RNA (2001), *7*:1034–1042. Cambridge University Press. Printed in the USA.

Position and sequence requirements for poly(A) length regulation by the poly(A) limiting element

JAYDIP DAS GUPTA,[1,3]* HAIDONG GU,[1,2]* and DANIEL R. SCHOENBERG[1]

[1]Department of Molecular and Cellular Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA
[2]Ohio State Biochemistry Program and The Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210, USA

ABSTRACT

The poly(A)-limiting element (PLE) is a *cis*-acting sequence that acts to limit poly(A) tail length on pre-mRNA to <20 nt. Functional PLEs are present in a number of genes, underscoring the generality of this control mechanism. The current study sought to define further the position requirements for poly(A) length regulation and the core sequence that comprises a PLE. Increasing the spacing between the PLE and the upstream 3′ splice site or between the PLE and the downstream AAUAAA had no effect on poly(A) length control. However, moving the PLE from the terminal exon to either an upstream exon or intron eliminated poly(A) length control. Poly(A) length control was further evaluated using a battery of constructs in which the PLE was maintained in the terminal exon, but where upstream introns were either deleted, modified, or replaced with a polypyrimidine tract. Poly(A) length control was retained in all cases, indicating that the key feature is the presence of the PLE in the terminal exon. A battery of mutations demonstrated the importance of the 5′ pyrimidine-rich portion of the element. Finally, UV crosslinking experiments identified an ~62-kDa protein in Hela nuclear extract that binds to a wild-type 23-nt PLE RNA oligonucleotides but not to a mutated nonfunctional form of the element.

Keywords: 3′ processing; polyadenylation; poly(A)-limiting element; pre-mRNA processing

INTRODUCTION

With the exception of histone mRNA, most vertebrate mRNAs possess 3′ poly(A) tails ranging from 100 to 200+ residues in length. Poly(A) addition is one of the last steps in pre-mRNA processing, occurring prior to splicing of the terminal intron and nuclear export (Bauren & Wieslander, 1994). Based on the exon definition model, the 3′ processing complex substitutes for a 5′ splice site to define the end of the terminal exon (Berget, 1995), and there is good evidence linking 3′ processing to splicing of the terminal intron (Lou et al., 1996; Cooke et al., 1999; Custodio et al., 1999). This interaction between splicing and 3′ processing plays a key role in regulating gene expression, particularly in cases involving exon selection linked to selection of alternative 3′ ends (Lou et al., 1999).

Reprint requests to: Daniel R. Schoenberg, Ph.D., Department of Molecular & Cellular Biochemistry, The Ohio State University, 333 Hamilton Hall, 1645 Neil Avenue, Columbus, Ohio 43210-1218, USA; e-mail: schoenberg.3@osu.edu.

*These authors contributed equally to this work.

[3]Current address: Department of Immunology, Lerner Research Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, Ohio 44195-0001.

Pre-mRNA 3′ processing is regulated by *cis*-acting elements in the message body (reviewed in Zhao et al., 1999). Upstream elements, or USEs, in HIV-1 proviral RNA (Gilmartin et al., 1995), adenovirus late RNA (Sittler et al., 1995), SV40 late RNA (Schek et al., 1992), and C2 complement gene (Moreira et al., 1998) can act as enhancers to stimulate 3′ processing of their respective RNAs. Conversely, U1A binding to a pair of U1A binding sites adjacent to AAUAAA in U1A pre-mRNA blocks poly(A) addition by inhibiting poly(A) polymerase without affecting the initial cleavage step (Gunderson et al., 1994). Elements upstream of AAUAAA can also regulate poly(A) tail length. Poly(A) length regulation has been best characterized for mRNAs that undergo deadenylation or poly(A) addition during various stages of egg maturation and embryonic development (reviewed in Richter, 1999). This form of regulation is cytoplasmic, and involves the interaction of the cytoplasmic polyadenylation element binding protein (CPE-BP) with a U-rich cytoplasmic polyadenylation element (CPE) upstream of AAUAAA (Wu et al., 1997). Embryo-specific deadenylation is controlled by a distinct upstream element, EDEN, through its interaction with EDEN-BP (Paillard et al., 1998). The deadenylation ef-

ficiency of the EDEN element is enhanced by the presence of upstream AUU repeats, suggesting an interplay between multiple *cis*-acting elements in regulating deadenylation (Audic et al., 1998).

In the course of studying the estrogen regulation of mRNA stability in *Xenopus*, we found that albumin and several other serum protein mRNAs have an unusually short, 17-nt poly(A) tail (Schoenberg et al., 1989; Pastori et al., 1991). Using an intron-specific primer in an RT-PCR assay for poly(A) tail length (Salles et al., 1999), we subsequently identified the same short poly(A) tail on albumin pre-mRNA (Rao et al., 1996), suggesting length control is intrinsic to 3′ processing rather than the result of nuclear deadenylation. The regulation of poly(A) tail length depends on the presence of a *cis*-acting poly(A)-limiting element, or PLE (Das Gupta et al., 1998). Two such elements, termed PLE A and B, are located in the terminal albumin gene exon, and elements related to the albumin gene PLE B have been identified in the terminal exons for the genes encoding transferrin and the transcription factor HIV-EP2/Schnurri 2 (Gu et al., 1999). A database search identified several hundred genes with PLE-like elements, and the mRNAs of those genes that we have examined to date have $<$20 nt poly(A) tails. These results indicate that nuclear regulation of poly(A) length is a general phenomenon occurring on numerous mRNAs. Given the importance of the poly(A) tail in nucleo-cytoplasmic export, mRNA turnover, and translation initiation (Baker, 1997), PLE regulation of poly(A) length can have a significant impact on gene expression.

The present study sought to determine the spatial requirements for a PLE to act in regulating poly(A) tail length, and to identify the nucleotides within the PLE sequence required for its function. Our results indicate that the PLE must be in the terminal exon to regulate poly(A) tail length, and poly(A) length control is independent of the distance between the PLE and AAUAAA or an upstream intron. A series of targeted mutations determined the importance of the pyrimidine-rich 5′ portion of the PLE for poly(A) length control, and using wild-type and mutant RNA oligonucleotides in UV-crosslinking experiments we identified an ~62-kDa protein that binds specifically to the PLE.

