Identification of two *cis*-acting elements that independently regulate the length of poly(A) on *Xenopus* albumin pre-mRNA

JAYDIP DAS GUPTA, HAIDONG GU, ELENA CHERNOKALSKAYA,¹
XIAOPING GAO,² and DANIEL R. SCHOENBERG

Department of Pharmacology and Comprehensive Cancer Center, The Ohio State University College of Medicine,
Columbus, Ohio 43210-1239, USA

ABSTRACT

Unlike most eukaryotic mRNAs studied to date, *Xenopus* serum albumin mRNA has a short (17-residue), discrete poly(A) tail. We recently reported that this short poly(A) tail results from regulation of the length of poly(A) on albumin pre-mRNA. The purpose of the present study was to locate the *cis*-acting element responsible for this, the poly(A)-limiting element or PLE. An albumin minigene consisting of albumin cDNA joined in exon 13 to the 3' end of the albumin gene produced mRNA with <20 nt poly(A) when transfected into mouse fibroblasts. This result indicates both that *cis*-acting sequences that regulate poly(A) length are within this construct, and that nuclear regulation of poly(A) length is conserved between vertebrates. Poly(A) length regulation was retained after replacing the terminal 53 bp and 3' flanking region of the albumin gene with a synthetic polyadenylation element (SPA). Conversely, fusing albumin gene sequence spanning the terminal 53 bp of the albumin gene and 3' flanking sequence onto the human β -globin gene yielded globin mRNA with a 200-residue poly(A) tail. These data indicate that the PLE resides upstream of the sequence elements involved in albumin pre-mRNA 3' processing. Poly(A) length regulation was restored upon fusing a segment bearing albumin intron 14, exon 15, and 3' flanking sequence onto the β -globin gene. We demonstrate that exon 15 contains two PLEs that can act independently to regulate the length of poly(A).

Keywords: albumin mRNA; mRNA 3'; processing; polyadenylation; poly(A) length regulation

INTRODUCTION

With the exception of histone mRNA, all eukaryotic mRNAs are polyadenylated during nuclear pre-mRNA processing. 3' Processing begins with the ordered assembly of a large macromolecular complex consisting of cleavage and polyadenylation specificity factor (CPSF) (Murthy & Manley, 1992; Bienroth et al., 1993), cleavage stimulation factor (CstF) (Takagaki et al., 1992), two cleavage factors (Rüegsegger et al., 1996), and poly(A) polymerase on the pre-mRNA. The pre-mRNA is cleaved approximately 20 nt downstream from AAUAAA and 25 nt upstream from a conserved U-rich element to generate the substrate for polyadenylation (Macdonald et al., 1994). The subsequent addition of poly(A) proceeds through two distinct steps. The first is

the slow addition of 10+ residues and requires only CPSF and poly(A) polymerase (Bienroth et al., 1993). The second is the rapid addition of 200–250 adenosines and requires the participation of nuclear poly(A)-binding protein II (PAB II) (Wahle, 1991).

3' Processing appears to be unregulated for many mRNAs. However, recent findings, primarily with viral RNAs, indicate that sequences upstream of AAUAAA can have significant effects (reviewed in Virtanen, 1995). For example, SV40 contains an upstream element (USE) consisting of three repeats of the motif AUUU GURA, which forms a specific complex with U1 snRNP by binding U1A in complex with U1 snRNA (Lutz et al., 1996). In addition, U1A protein itself can interact with CPSF to stabilize CPSF binding to AAUAAA (Lutz et al., 1996). In the case of HIV-1, sequences immediately adjacent to the 3' poly(A) site reduce the efficiency of 3' processing. However, a USE upstream of TAR interacts with the 160-kDa subunit of CPSF to enhance 3' processing (Gilmartin et al., 1995). In this case, the TAR stem-loop serves to bring the USE and AAUAAA into proximity. A different form of regulation is observed with U1A pre-mRNA. This RNA contains a

Reprint requests to: Daniel Schoenberg, Department of Pharmacology, The Ohio State University College of Medicine, 333 West 10th Avenue, Columbus, Ohio 43210-1239, USA; e-mail: schoenberg.3@osu.edu.

¹Present address: ESA, Inc., 22 Alpha Road, Chelmsford, Massachusetts 01824, USA.

