L monoclonal antibody 4D11 indicated that hnRNP L was present at high levels in the TAP and Aly/REF immunoprecipitates (Fig. 7G, lanes 2 and 4, respectively) but not in the FLAG and SRm160 immunoprecipitates (Fig. 7G, lanes 1 and 3, respectively). The approximately 2% of hnRNP L associated with TAP and Aly/REF is quite significant, given that hnRNP L is a highly abundant protein in cells (43). The fact that hnRNP L bound REF but not SRm160 implies that hnRNP L associated with only some specific subcomponents of exon-exon junction complexes, not all of them. Thus, we conclude that hnRNP L probably associates within mammalian cells with TAP and Aly/REF but not SRm160.

## DISCUSSION

We determined here mechanisms by which an HSV-TK PPE enhances cytoplasmic accumulation of intronless RNAs. First, we identified a 49-nt subelement of the previously reported TK119 PPE that can function efficiently in stabilization, polyadenylation, and export of an intronless transcript when present near the 5' end of the RNA in two or more copies (Fig.

1 and 3). This 2×TK49 PPE bound only hnRNP L (Fig. 6), with a strong positive correlation existing between hnRNP L binding (Fig. 6) and enhancement of proper 3' end formation in situ (Fig. 3), polyadenylation in vitro (Fig. 4), and cytoplasmic accumulation of the RNA (Fig. 2, 3, and 5). Thus, binding of hnRNP L was both necessary and sufficient to recruit all of the posttranscriptional processing and export factors needed to enable this PPE to perform all of its multiple functions in intronless mRNA biogenesis. We further showed that hnRNP L specifically bound the mRNA export factor TAP and the exon-exon junction complex protein Aly/REF both in vitro and in vivo (Fig. 7), with the presence of the 2×TK49 PPE increasing the kinetics and efficiency of nucleocytoplasmic export of an intronless RNA in *Xenopus* oocytes (Fig. 5). Thus, we conclude that this PPE enhances intron-independent gene expression via binding hnRNP L, with hnRNP L functioning, in part, as an adaptor protein to recruit Aly/REF- and TAP/p15-containing complexes to PPE-containing RNAs to mediate their export.

**PPE-like elements can substitute for introns.** β-Globin primary transcripts lacking both of β-globin's introns are defective in RNA stability, proper 3' end formation, and nucleocytoplasmic export (3, 5, 33, 46, 50) (Fig. 1). Insertion of PPEs dramatically enhanced both cytoplasmic accumulation and 3' end processing, doing so in the absence of intron excision (Fig. 1 and 3). Given that introns function in modulating pre-mRNA stability, 3' end cleavage and polyadenylation, nucleocytoplasmic export, surveillance, and translation, we conclude that PPE-like elements can substitute for introns to provide at least some of these functions (Fig. 1 through 5).

**Sequences necessary for PPE function.** Previous deletional and mutational analyses indicated that the central region of TK119 PPE, including the CA repeat region, was important for binding of hnRNP L (30, 32). Additional mutational analysis presented here indicates that the 3' end of TK49 PPE, including its CA repeat region, was crucial for 2×TK49 PPE to function in transient transfection assays (Fig. 2). Regeneration of the putative stem with complementary base substitution mutations failed to restore PPE's functional activities (Fig. 2 and 3). On the other hand, mutations in the loop region did not disrupt 2×TK49 PPE's ability to enhance intron-independent gene expression (Fig. 2 and 3). Taken together with the findings that hnRNP L binds CA repeat regions (14, 15, 21, 22, 48) and 2×TK49 binds only hnRNP L (Fig. 6), these data indicate that the CA repeat region is probably essential for PPE functioning.

The 1×TK49 PPE bound hnRNP L efficiently (Fig. 6) yet failed to enhance intron-independent gene expression (Fig. 1). The most likely explanation for this finding is that the PPE needs to be present in two or more copies to function. Consistent with this conclusion are the findings that (i) 2×TK119 functions better than 1×TK119 (32) and (ii) 4×TK49 functions even better than 2×TK49 (Fig. 1D, lanes 11 to 14). Quite likely, the presence of multiple hnRNP L binding sites in the RNA enables stable, synergistic protein-protein interactions among multiple hnRNP L molecules, the proteins with which hnRNP L interacts (Fig. 7 and 8), and in turn, their interacting factors. Likewise, the constitutive transport element (2, 10), posttranscriptional regulatory element (7, 8), and H2a element (18) all need to be present in two or more copies to function.