RESULTS

Poly(A) length regulation is independent of the spacing between the PLE and adjacent RNA processing signals

The PLEs present in the albumin and transferrin genes are relatively close to both the upstream 3′ splice site and downstream AAUAAA, and the constructs examined in our previous studies (Das Gupta et al., 1998; Gu et al., 1999) had similarly close spacing between the PLE and adjacent pre-mRNA processing signals. In contrast, the PLE identified in the terminal exon of the HIV-EP2/Schnurri 2 gene is 1,649 nt upstream of AAUAAA. To determine whether the spacing between the PLE and the upstream 3′ splice site or the downstream AAUAAA influences poly(A) length regulation, fragments of ΦX174 DNA ranging in size from 200 to 500 bp were inserted either upstream or downstream of this element as diagramed in Figure 1. In this and all subsequent figures the control labeled glo-SPA (lane 9) corresponds to the CMV

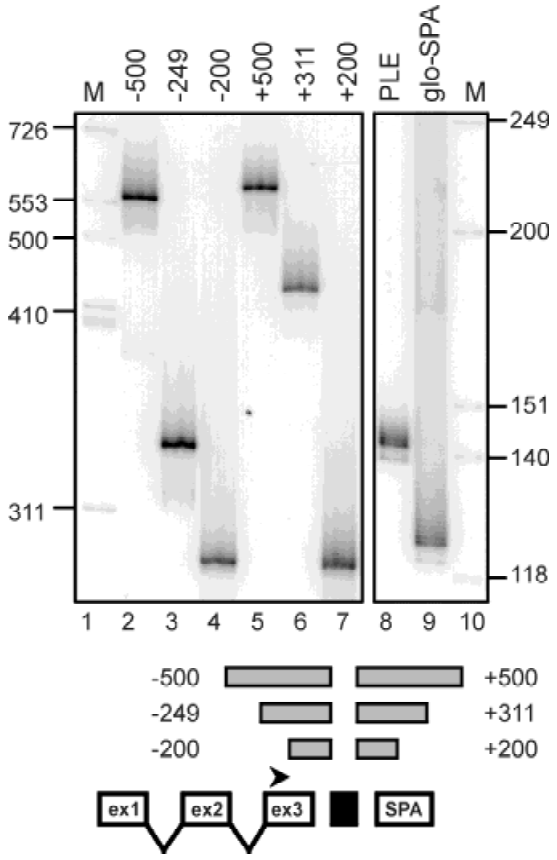

FIGURE 1. Spacing effects on poly(A) length regulation. *Hinf*I fragments of ΦX174 DNA ranging in size from 200 to 500 bp were inserted between the stop codon at the end of exon 3 and the PLE, or between the PLE and SPA as shown on the bottom of the figure. LM(tk-) cells were transfected with these plasmids and poly(A) tail length was determined by RT-PCR using a 5′ [^{32}P]labeled primer from globin exon 3 as described previously (Das Gupta et al., 1998) on nuclear RNA isolated 24 h after transfection. The different sizes of the PCR products result from the different lengths of DNA inserted between the PLE and either the upstream 3′ splice site or SPA. In the nomenclature used, −500 (lane 2), −249 (lane 3), and −200 (lane 4) correspond to inserts that increased the distance between the PLE and stop codon at the end of exon 3. Similarly, +500 (lane 5), +311 (lane 6), and +200 (lane 7) correspond to inserts that increased the distance between the PLE and SPA. The more slowly migrating PCR products from constructs with inserted DNA are grouped in the panel on the left, and the faster migrating controls of vector with (PLE, lane 8) and without a PLE (glo-SPA, lane 9) are shown in the adjacent panel. The same 5′-end-labeled, *Hinf*I-digested ΦX174 DNA size marker was electrophoresed in lanes 1 and 10. The location of the [^{32}P]labeled primer used for RT-PCR poly(A) analysis is shown with an arrowhead in the diagram on the bottom of the figure.

promoter-driven globin mRNA control bearing the synthetic polyadenylation element but no PLE. Poly(A) tails of <20 nt were obtained on all PLE-containing nuclear RNAs regardless of the linear spacing between the element and the upstream 3′ splice site (Fig. 1, lanes 2–4), or downstream AAUAAA (Fig. 1, lanes 5–7), indicating that such spacing plays no detectable role in poly(A) length control.

The PLE must be in the terminal exon to regulate poly(A) tail length

The functional PLEs identified to date all reside in the terminal exons of their respective genes. To determine whether this is significant for poly(A) length control, we prepared the constructs shown in Figure 2 where an albumin PLE was moved upstream into exon 2 (in a manner that maintained the reading frame) or into intron 2. The length of the poly(A) tail on the reporter globin mRNA was then determined on RNA isolated from nuclei of transfected cells. The results in Figure 2, lanes 4 (exon 2) and 5 (intron 2) indicate that moving the PLE to either of these upstream sites abrogated poly(A) length regulation, resulting in globin mRNAs whose poly(A) tails were indistinguishable from the 100–200-nt poly(A) present on the control (−PLE) mRNA (Fig. 2, lane 3, glo-SPA). The PCR product from the CMV-glo-PLE A-SPA control (denoted CMV-glo-(−124/−104)-SPA in Das Gupta et al., 1998) (Fig. 2, lane 2) is labeled as "exon 3" in this figure to be consistent with the nomenclature used for the other two constructs. Thus, to function in regulating poly(A) tail length, the PLE must be located in the terminal exon.