²Present address: Genpower Inc., 6106 Hana Road, Edison, New Jersey, New Jersey 08817, USA.

complex structure in its 3' UTR that can bind two molecules of U1A protein (van Gelder et al., 1993). The complex then interacts with poly(A) polymerase to block poly(A) addition onto the cleaved transcript (Gunder-son et al., 1994). Thus, viral RNAs show regulation of the level of efficiency of 3' processing, and U1A RNA is regulated by the all-or-none addition of poly(A).

Other than studies on the mechanism of action of PAB II (Wahle, 1995), there is little direct evidence for regulation of the length of poly(A) added during pre-mRNA 3' processing. We became interested in this possibility from studies in this laboratory on the regulation of serum albumin mRNA stability in *Xenopus* liver. Estrogen administration to male *Xenopus* results in the coordinate destabilization of the mRNAs for the major serum proteins (Riegel et al., 1986; Schoenberg et al., 1989; Pastori et al., 1991a, 1992). Unlike a number of mRNAs, which undergo poly(A) shortening associated with their destabilization (Shyu et al., 1991; Stoeckle & Guan, 1993; Lagnado et al., 1994; Schiavi et al., 1994; Caponigro & Parker, 1995), albumin possesses a uniquely short, discrete poly(A) tail that is unchanged in the process of its degradation (Schoenberg et al., 1989). It appears that this short poly(A) serves to "mark" albumin for destabilization by an estrogen-induced endonuclease (Pastori et al., 1991b; Dompenciel et al., 1995). To determine the source of this unusual feature on albumin mRNA, we used an RT-PCR protocol to assay polyadenylation on albumin mRNA and pre-mRNA. Unexpectedly, we found that the short poly(A) tail on albumin results from regulation of poly(A) length on albumin pre-mRNA (Rao et al., 1996). In the present study, we map the *cis*-acting sequence element responsible for this, the poly(A)-limiting element or PLE, to a region upstream of the 3' processing signals at the 5' end of the terminal exon.

RESULTS

Albumin mRNA expressed from a minigene in transfected cells has short poly(A)

Our original quantification of poly(A) length on albumin mRNA used an assay [RNase A + T1 digestion of 3' end-labeled albumin mRNA (Schoenberg et al., 1989)] that allowed for precise measurement of the 17-nt poly(A) tail. However, the difficulty in obtaining sufficient quantities of pre-mRNA to perform poly(A) analysis led us in an earlier study (Rao et al., 1996) to adapt an RT-PCR assay for poly(A) length to determine the degree of polyadenylation of albumin pre-mRNA. Baurén and Wieslander (1994) and Nesic et al. (1995) showed that 3' processing and polyadenylation occur prior to splicing of the terminal intron. Based on this, we defined albumin pre-mRNA as that species of polyadenylated albumin mRNA bearing the terminal intron, and used a primer to this intron in combination with an oli-

go(dT) primer/adaptor to show that albumin pre-mRNA has the same short poly(A) tail as found on fully processed albumin mRNA (Rao et al., 1996). Unlike the discrete band seen by RNase A+T1 assay, the 17-nt albumin poly(A) tail appears as a smear or a group of closely spaced bands on high-resolution gels of RT-PCR products. This is because the oligo(dT) primer/adaptor can initially hybridize with any stretch of 12 or fewer A residues within the 17-nt albumin poly(A) tail. In light of this, poly(A) tails the length of that found on albumin mRNA by RT-PCR will be referred to here as <20 nt to match the observed data.

To localize the *cis*-acting element responsible for regulating poly(A) length, it was necessary to study this process in transfected cells. Because of the technical difficulty of working with transfected *Xenopus* hepatocytes, we first sought to determine whether poly(A) length regulation could be replicated in a cell culture model. For this, we constructed an albumin minigene consisting of albumin cDNA fused at an *EcoR* I site present in exon 13 of the albumin gene with a 2.3-kb genomic *EcoR* I fragment bearing the remaining portion of exon 13, the two terminal introns, exons 14 and 15, and 939 bp of 3' flanking sequence (see diagram in Fig. 1). Transcription of this and all other vectors used in this study was controlled by the CMV immediate early promoter.