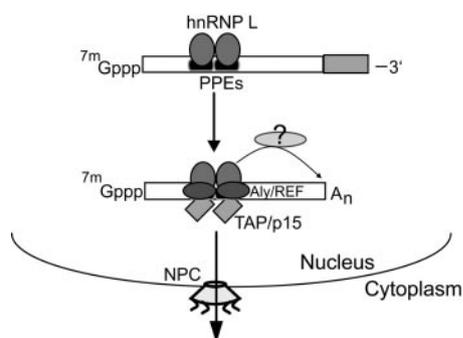


FIG. 8. Proposed role of hnRNP L in mediating nucleocytoplasmic export of PPE-containing intronless mRNAs. hnRNP L functions as an adaptor molecular, binding both PPEs and the mRNA export factors TAP and Aly/REF, bringing multiple TAP-p15 complexes to RNAs containing multiple PPEs to efficiently mediate their export. Likely, hnRNP L also associates with other as-yet-unidentified factors, recruiting the 3' end cleavage/polyadenylation machinery to the RNA as well.

#### hnRNP L binding is sufficient for PPE's multiple functions.

Previously, a correlation was shown to exist between hnRNP L binding and PPE function (32). Here, we showed that hnRNP L is the only protein that binds 2×TK49 PPE at a detectable level (Fig. 6B) and that all mutants defective in hnRNP L binding are also defective in all of PPE's functions (Fig. 2, 3, 4, and 6). Thus, hnRNP L binding is both necessary and sufficient to bring in all of the posttranscriptional processing and export factors for PPE-containing RNAs to be efficiently stabilized, polyadenylated, and exported to the cytoplasm, i.e., hnRNP L is performing multiple functions in biogenesis of this intronless transcript. Consistent with this conclusion is hnRNP L's already documented functions in regulation of splicing (22), stabilization (15, 21, 48), and translation (14).

**hnRNP L recruits TAP and Aly/REF.** Here, we provide the first evidence that hnRNP L enhances intronless mRNA biogenesis, in part, by recruiting TAP and Aly/REF (Fig. 7). TAP is a general mRNA export factor in mammalian cells (23, 29, 45). It is usually recruited by Aly/REF to the exon-exon junction complexes bound to mRNAs following excision of introns (27, 28). Some viral RNA elements, such as Mason-Pfizer monkey virus constitutive transport element, bind TAP directly, exporting their mRNA without intron excision (12). The mouse histone H2a element enhances mRNA export without splicing via binding SR proteins, such as 9G8 and SRp20, which function as adaptors for the binding of TAP (19, 20). HSV-1's ICP27 interacts both directly with TAP and indirectly with the RNA export adaptor Aly/REF to mediate nucleocytoplasmic export of intronless HSV-1 RNAs (4). Here, we show that hnRNP L binds both TAP and Aly/REF, enhancing the nucleocytoplasmic export of PPE-containing intronless mRNAs in the absence of introns and their excision. Given that TAP binds Aly/REF, it remains unclear whether hnRNP L directly binds both or only one of these proteins. Taking together all of these findings, we conclude that hnRNP L binding to PPE-containing RNAs enhances their nucleocytoplasmic export via functioning as an adaptor for binding TAP and Aly/REF (Fig. 8). Remaining unknown is whether other CA repeat-containing transcripts also load TAP by binding hnRNP L.

**hnRNP L enhances polyadenylation.** We also showed here that the presence of the PPE enhances proper 3' end forma-

tion and polyadenylation (Fig. 3 and 4). We speculate that hnRNP L likely recruits the polyadenylation machinery directly or indirectly to PPE-containing transcripts, thereby facilitating 3' end formation regardless of introns (Fig. 8). A precedent for such a dual-functioning hnRNP is Nab2p, a nuclear poly(A)-binding protein involved in both control of poly(A) tail length and nuclear export of mRNA in yeast (16). Recently, we identified an RNA element within the transcribed region of the cellular intronless *c-Jun* gene, named CJE for *c-Jun* enhancer, which also enhances both 3' end formation and cytoplasmic accumulation of intronless transcripts (13a). As yet unknown is the *trans*-acting factor(s) with which the CJE interacts and the mechanisms by which it functions. Likely, most, if not all, intronless pre-mRNAs contain sequence elements that bind hnRNPs, thereby recruiting cofactors that facilitate stabilization, proper 3' end formation, and nucleocytoplasmic export by mechanisms analogous or identical to the ones normally employed by intron-containing RNAs.