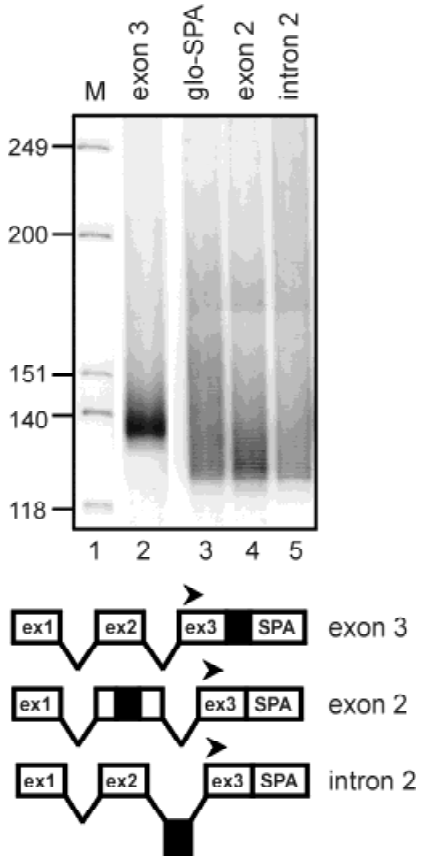

FIGURE 2. Position dependence for PLE regulation of poly(A) length. LM(tk-) cells were transfected with plasmids diagramed on the bottom of the figure in which the PLE was located in the terminal exon (between the stop codon at the end of exon 3 and SPA, lane 2, exon 3), moved in frame into exon 2 (lane 4, exon 2), or moved into intron 2 (lane 5, intron 2). The construct labeled exon 3 is the PLE A construct CMV-glo-(−124/−104)-SPA described previously (Das Gupta et al., 1998). Lane 3 is a control of vector lacking an inserted PLE (lane 3, glo-SPA). Lane 1 (M) contains a size marker of 5′-end-labeled, *Hinf* I-digested ΦX174 DNA. The location of the [^{32}P]labeled primer used for RT-PCR poly(A) analysis is shown with a filled arrow in the diagram on the bottom of the figure.

Relationship between the presence of upstream introns and poly(A) length control

Antoniou and coworkers (Custodio et al., 1999) showed that 3′ processing of native β-globin pre-mRNA is relatively unaffected by mutations that inactivate the splicing of introns 1 or 2, but is dependent on the presence of the polypyrimidine sequence in the distal portion of globin intron II (Antoniou et al., 1998). This distal polypyrimidine tract alone can effectively replace intron II in stimulating 3′ processing in stably transfected MEL cells when placed between exons 2 and 3 in the βPY construct. The relationship between proximal intron splicing and PLE regulation of poly(A) length was examined using a series of previously described constructs (Antoniou et al., 1998; Custodio et al., 1999), which we modified to be consistent with those used in our prior studies (Das Gupta et al., 1998; Gu et al., 1999). These constructs, which are shown in Figure 3, all are expressed under control of the CMV promoter and contain both PLE B and the SPA synthetic polyadenylation element. βIVS-I and βIVS-II consist of globin genes bearing only intron I or intron II, respectfully. Thus, βIVS-I lacks the sequences shown by Custodio et al. (1999) to be involved in coupling 3′ processing to splicing of intron II. βΔRS contains an intact intron I, but the central portion of intron II has been deleted, retaining 87 bp of the 5′ portion of the intron and 170 bp of the 3′ portion of the intron including the terminal polypyrimidine tract. cβΔRS is identical to βΔRS except that it lacks intron I. In βPY intron II is replaced with an in-frame 21-nt polypyrimidine tract insertion.

The effects of these changes in globin gene structure on poly(A) length regulation are shown in Figure 4. In no cases did we observe RNA that failed to undergo 3′ processing, indicating that the strong SPA element in our constructs obviates the requirement of intron II sequences for 3′ processing. We have done this analysis by both standard RT-PCR for poly(A) length and by

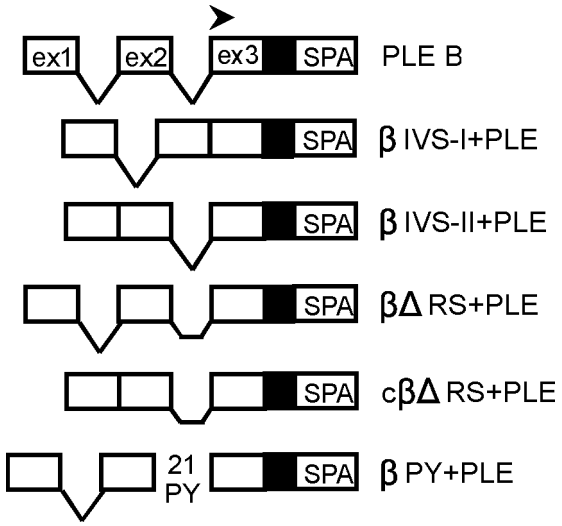

FIGURE 3. Organization of β-globin constructs evaluated for their effect on poly(A) length control. A series of β-globin gene constructs described by papers from Antoniou and co-workers (Antoniou et al., 1998; Custodio et al., 1999) were modified to fit the CMV-globin-SPA expression vector bearing PLE B as indicated. In βIVS-I and II, introns II and I, respectively, have been deleted, with the corresponding exons joined at the correct junctions. βPY differs from βIVS-I in having a 21-bp polypyrimidine tract inserted at the junction of exons 2 and 3. In $\beta\Delta$RS and c$\beta\Delta$RS, the central portion of intron 2 has been deleted, leaving 87 bp of the 5′ portion of the intron and 170 bp of the 3′ portion. The latter construct also lacks intron 1. The effect of these changes on pre-mRNA splicing was confirmed by RT-PCR of nuclear RNA (data not shown).