The experiment in Figure 1A shows the results of RT-PCR analysis of poly(A) length on albumin pre-mRNA extracted from *Xenopus* liver (lane 2), and albumin pre-mRNA expressed in LM(tk⁻) cells transfected with the albumin minigene (lane 3). In this experiment, RT-PCR was performed using an unlabeled oligo(dT) primer/adaptor and a radiolabeled upstream primer to the terminal intron of albumin pre-mRNA (intron 14, open arrowhead). Identical results were obtained from both preparations, indicating that poly(A) length regulation is accurately reproduced in transfected mammalian cells. As an additional control, poly(A) analysis was performed on the same preparations using a primer in the terminal albumin exon (exon 15, filled arrowhead), because results using this primer will measure poly(A) length on both pre-mRNA and fully processed mRNA, with the latter being the predominant species in the preparation. The results in Figure 1B show albumin mRNA expressed in transfected L cells has the same <20 nt poly(A) tail as authentic albumin mRNA from liver, thus confirming the results obtained using the intron-specific primer.

The PLE is not located within the vicinity of the 3' processing signals

Most of the sequence elements involved in regulating 3' processing are found within 50 nt of the cleavage site (Virtanen, 1995). To determine whether the PLE was in this portion of albumin pre-mRNA, the terminal

FIGURE 1. Albumin pre-mRNA expressed from a minigene has <20 nt poly(A) in transfected L cells. **A:** Total RNA from *Xenopus* liver (lane 2) or from L cells transfected with the albumin minigene (lane 3) RNA were analyzed for poly(A) length by RT-PCR using a [32 P]-labeled albumin intron 14 primer (indicated with an open arrowhead at bottom), and unlabeled oligo(dT) primer/adaptor, and the products separated on a denaturing 6% acrylamide/urea gel. This primer combination was expected to yield a 193-bp limit product (open arrowhead). **B:** RNAs in A were analyzed as above using a radiolabeled primer from albumin exon 15 (filled arrowhead at bottom). Both samples were expected to yield a 145-bp limit product for PCR. **C:** Poly(A) tail length was determined on RNA expressed from the hybrid CMV-albumin minigene-SPA construct using the same exon 15 primer as in B.

53 bp and 3' flanking sequence of the albumin gene were replaced with a synthetic poly(A) sequence (SPA, Levitt et al., 1989) to generate the albumin minigene-SPA. The structure of this construct is shown below

Figure 1C and the sequence of the 3' portion of the hybrid gene is shown in Figure 2. Replacing the 3' 53 bp of the albumin gene with the SPA element resulted in a shorter product than that made from the albumin minigene. However, this product had the same <20 nt poly(A) tail as mRNA expressed from the minigene or liver albumin mRNA, indicating that the PLE resides upstream of the 3' processing signals in either the exon sequence of the albumin gene or in one of the two terminal introns present in these constructs. It should be noted that all of the RT-PCR products in this study were sequenced (e.g., Fig. 2) to identify the locations of the cleavage sites on the expressed mRNAs. These results also confirmed the RT-PCR data showing poly(A) tails of <20 nt on the products from the albumin minigene and minigene-SPA constructs.

To characterize further the elements involved in poly(A) length regulation, we developed a vector (CMV-glo-TAA, Fig. 3) in which a portion of exon 3 and the 3' UTR of the human β -globin gene were replaced by an in-frame stop codon and a multiple cloning site. As with the albumin minigene constructs, transcription is controlled by the CMV immediate-early promoter. We first used this vector to confirm that the 53 bp deleted from the end of exon 15 in Figure 1C are not involved in regulating poly(A) length. In the experiment in Figure 4, 3' processing of β -globin mRNA was controlled by the SPA element (CMV-glo-SPA, lane 2) or albumin sequences spanning from -53 to +138 relative to the pre-mRNA cleavage site [CMV-glo-alb(-53/+138), lane 3]. mRNA expressed from CMV-glo-SPA had the long, heterogeneous poly(A) tail expected for an RNA whose 3' processing was controlled by this optimized element. mRNA expressed from CMV-glo-alb(-53/+138) had a similarly long, heterogeneous poly(A) tail. This result confirmed that the PLE is not in the terminal 53 bp of albumin pre-mRNA.