**hnRNP L likely also enhances stabilization of the pre-mRNA.** cDNA variants of intron-requiring genes are defective in stabilization of the primary transcript as well as polyadenylation and nucleocytoplasmic export (3, 5, 32, 46, 50) (Fig. 1). We confirm here the previous report (24) that hnRNP L binds itself (Fig. 7). We speculate that hnRNP L binding to itself and other hnRNPs likely contributes to stabilization of the primary transcript as well as synergistic recruitment to the RNA of factors involved in processing and export. Consistent with this hypothesis are the findings that (i) the PPE enhances nuclear accumulation of pre-mRNA synthesized from the intronless variant of the  $\beta$ -globin gene (Fig. 1 to 3) and (ii) the PPE functions well near the 5' end of the primary transcript but poorly, if at all, when located near the site of polyadenylation (13a, 32; D. Willard and J. E. Mertz, unpublished). Hui et al. (21) also found that hnRNP L stimulates stability by binding to a CA repeat region in endothelial cell nitric oxide synthase pre-mRNA. Moreover, the human vascular endothelial growth factor mRNA contains a CA-rich element in the 3' untranslated region that binds hnRNP L and mediates mRNA stability under hypoxic growth conditions (48). Thus, these RNA stability effects are likely all due to recognition of CA-rich RNA elements by hnRNP L. However, given coupling with other steps in mRNA biogenesis, it would be very difficult to perform a definitive and direct experiment to validate this hypothesis.

In summary, we conclude that the binding of hnRNP L is both necessary and sufficient to recruit all of the cellular factors needed to enable stabilization, 3' end cleavage/polyadenylation, nucleocytoplasmic export, and cytoplasmic accumulation of PPE-containing intronless transcripts, doing so, in part, via association with the mRNA export factor TAP and the exon-exon junction complex protein Aly/REF. Importantly, these findings indicate that many, if not most, intronless pre-mRNAs are likely processed via pathways similar to the ones employed in the processing of intron-containing ones, with factors such as hnRNP L and 9G8 (19) substituting for exon-exon junction complexes in recruiting the processing machineries to the pre-mRNAs.

#### ACKNOWLEDGMENTS

We thank Robin Reed, Sung Key Jang, Yingqun Huang, Elisa Izauralde, Stanley Lemon, Maurice Swanson, Benjamin J. Blencowe, and

Edward Seto for providing plasmids. We thank Gideon Dreyfuss for plasmid and antibody. We especially thank Elsebet Lund for assistance with the *Xenopus* oocyte experiment. We thank James Dahlberg, Jeff Ross, Paul Lambert, Jeff Habig, and members of the Mertz laboratory for helpful discussions and comments on the manuscript.

This work was supported by Public Health Service research grants CA22443 and CA14520 from the National Cancer Institute and funds to the McArdle Laboratory for Cancer Research.