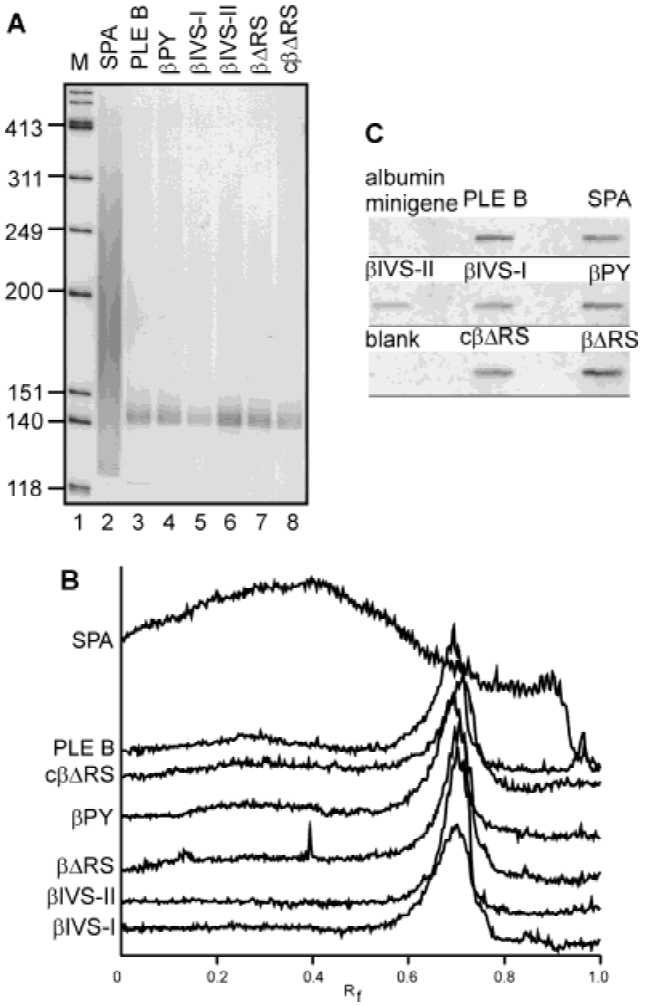

FIGURE 4. Relationship between individual introns and poly(A) length regulation. **A:** Poly(A) tail length on nuclear RNA isolated from transfected LM(tk-) cells was analyzed by PAT assay as described in Materials and Methods. Lane 1 (M) contains a size marker of 5′-end-labeled, *Hinf* I-digested ΦX174 DNA, and RNAs expressed from the different globin plasmids are indicated on the top of the gel. **B:** The gel in **A** was scanned and the individual scans for each RNA are shown. **C:** The RNAs analyzed in **A** were applied to a nylon membrane using a slot-blot apparatus and probed with [^{32}P]labeled globin cDNA. Controls included RNA from cells transfected with a plasmid expressing albumin mRNA (albumin minigene, top left) and no added RNA (bottom left).

PAT assay with similar results. The autoradiogram of the PAT assay for poly(A) length of these mRNAs is shown in Figure 4A and densitometric scans of this autoradiogram are shown in Figure 4B. Figure 4C is a slot blot showing the relative amount of each globin mRNA reporter analyzed in Figure 4A,B.

The constructs possessing a full-length or truncated intron II (PLE B, Fig. 4A, lane 3, βIVS-II, lane 6, $\beta\Delta$RS, lane 7, c$\beta\Delta$RS, lane 8) all produced pre-mRNA with $<$20 nt poly(A) tails regardless of the presence or absence of intron 1. When globin intron 2 was deleted in the βIVS-I (Fig. 4A, lane 5) the resulting pre-mRNA had the same $<$20-nt poly(A) tail as seen on pre-mRNAs bearing intron 2. This finding agrees with the results obtained in Figure 1, and indicates that the distance between the PLE and an upstream intron does not affect poly(A) length control. In addition these data show that intron-specific sequences are not involved in linking splicing to regulation of poly(A) length, as the constructs deleted for most of intron II ($\beta\Delta$RS, c$\beta\Delta$RS) or lacking intron 2 altogether (βIVS-I) all behaved equally. Replacing intron 2 with a polypyrimidine tract in the βPY construct (Fig. 4A, lane 4) also had no effect on poly(A) length control. Thus it appears that in contrast to the link between intron II splicing and 3′ processing observed by Custodio et al. (1999), the PLE acts independently of sequences in the upstream introns to regulate poly(A) tail length.

Mutational mapping of functional elements within the PLE

The PLEs examined to date appear palindromic, with a pyrimidine-rich 5′ half and a purine-rich 3′ half. To define residues within the PLE that are critical for poly(A) length regulation, we synthesized DNA primers containing the sequences indicated on the bottom of Figure 5 and inserted the resulting mutated elements

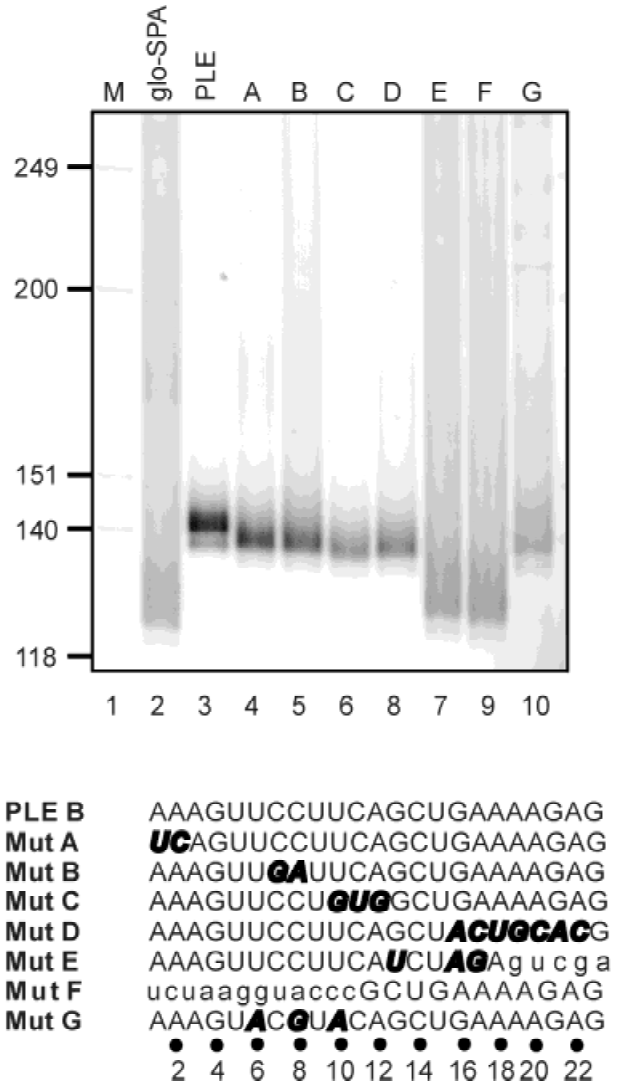

FIGURE 5. Mutational inactivation of PLE B. Mutations in PLE B construct are diagramed on the bottom of the figure with individual base changes indicated in bold italics. Vector sequences replacing deleted portions of PLE B in Mut E and F are shown in lower case. The numbering at the bottom and in the text refers to positions within the originally identified PLE B sequence (Das Gupta et al., 1998). Poly(A) tail length on nuclear RNA isolated from transfected LM(tk-) cells was analyzed as in Figure 1, and compared to controls of vector alone (glo-SPA, lane 2) and vector containing wild-type PLE B (PLE, lane 3). Lane 1 (M) contains a size marker of 5′-end-labeled, *Hinf* I-digested ΦX174 DNA.

between the end of globin exon 3 and SPA in the same manner as the original PLE B construct. Poly(A) tail length on the reporter globin mRNA was then determined for RNA isolated from the nuclei of transfected cells.