Scanning densitometry was used to quantify poly(A) length and to provide a graphical comparison of poly(A) length on mRNA transcribed from CMV-glo-SPA and CMV-glo-alb(-59/+138) versus the <20 nt poly(A) tails seen on albumin pre-mRNA and mRNA. In the panel on the right side of Figure 4, the poly(A) tails on albumin pre-mRNA and total mRNA from control lanes of Figure 5 (see below) appear as discrete peaks that begin and return to baseline values. Background subtraction was not used here to avoid introducing bias into the data. In contrast, both CMV-glo-SPA and CMV-glo-alb(-59/+138) mRNA have poly(A) ranging from 12 nt (the limit for detection by this assay) to >250 nt.

Identification of multiple poly(A)-limiting elements in exon 15

The next series of constructs examined the effect of sequences further upstream of position -53 for their ability to regulate poly(A) length. In the experiment in

synthetic poly(A) element (SPA)

UCGACAAAUAAAGAUCUUUUUUUCAUUAGAUCUGUGUGUUGGUUUUUUGUGUGGCAUG

albumin -53/+138

GUCGACUUGCUUAGAGCAUUCACUGUGUGUUGUAAUAAAUAAGCAUUUAAAAAAAAUAU
 AUAGAGAGUGAAAAUGUAUUUAUCUGUUGUUUCAGUGUCACAUACACACCAGAACUAUAUG
 GUUUACAUAUUUUAGUCCGAUUUCCGUGAUCUCUUUCUACCUAGAAAAUUAAACAUCAGU
 UAGAUAGAAGGUACAG

albumin minigene SPA

CCUUCAAAACUCACUGAGGAACACCUUCCAUCUCUCAAAACACAAGAAAAAAAAAGUCCUUCA
 GCUGAAAAGAGCAU UCGACAAAUAAAGAUCUUUUUUUCAUUAGAUCUGUGUGUUGGUU
 UUUUGUGUGGCAUG

albumin exon 15

CCUUCAAAACUCACUGAGGAACACCUUCCAUCUCUCAAAACACAAGAAAAAAAAAGUCCUUCA
 GCUGAAAAGAGCAUUUGCUUAGAGCAUUCACUGUGUGUUGUAAUAAAUAAGCAUUUAA
 AAAAAUA

FIGURE 2. Sequences of the 3' ends of mRNAs expressed from the albumin minigene, the minigene-SPA, and globin-SPA constructs. The sequences of the synthetic polyadenylation element (SPA, Levitt et al., 1989) and the albumin gene spanning from 53 bp upstream of the poly(A) addition site to 139 bp downstream are shown in the first group. Below this are the sequences of the minigene-SPA construct, in which the terminal 53 bp of albumin exon 15 and 3' flanking sequence were replaced with SPA, and albumin exon 15. The AAUAAA polyadenylation signals are underlined, and the sites of cleavage and poly(A) addition are shown with a black dot. Locations of the cleavage sites were determined by sequence analysis of RT-PCR products generated as described for the experiments in Figure 1.

Figure 5, a portion of the albumin gene from 20 bp upstream of the 5' splice site of the terminal intron (intron 14) to 212 nt downstream of the cleavage site was inserted into CMV-glo-TAA. The starting material for this construction was the albumin minigene rather than the native albumin gene, because the former has a 19-bp tag from exon 2 of the β -globin gene inserted in a unique Sex A1 site located precisely at the intron 14/5' splice site to differentiate mRNA transcribed from this vector from endogenous albumin mRNA in transfected *Xenopus* hepatocytes. Thus, intron 14 and exon 15 are the only albumin sequences present in mRNA transcribed from CMV-glo-alb(intron 14/+212). RT-PCR analysis in Figure 5 of RNA extracted from transfected L cells using primers to both intron 14 and exon 15 showed that the hybrid globin pre-mRNA transcribed from this construct has the same <20 nt poly(A) tail as albumin mRNA extracted from *Xenopus* liver. This localized the PLE to either intron 14 or to the 5' 75 bp of the 128-bp exon 15 that remained intact in the minigene-SPA construct.

The deletion constructs shown on the bottom of Figure 6A were transfected into L cells, and poly(A) length was determined for the expressed pre-mRNAs using a radiolabeled upstream primer to globin intron 2. Delet-

FIGURE 3. Map of CMV-glo-TAA. The vector CMV-glo-TAA consists of the human β -globin gene terminated at a unique *EcoR* I site in exon 3 fused to CMV exon 1 and the CMV promoter. The first AUG is that of β -globin. An in-frame stop codon is followed by unique sites for Asp 718, *Kpn* I, *Ava* I, *Sma* I, *Xma* I, *Xba* I, and *Sal* I.