## REFERENCES

- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1987. Preparation of nuclear and cytoplasmic extracts from mammalian cells, p. 12.1.1–12.1.9. *In* V. B. Chanda (ed.), *Current protocols in molecular biology*. John Wiley & Sons, Inc., New York, N.Y.
- Bray, M., S. Prasad, J. W. Dubay, E. Hunter, K.-T. Jeang, D. Rekosh, and M.-L. Hammarskjöld. 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc. Natl. Acad. Sci. USA* **91**:1256–1260.
- Buchman, A. R., and P. Berg. 1988. Comparison of intron-dependent and intron-independent gene expression. *Mol. Cell. Biol.* **8**:4395–4405.
- Chen, I.-H. B., K. S. Sciabica, and R. M. Sandri-Goldin. 2002. ICP27 interacts with the RNA export factor Aly/REF to direct herpes simplex virus type 1 intronless mRNAs to the TAP export pathway. *J. Virol.* **76**:12877–12889.
- Collis, P., M. Antoniou, and F. Grosveld. 1990. Definition of the minimal requirements within the human  $\beta$ -globin gene and the dominant control region for high level expression. *EMBO J.* **9**:233–240.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
- Donello, J. E., A. A. Beeche, G. J. Smith III, G. R. Lucero, and T. J. Hope. 1996. The hepatitis B virus posttranscriptional regulatory element is composed of two subelements. *J. Virol.* **70**:4345–4351.
- Donello, J. E., J. E. Loeb, and T. J. Hope. 1998. Woodchuck hepatitis virus contains a tripartite posttranscriptional regulatory element. *J. Virol.* **72**:5085–5092.
- Dreyfuss, G., M. J. Matunis, S. Piñol-Roma, and C. G. Burd. 1993. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* **62**:289–321.
- Ernst, R. K., M. Bray, D. Rekosh, and M.-L. Hammarskjöld. 1997. A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Mol. Cell. Biol.* **17**:135–144.
- Good, P. J., R. C. Welch, W.-S. Ryu, and J. E. Mertz. 1988. The late spliced 19S and 16S RNAs of simian virus 40 can be synthesized from a common pool of transcripts. *J. Virol.* **62**:563–571.
- Grüter, P., C. Taberner, C. von Kobbe, C. Schmitt, C. Saavedra, A. Bachi, M. Wilm, B. K. Felber, and E. Izaurralde. 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell* **1**:649–659.
- Guang, S. 2004. Effects of cis elements and trans factors on intron-independent gene expression. Ph.D. thesis. University of Wisconsin, Madison.
- Guang, S., and J. E. Mertz. 2005. Pre-mRNA processing enhancer (PPE) elements from intronless genes play additional roles in mRNA biogenesis than do ones from intron-containing genes. *Nucleic Acids Res.* **33**:2215–2226.
- Hahm, B., Y. K. Kim, J. H. Kim, T. Y. Kim, and S. K. Jang. 1998. Heterogeneous nuclear ribonucleoprotein L interacts with the 3' border of the internal ribosomal entry site of hepatitis C virus. *J. Virol.* **72**:8782–8788.
- Hamilton, B. J., R. C. Nichols, H. Tsukamoto, R. J. Boado, W. M. Pardridge, and W. F. Rigby. 1999. hnRNP A2 and hnRNP L bind the 3' UTR of glucose transporter 1 mRNA and exist as a complex *in vivo*. *Biochem. Biophys. Res. Commun.* **261**:646–651.
- Hector, R. E., K. R. Nykamp, S. Dheur, J. T. Anderson, P. J. Non, C. R. Urbinati, S. M. Wilson, L. Minvielle-Sebastia, and M. S. Swanson. 2002. Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J.* **21**:1800–1810.
- Huang, Y., K. M. Wimler, and G. G. Carmichael. 1999. Intronless mRNA transport elements may affect multiple steps of pre-mRNA processing. *EMBO J.* **18**:1642–1652.
- Huang, Y., and J. A. Steitz. 2001. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell* **7**:899–905.
- Huang, Y., R. Gattoni, J. Stévenin, and J. A. Steitz. 2003. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell* **11**:837–843.
- Huang, Y., T. A. Yario, and J. A. Steitz. 2004. A molecular link between SR protein dephosphorylation and mRNA export. *Proc. Natl. Acad. Sci. USA* **101**:9666–9670.
- Hui, J., G. Reither, and A. Bindereif. 2003. Novel functional role of CA repeats and hnRNP L in RNA stability. *RNA* **9**:931–936.
- Hui, J., K. Stangl, W. S. Lane, and A. Bindereif. 2003. hnRNP L stimulates splicing of the eNOS gene by binding to variable-length CA repeats. *Nat. Struct. Biol.* **10**:33–37.
- Izaurralde, E. 2002. A novel family of nuclear transport receptors mediates the export of messenger RNA to the cytoplasm. *Eur. J. Cell Biol.* **81**:577–584.