Mutations A–C and G targeted the pyrimidine-rich 5′ half of PLE B. Mutation A, which changed the 5′-most AA dinucleotide to UC, had little effect on poly(A) length regulation (Fig. 5, lane 4, Mut A), indicating that the functional element within PLE B is smaller than the 23 nt originally identified. The importance of the 5′ half of the element was demonstrated by mutation F (Mut F, Fig. 5, lane 9), where poly(A) length control was lost after deleting this pyrimidine-rich region. The large number of pyrimidines in this region may function redundantly, as converting the CC dinucleotide at positions 7 and 8 to GA (Fig. 5, lane 5, Mut B) appeared to have an intermediate effect, where the majority of mRNAs had a $<$20-nt poly(A) tail, but there was also smearing to greater lengths. However, it is difficult to make unequivocal judgments about leaky regulation of poly(A) tail length using the RT-PCR assay. The requirement for multiple adjacent pyrimidines here was tested further in mutation G, where the 5′ pyrimidine-rich region was disrupted by replacing every other pyrimidine with a purine (UUCCUUC $\rightarrow$ UACGUAC, Fig. 5, lane 10). These changes inactivated the PLE, underscoring the necessity for a stretch of adjacent pyrimidine residues in this portion of the element.

Changing UCA at positions 10–12 to GUG, in a region seemingly between the purine and pyrimidine-rich elements had no effect on poly(A) length regulation (Fig. 5, lane 6, Mut C). Similarly, mutating GAAAAG at positions 16–22 to ACUGCA had little effect on poly(A) length control (Fig. 5, lane 7, Mut D). It should be noted that this mutation retained purines at positions 16, 19, and 21. While Mut E (Fig. 5, lane 8) retained the pyrimidine-rich 5′ half, changing the purine-rich 3′ half inactivated the PLE. Based on the alignment of the identified PLEs with the mutation data (see Discussion and Fig. 7), we think the critical change here was the G $\rightarrow$ U mutation at position 13. In summary, the mutations tested here underscore the importance of the 5′ pyrimidine-rich region, but the exact residues required within the 3′ purine-rich for PLE-function need to be resolved further.

Identification of a nuclear PLE-binding protein

Having defined the sequence requirements for PLE function, we next sought to determine whether a specific PLE-binding protein (PLEBP) could be identified. In the experiment in Figure 6A increasing amounts of Hela nuclear extract were incubated with a 23-nt 5′-[^{32}P]labeled RNA oligonucleotide spanning the PLE B sequence shown in Figure 5 or a similarly labeled RNA oligonucleotide spanning the Mut G sequence. Bound protein was identified by UV crosslinking, followed by RNase digestion, SDS-PAGE, and autoradiography of the dried gel. This identified an $\sim$62-kDa PLEBP that bound to the PLE but not to Mut G. To determine the specificity of binding, we performed the competition binding experiment shown in Figure 6B. In this experiment nuclear extract was prebound on ice for 30 min to the 5′-[^{32}P]PLE B RNA oligonucleotide, followed by the addition of the indicated amounts of unlabeled competitor PLE B or Mut G RNA oligonucleotide and a

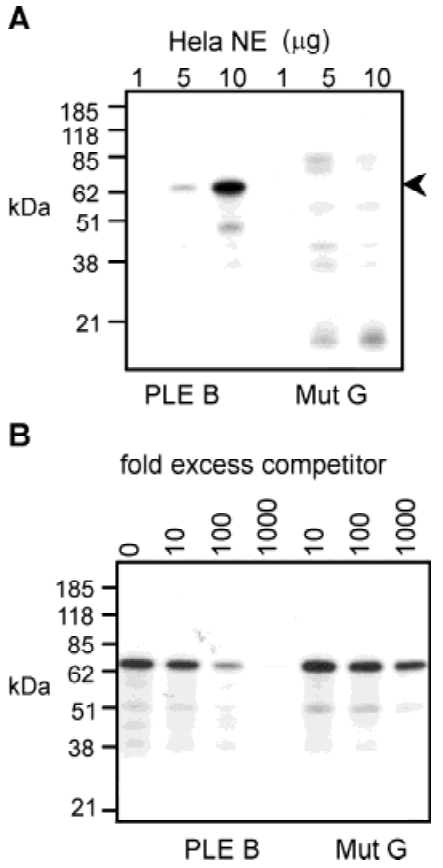

FIGURE 6. Identification of a candidate PLE-binding protein. **A:** Increasing amounts of Hela nuclear extract were incubated with a 5′-[^{32}P]labeled 23-nt RNA oligonucleotide for PLE B (left half) or a similar RNA oligonucleotide bearing the Mut G sequence shown in Figure 5. The mixtures were UV crosslinked, digested with RNase A, and analyzed by SDS-PAGE and autoradiography. The arrowhead indicates the position of the PLE-binding protein. **B:** Ten micrograms of Hela nuclear extract were incubated on ice for 30 min with 5′-[^{32}P]labeled PLE B RNA oligonucleotide. To this was added the indicated fold-excess of either unlabeled PLE B or Mut G RNA oligonucleotides. After an additional 30 min on ice, the mixtures were UV crosslinked and analyzed as in **A**.

subsequent 30 min of incubation to displace bound RNA. The degree of binding was determined as in Figure 6A by UV crosslinking and SDS-PAGE. Whereas the majority of bound [^{32}P]PLE B was competed by a 100-fold excess of its cognate unlabeled oligonucleotide, no competition was observed with even a 1,000-fold excess of Mut G RNA oligonucleotide. These results are consistent with involvement of a specific nuclear RNA-binding protein in the PLE regulation of poly(A) tail length.