FIGURE 4. Sequence in the terminal 53 nt of albumin exon 15 does not regulate poly(A). Total RNA from L cells transfected with CMV-glo-SPA (lane 2) or CMV-glo-alb(-53/+138) (lane 3) were assayed for poly(A) length by RT-PCR using a [32 P]-labeled globin exon 3 primer (filled arrowhead, lower left) and unlabeled oligo(dT) primer/adaptor, and the products were separated by electrophoresis on a denaturing 6% acrylamide/urea gel. The expected minimal sizes for PCR products obtained by the oligo(dT) primer/adaptor binding adjacent to the poly(A) addition site were 152 bp for CMV-glo-alb(-53/+138) and 120 bp for CMV-glo-SPA. The right panel shows these data and the results from lanes 3 and 5 of Figure 5 analyzed by scanning densitometry. Positions of the size markers are indicated across the top of the scan.

ing intron 14 (-124/+212 construct) had no effect on poly(A) length regulation (Fig. 6A, lane 2). Poly(A) length regulation was also retained in deletions -103/+212 and -80/+212. It was not until exon 15 was deleted to the last 53 bp that poly(A) length regulation was lost. It should be noted that the -53/+212 construct was used in Figure 4. In that experiment, poly(A) length was determined using a globin exon 3 primer instead of the globin intron 2 primer used in Figure 6A. These data localized a functional PLE to the 27 bp between -80 and -54 within albumin exon 15.

In the course of mapping the location of the PLE, we prepared a deletion of CMV-glo-alb(intron 14/+212) in which all but the first 20 bp of exon 15 was replaced by the synthetic poly(A) site. In this hybrid globin gene (CMV-glo-intron 14-SPA), the only albumin sequences remaining were intron 14 and the first 20 bp of exon 15. Figure 6B shows the results of poly(A) analysis of mRNA expressed from this construct analyzed using a primer to globin exon 3 to insure the fidelity of splicing of the albumin intron. To our surprise, this construct also yielded mRNA with a <20 nt poly(A) tail. Careful ex-

FIGURE 5. Regulated polyadenylation in constructs bearing albumin intron 14 and exon 15. L cells were transfected with a hybrid construct that had the β -globin gene fused to the α' portion of the albumin gene spanning from the 5' splice site of intron 14 to 212 bp of 3' flanking sequence. RT-PCR poly(A) analysis was performed on total RNA from transfected cells (lanes 2 and 4) and compared to a control for albumin mRNA in liver (lanes 3 and 5). Analyses in lanes 2 and 3 used a radiolabeled primer for intron 14 (open arrowhead, bottom) to measure poly(A) length in pre-mRNA, and those in lanes 4 and 5 used a primer for exon 15 (filled arrowhead, bottom) to measure poly(A) length on both pre-mRNA and fully processed mRNA. The limit size for products amplified with the intron 14 primer is 193 bp, whereas the limit size for products amplified with the exon 15 primer is 145 bp. Albumin intron 14 is identified by the asterisk.

amination of exon 15 (see Fig. 7 and Discussion) identified a region of sequence similarity between the $-80/-53$ region containing a functional PLE and the upstream $-124/-104$ region. If the $-124/-104$ sequence could also impart poly(A) length regulation onto globin mRNA, it would indicate the presence of multiple, independent poly(A)-limiting elements. The most stringent test was to insert the 21 bp between -124 and -104 into CMV-glo-SPA so that this was the only

albumin sequence present. The data in Figure 6C, lane 3 show that pre-mRNA expressed from this construct has a <20 nt poly(A) tail, whereas pre-mRNA expressed from CMV-glo-SPA has the long, heterogeneous poly(A) tail observed in other experiments. The same result was obtained when poly(A) analysis was performed using a primer to globin exon 3 (Fig. 6C, lanes 4 and 5). In the latter sample, the greater separation achieved with these smaller products produced single-base resolution of the short poly(A) tail. These data indicate that albumin pre-mRNA has two independently functional poly(A)-limiting elements within the terminal exon, one between -124 and -104 , and a second between -80 and -53 .