- Kim, J. H., B. Hahm, Y. K. Kim, M. Choi, and S. K. Jang. 2000. Protein-protein interaction among hnRNPs shuttling between nucleus and cytoplasm. *J. Mol. Biol.* **298**:395–405.
- Krecic, A. M., and M. S. Swanson. 1999. hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* **11**:363–371.
- Le Hir, H., E. Izaurralde, L. E. Maquat, and M. J. Moore. 2000. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**:6860–6869.
- Le Hir, H., D. Gatfield, E. Izaurralde, and M. J. Moore. 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**:4987–4997.
- Le Hir, H., A. Nott, and M. J. Moore. 2003. How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* **28**:215–220.
- Lei, E. P., and P. A. Silver. 2002. Protein and RNA export from the nucleus. *Dev. Cell* **2**:261–272.
- Liu, X. 1994. Effects of intron and exon sequence elements on intron-dependent and intron-independent gene expression. Ph.D. thesis. University of Wisconsin, Madison.
- Liu, X., and J. E. Mertz. 1993. Polyadenylation site selection cannot occur *in vivo* after excision of the 3'-terminal intron. *Nucleic Acids Res.* **21**:5256–5263.
- Liu, X., and J. E. Mertz. 1995. hnRNP L binds a cis-acting RNA sequence element that enables intron-dependent gene expression. *Genes Dev.* **9**:1766–1780.
- Liu, X., and J. E. Mertz. 1996. Sequence of the polypyrimidine tract of the 3'-terminal 3' splicing signal can affect intron-dependent pre-mRNA processing *in vivo*. *Nucleic Acids Res.* **24**:1765–1773.
- Lu, S., and B. R. Cullen. 2003. Analysis of the stimulatory effect of splicing on mRNA production and utilization in mammalian cells. *RNA* **9**:618–630.
- Luo, M.-J., and R. Reed. 1999. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Natl. Acad. Sci. USA* **96**:14937–14942.
- Lykke-Andersen, J., M.-D. Shu, and J. A. Steitz. 2001. Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science* **293**:1836–1839.
- Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**:911–940.
- McCracken, S., M. Lambermon, and B. J. Blencowe. 2002. SRm160 splicing coactivator promotes transcript 3'-end cleavage. *Mol. Cell. Biol.* **22**:148–160.
- Niwa, M., S. D. Rose, and S. M. Berget. 1990. *In vitro* polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev.* **4**:1552–1559.
- Niwa, M., and S. M. Berget. 2001. Polyadenylation precedes splicing *in vitro*. *Gene Expr.* **1**:5–14.
- Otero, G. C., and T. J. Hope. 1998. Splicing-independent expression of the herpes simplex virus type 1 thymidine kinase gene is mediated by three cis-acting RNA subelements. *J. Virol.* **72**:9889–9896.
- Pasquinielli, A. E., R. K. Ernst, E. Lund, C. Grimm, M. L. Zapp, D. Rekosh, M.-L. Hammarskjöld, and J. E. Dahlberg. 1997. The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO J.* **16**:7500–7510.
- Piñol-Roma, S., M. S. Swanson, J. G. Gall, and G. Dreyfuss. 1989. A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *J. Cell Biol.* **109**:2575–2587.
- Piñol-Roma, S., and G. Dreyfuss. 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**:730–732.
- Reed, R., and E. Hurt. 2002. A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108**:523–531.
- Ryu, W.-S. 1989. Processing of transcripts made from intron-containing and intronless protein-coding genes. Ph.D. thesis. University of Wisconsin, Madison.
- Ryu, W.-S., and J. E. Mertz. 1989. Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. *J. Virol.* **63**:4386–4394.
- Shih, S.-C., and K. P. Claffey. 1999. Regulation of human vascular endothelial growth factor mRNA stability in hypoxia by heterogeneous nuclear ribonucleoprotein L. *J. Biol. Chem.* **274**:1359–1365.
- Stutz, F., A. Bachi, T. Doerks, I. C. Braun, B. Séraphin, M. Wilm, P. Bork, and E. Izaurralde. 2000. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* **6**:638–650.
- Yu, X.-M., G. W. Gelembiuk, C.-Y. Wang, W.-S. Ryu, and J. E. Mertz. 1991. Expression from herpesvirus promoters does not relieve the intron requirement for cytoplasmic accumulation of human  $\beta$ -globin mRNA. *Nucleic Acids Res.* **19**:7231–7234.
- Zhou, Z., M.-J. Luo, K. Straesser, J. Katahira, E. Hurt, and R. Reed. 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* **407**:401–405.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**:3406–3415.