DISCUSSION

The current study sought to characterize further the regulation of poly(A) tail length by the PLE. The results in Figuers 1 and 2 show that the PLE must be in the terminal exon to function in regulating poly(A) length, and that the linear distance between the PLE and either the upstream 3′ splice site or the downstream AAUAAA does not influence poly(A) length regulation. These data are in good agreement with our previous observations on the PLE in the HIV-EP2/Schnurri 2 gene, where the 3′ end of the PLE resides 10 nt upstream of the translation stop codon and 1,626 nt upstream of AAUAAA in a large terminal exon (Gu et al., 1999). The absence of a spacing effect is not entirely surprising, because RNA folding could juxtapose the PLE with AAUAAA.

A link between pre-mRNA 3′ processing and splicing of the terminal intron has been reported for several mRNAs (Lou et al., 1996; Cooke et al., 1999; Zhao et al., 1999). Because previous work was performed using intron-containing mRNAs (Das Gupta et al., 1998) we wished to determine whether there is an impact of upstream introns on the regulation of poly(A) length. This was done using a panel of β-globin gene splice-site mutations (shown in Fig. 3 and kindly provided by Dr. Michael Antoniou) whose effects on 3′ processing have been previously characterized (Antoniou et al., 1998; Custodio et al., 1999). In the experiments performed here, these mutant genes were substituted for the β-globin reporter gene in a manner that retained the PLE in the terminal exon and the strong SPA element controlling 3′ processing. Deleting either intron II in the βIVS-I construct or intron I in the βIVS-II construct had no effect on poly(A) length regulation, a result that further supports the finding that there is no effect of distance between an upstream intron and the PLE on regulated polyadenylation. Previous work by Antoniou et al. (1998) showed that sequences within the distal portion of intron II are required for efficient 3′ processing of β-globin pre-RNA expressed in its native context from the β-globin promoter. The efficient 3′ processing observed here in the absence of intron II likely results from our use of the optimized SPA element (Levitt et al., 1989) in place of the native globin 3′ processing signals. Because equivalent results were obtained with βIVS-I and βIVS-II, it is unlikely that sequences unique to either intron are involved in poly(A) length regulation. This is supported by results obtained with the constructs c$\beta\Delta$RS and $\beta\Delta$RS, in which all but 257 bp comprising the very 5′ and 3′ portions of intron II were deleted. Attempts to inactivate splicing of intron II consistently yielded alternatively spliced products with predominantly short poly(A) (data not shown), making it difficult to use that approach to determine whether there is a direct relationship between intron splicing and poly(A) length regulation. However, the PLE regulates poly(A) tail length on both intronless and intron-containing luciferase mRNA (V. Betapudi & D.R. Schoenberg, in prep.), indicating that neither the presence or splicing of an upstream intron is required for poly(A) length control.

Our results are consistent with a recurring theme of elements upstream of the poly(A) site impacting 3′ processing by serving as binding sites for proteins that interact with poly(A) polymerase (Gunderson et al., 1997, 1998; Vagner et al., 2000). Because the PLE must be in the terminal exon to function in regulating poly(A) length, a PLE-binding protein may regulate the activity of poly(A) polymerase either through direct binding to poly(A) polymerase or by binding to PAB II in a manner analogous to influenza virus NS1 protein (Chen et al., 1999). Vagner et al. (2000) demonstrated that U2AF65 interacts directly with poly(A) polymerase to couple 3′ processing to splicing. Thus when the PLE is in exon 2 or intron II, U2AF65 binding to PAP or the assembly of a spliceosome complex on the downstream intron may block access of a bound PLEBP to PAP or PAB II, resulting in the default addition of a 100–200-nt poly(A) tail. The characterization of this mechanism must await the identification of the PLEBP detected in Figure 6 and development of an in vitro assay for regulated polyadenylation. Efforts are underway to purify this protein, and to date we have ruled out PTB, hnRNP H, hnRNP F, and PAB II as candidate PLEBPs.

The PLEs characterized to date have two or three As at the 5′ end, a stretch of up to eight pyrimidines (interrupted by a single A in some cases), followed by a stretch of up to nine purines with a predominance of A residues. The experiment in Figure 5 sought to determine functional characteristics of this repeating structure. An alignment of albumin gene PLEs, the PLEs from the transferrin and HIV-EP2/Schnurri 2 genes, and a core PLE derived from our mutation data is shown in Figure 7. The most conclusive findings came from mutations within the 5′ pyrimidine-rich portion of the element (Mut B, F, G). Either deleting this region (Fig. 5, Mut F), or changing every other pyrimidine to a purine (Fig. 5, Mut G) completely inactivated the PLE, whereas changing two central pyrimidines to purines had a more modest effect (Fig. 5, Mut B). Mutations in the 3′ purine-rich portion of the element proved less informative. Although Mut E changed several of these residues, the alignment in Figure 7 suggests that the G → U mutation at position 13 was the change that was most likely responsible for inactivating the PLE (see Fig. 5, Mut E and lane 8). Because disrupting the contiguous stretch of purines in the 3′ half of the PLE by mutating the sequence GAAAAGA to ACUGCA had no effect on poly(A) length regulation (Fig. 5, Mut D, lane 7), it is possible that the 3′ purine-rich portion of the PLE is dispensable, or that the roles of individual bases are more subtle than could be determined by this approach.

PLE A	AACUCACUGAGGAACA
PLE B	AGUUCCUUCAGCUGAAAAGAG
transferrin	AGUUUAUUCUCAGAUGUGGGAGG
HIV-EP2	AGAUCCUUCAUCAGAAAAGAG
core PLE	AGUUCCUUYRGCURNRNRRR

FIGURE 7. A consensus PLE. The sequences of two albumin gene PLEs are aligned with the PLEs from the transferrin gene, the HIV-EP2/Schnurri 2 gene, and a core PLE compiled from the mutation data in Figure 5 and the sequences of the above PLEs.

MATERIALS AND METHODS

Plasmid constructions

The basic expression vector CMV-glo-SPA was described previously (Das Gupta et al., 1998). To insert a PLE into exon 2 of CMV-glo-SPA (Fig. 2), a *Bst*EII site was introduced at position 1365, and the resulting plasmid was digested with *Bst*EII and end-filled with Klenow fragment of DNA polymerase. PLE A was prepared by annealing primers XG-3(5′-CAAACTCACTGAGGAACACC) and EC73 (5′-CAGATCTA GAAGGTGTTCCTCAGTG), followed by end-filling with Klenow fragment of DNA polymerase. This was ligated to the above cut vector and both the orientation of the insert and the maintenance of correct reading frame were confirmed by DNA sequencing. PLE A was introduced directly into intron II (position 2289 of CMV-glo-SPA) using the GeneEditor system (Promega) and primer JOY29 (5′-GGCTGGATTATTCTG AGTCCAAACTCACTGAGGAACACCTTCGCTAGGCCCT TTTGCTAATC). The plasmids βIVS-Iμ, βIVS-IIμ, βPYμ, $\beta\Delta$RSμ, and c$\beta\Delta$RSμ containing mutated globin genes with the locus control region were a gift from Dr. Michael Antoniou (GST School of Medicine, London). The portion of the β-globin gene in CMV-glo-PLE B-SPA spanning the *Nco*I site in exon 1 to the *Bst*XI site in exon 3 was replaced with the corresponding portion of these mutated globin genes to generate a series of plasmids with alterations in the splicing of either intron I or intron II. For experiments to examine the effect of changing the distance between the PLE and the upstream 3′ splice site or downstream SPA element (Fig. 1), the plasmid CMV-glo-PLE A-SPA was digested with either *Asp*718 or *Xba*I, followed by incubation with Klenow fragment of DNA polymerase to fill in unpaired ends. The spacer DNA consisted of *Hinf*I fragments of ΦX174 DNA that were end-filled using Klenow fragment of DNA polymerase and separated on an agarose gel prior to ligation to the restricted vector. Mutations within PLE B (Fig. 5) were prepared by annealing primers bearing the indicated sequences, end-filling with Klenow fragment of DNA polymerase, and digestion with *Xba*I. These were cloned into CMV-glo-SPA that was digested with *Sma*I and *Xba*I. The following primer pairs were used: mutation A, JOY42 (5′-TCAGTTCCTTCAGCTGAAAAGAG) and JOY43 (5′-GGTCTAGACTCTTTTCAGCTGAA); mutation B, JOY44 (5′-AAAGTTGATTCAGCTGAAAAGAG) and JOY43; mutation C, JOY45 (5′-AAAGTTCCTGTGGCTGAAAAGAG) and JOY46 (5′-GGTCTAGACTCTTTTCAGCCACA); mutation D, JOY47(5′-AAAGTTCCTTCAGCTACTGCACG) and JOY48-(5′-GGTCTAGACGTGCAGTAGCTGAA); mutation E, JOY49 (5′-AAAGTTCCTTCA) and JOY50 (5′-GGTCTAGATGAA GGAACTTT); mutation F, JOY51 (5′-GCTGAAAAGAG) and JOY52 (5′-GGTCTAGACTCTTTTCAGC); mutation G, JOY 53 (5′-AAAGTACGTACAGCTGAAAAGAG) and JOY54 (5′-GTCTAGACTCTTTTCAGCTGTA).

UV crosslinking to identify PLEBP

Hela S3 cells were obtained from the Cell Culture Center. Nuclear extract was prepared as described by Rüegsegger et al. (1996). The RNA oligonucleotides PLE B (5′-AAAG UUCCUUCAGCUGAAAAGAG) and Mut G (5′-AAAGUAC GUACAGCUGAAAAGAG) were obtained from Dharmacon. One hundred femtomoles of each oligonucleotide was labeled at the 5′ end using [^{32}P]ATP and T4 kinase. These were mixed with 1, 5, or 10 μg of HeLa nuclear extract in 2 mM Tris-HCl, pH 7.6, 0.2 mM magnesium acetate, 0.2 mM dithiothreitol, 14 mM KCl, 2% glycerol, 0.2 mM EDTA, 8 μM EGTA and 500 ng/μL Heparin in a total volume of 10 μL. The RNA–protein mixture was incubated on ice for 30 min and then placed 3 cm from a 254 nm light source and irradiated for 2 min in a Stratalinker (Stratagene). Ten micrograms RNase A were then added to the irradiated mixture to digest the RNA. Crosslinked protein was separated on a 10% SDS-PAGE and identified by autoradiography of the dried gel. To determine the effectiveness of Mut G for binding to PLEBP62, 100 fmol of 5′-[^{32}P]labeled PLE B RNA oligo were mixed with 10 μg of HeLa nuclear extract under the same binding conditions as described above and incubated on ice for 30 min. Unlabeled competitor (PLE B or Mut G) was added to the binding reaction after the complex formation and incubated on ice for another 30 min. The RNA–protein–competitor mixture was then UV irradiated, RNase A treated, and visualized as described above.

Manipulation of cells

Lm(tk-) cell were obtained from American Type Culture Collection. They were maintained in Dulbecco's Minimal Essential Medium plus 5% fetal calf serum. Cells were transfected with Superfect reagent (Qiagen), using the manufacturer's protocol. Nuclei were isolated from cells harvested 24 h after transfection, and RNA was extracted as described by Sambrook et al. (1989). The extracted RNA was treated with RNase-free DNase prior to its use for cDNA synthesis.

RT/PCR assay for poly(A) tail length

Reverse transcription and PCR assays were done using oligo(dT) anchor XG2 (5′-GGGGATCCGCGGTTTTTTTTTT) and 5′-[^{32}P]labeled primer XG1 (5′-GGCAACGTGCTGGT CTGTGT) as described previously (Das Gupta et al., 1998), or by the PAT assay (Salles et al., 1999) using the same primers but with the added step of ligating phosphorylated oligo(dT) annealed to the RNA prior to addition of XG2. The results of the PAT assay data in Figure 4 were analyzed by scanning densitometry using a Protein Databases 420oe scanning densitometer and Quantity One software.

ACKNOWLEDGMENTS

We wish to thank Greg Gilmartin for his helpful comments on this article and Michael Antoniou for his gift of globin expression vectors. This work was supported by National Institutes of Health Grant GM55407 from the National Institute of General Medical Sciences to D.R.S. and by Center Grant P30 CA16058 from the National Cancer Institute to The Ohio State University Comprehensive Cancer Center.

Received March 5, 2001; accepted without revision April 5, 2001

REFERENCES

Antoniou M, Geraghty F, Hurst J, Grosveld F. 1998. Efficient 3′-end formation of human β-globin mRNA in vivo requires sequences within the last intron but occurs independently of the splicing reaction. *Nucleic Acids Res 26*:721–729.

Audic Y, Omilli F, Osborne HB. 1998. Embryo deadenylation element-dependent deadenylation is enhanced by a *cis* element containing AUU repeats. *Mol Cell Biol 18*:6879–6884.

Baker EJ. 1997. mRNA polyadenylation: Functional implications. In: Harford J, Morris DR, eds. *mRNA metabolism and post-transcriptional gene regulation*. New York: John Wiley & Sons, Inc. pp 85–104.

Bauren G, Wieslander L. 1994. Splicing of Balbiani ring-1 gene pre-messenger RNA occurs simultaneously with transcription. *Cell 76*:183–192.

Berget SM. 1995. Exon recognition in vertebrate splicing. *J Biol Chem 270*:2411–2414.

Chen Z, Li Y, Krug RM. 1999. Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3′-end processing machinery. *EMBO J 18*:2273–2283.

Cooke C, Hans H, Alwine JC. 1999. Utilization of splicing elements and polyadenylation signal elements in the coupling of polyadenylation and last-intron removal. *Mol Cell Biol 19*:4971–4979.

Custodio N, Carmo-Fonseca M, Geraghty F, Pereira HS, Grosveld F, Antoniou M. 1999. Inefficient processing impairs release of RNA from the site of transcription. *EMBO J 18*:2855–2866.

Das Gupta J, Gu H, Chernokalskaya E, Gao X, Schoenberg DR. 1998. Identification of two *cis*-acting elements that independently regulate the length of poly(A) on *Xenopus* albumin pre-mRNA. *RNA 4*:766–776.

Gilmartin GM, Fleming ES, Oetjen J, Graveley BR. 1995. CPSF recognition of an HIV-1 mRNA 3′-processing enhancer: Multiple sequence contacts involved in poly(A) site definition. *Genes & Dev 9*:72–83.

Gu H, Das Gupta J, Schoenberg DR. 1999. The poly(A)-limiting element is a conserved *cis*-acting sequence that regulates the length of poly(A) on nuclear pre-mRNAs. *Proc Natl Acad Sci USA 96*:8943–8948.

Gunderson SI, Beyer K, Martin G, Keller W, Boelens WC, Mattaj IW. 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell 76*:531–541.

Gunderson SI, Polycarpou-Schwarz M, Mattaj IW. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell 1*:255–264.

Gunderson SI, Vagner S, Polycarpou-Schwarz M, Mattaj IW. 1997. Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes & Dev 11*:761–773.

Levitt N, Briggs D, Gil A, Proudfoot NJ. 1989. Definition of an efficient synthetic poly(A) site. *Genes & Dev 3*:1019–1025.

Lou H, Gagel RF, Berget SM. 1996. An intron enhancer recognized by splicing factors activates polyadenylation. *Genes & Dev 10*:208–219.

Lou H, Helfman DM, Gagel RF, Berget SM. 1999. Polypyrimidine tract-binding protein positively regulates inclusion of an alternative 3′-terminal exon. *Mol Cell Biol 19*:78–85.

Moreira A, Takagaki Y, Brackenridge S, Wollerton M, Manley JL, Proudfoot NJ. 1998. The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3′ end formation by two distinct mechanisms. *Genes & Dev 12*:2522–2534.

Paillard L, Omilli F, Legagneux V, Bassez T, Maniey D, Osborne HB. 1998. EDEN and EDEN-BP, a *cis* element and an associated factor that mediate sequence-specific mRNA deadenylation in *Xenopus* embryos. *EMBO J 17*:278–287.

Pastori RL, Moskaitis JE, Buzek SW, Schoenberg DR. 1991. Coordinate estrogen-regulated instability of serum protein-coding messenger RNAs in *Xenopus laevis*. *Mol Endocrinol 5*:461–468.

Rao MN, Chernokalskaya E, Schoenberg DR. 1996. Regulated nuclear polyadenylation of *Xenopus* albumin pre-mRNA. *Nucleic Acids Res 24*:4078–4083.

Richter JD. 1999. Cytoplasmic polyadenylation in development and beyond. *Microbiol Mol Biol Rev 63*:446–456.

Rüegsegger U, Beyer K, Keller W. 1996. Purification and characterization of human cleavage factor Im, involved in the 3′ end processing of messenger RNA precursors. *J Biol Chem 271*: 6107–6113.

Salles FJ, Richards WG, Strickland S. 1999. Assaying the polyadenylation state of mRNAs. *Methods 17*:38–45.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Plainview, NY: Cold Spring Harbor Laboratory Press.

Schek N, Cooke C, Alwine JC. 1992. Definition of the upstream efficiency element of the simian virus-40 late polyadenylation signal by using in vitro analyses. *Mol Cell Biol 12*:5386–5393.

Schoenberg DR, Moskaitis JE, Smith JH Jr, Pastori RL. 1989. Extranuclear estrogen-regulated destabilization of *Xenopus laevis* serum albumin mRNA. *Mol Endocrinol 3*:805–814.

Sittler A, Gallinaro H, Jacob M. 1995. The secondary structure of the adenovirus-2 L4 polyadenylation domain: Evidence for a hairpin structure exposing the AAUAAA signal in its loop. *J Mol Biol 248*:525–540.

Vagner S, Vagner C, Mattaj IW. 2000. The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3′-end processing and splicing. *Genes & Dev 14*:403–413.

Wu L, Good PJ, Richter JD. 1997. The 36-kilodalton embryonic-type cytoplasmic polyadenylation element-binding protein in *Xenopus laevis* is ElrA, a member of the ELAV family of RNA-Binding proteins. *Mol Cell Biol 17*:6402–6409.

Zhao J, Hyman L, Moore C. 1999. Formation of mRNA 3′ ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev 63*:405–445.