DISCUSSION

Several years ago, we made the puzzling observation that albumin mRNA has a very short, discrete poly(A) tail (Schoenberg et al., 1989). Recent results confirmed this finding, and showed that this unusual feature results from regulation of poly(A) length on albumin pre-mRNA (Rao et al., 1996). The goals of the present study were to determine whether the regulation of poly(A) length on albumin mRNA was controlled by a specific *cis*-acting element, and, if so, to determine the location of this element.

The results in Figure 1 show that pre-mRNA transcribed from an albumin minigene containing the complete exon sequence, the two terminal introns, and a considerable portion of 3' flanking sequence has short (<20 nt) poly(A) in transfected murine fibroblasts. These data make several important points. First, they show that there is nothing unique about *Xenopus* hepatocytes with respect to the regulation of poly(A) tail length. Those factors involved in regulating poly(A) length are present in a totally unrelated cell line. Second, the fact that regulation of poly(A) length is observed in a cell that does not express albumin confirms that a specific *cis*-acting element, which we term the poly(A)-limiting element, or PLE, is involved in this process. Third, these data also suggest that nuclear regulation of poly(A) tail length is a phenomenon that may not be unique to *Xenopus* albumin. The number of mRNAs whose polyadenylation has been studied in detail is fairly limited. Hence, it is possible that albumin is just the first of what may be a larger class of mRNAs that show this form of regulation.

The data in Figure 6 identified PLEs between 53 and 80 bp, and 104 and 124 bp upstream of the cleavage site, both of which can act independently to regulate poly(A) length. The sequences of exon 15 and the $-53/-80$ and $-104/-124$ regions are shown in Figure 7. Both regions contain a similar eight-nucleotide sequence, CUGAA/GA/GAA/G. Experiments are now in progress to determine whether this element constitutes the PLE.

FIGURE 6. Identification of multiple PLEs in exon 15. **A:** Deletion mapping identifies a PLE between 80 and -54. A series of exon 15 constructs beginning 124, 103, 80, or 53 bp upstream of the cleavage site and terminating 212 bp downstream of the cleavage site were inserted into CMV-glo-TAA. Poly(A) length on the hybrid pre-mRNA expressed in transfected L cells was determined by RT-PCR using a primer to globin intron 2 (open arrowhead). The -53/+212 construct is the same as that analyzed in Figure 4. **B:** Poly(A) length regulation by a sequence in intron 14 or the 5' end of exon 15. A region spanning the terminal albumin intron (intron 14, asterisk) and the 5' 20 bp of exon 15 was cloned into CMV-glo-SPA. Poly(A) analysis was performed using a primer to globin exon 3 (filled arrowhead) to detect the product bearing spliced intron 14. **C:** Identification of a PLE between -124 and -104. The 21 bp spanning -124/-104 of albumin exon 15 (shaded) was inserted between the stop codon and the synthetic poly(A) element of CMV-glo-SPA. The length of poly(A) on the hybrid pre-mRNA expressed in transfected L cells was determined using radiolabeled primers to globin intron 2 (lanes 2 and 3, open arrowhead) or exon 3 (lanes 4 and 5, filled arrowhead).

Several levels of regulation have been found for pre-mRNA 3' processing. A number of studies have shown a link between splicing and 3' processing (Niwa et al., 1990; Niwa & Berget, 1991; Nesic & Maquat, 1994; Cooke & Alwine, 1996), both in vivo and in vitro. Elements upstream of AAUAAA in a number of viral RNAs

can enhance the efficiency of 3' processing. With these mRNAs, the predominant mode of action appears to be stabilization of CPSF binding (Lutz & Alwine, 1994; Gil-martin et al., 1995). In the case of the 3' processing of the pre-mRNA for U1 snRNP A protein, binding of two molecules of U1A to the 3' UTR of U1A pre-mRNA has

Although our data are consistent with regulation of poly(A) addition, we cannot formally rule out the possibility that albumin pre-mRNA receives a >200-nt poly(A) tail and, prior to splicing of the terminal intron, the PLE directs deadenylation to the 17 nt seen here and in our previous study. Paillard et al. (1998) identified an element (EDEN) consisting of several U(A/G) repeats in the 3' UTRs of Eg5 and *c-mos* mRNA, which is responsible for the deadenylation of these mRNAs at fertilization. Although 3' UTR of albumin mRNA has the sequence UGUGUGUUG just upstream of AAUAAA, it is unlikely that such a mechanism is functioning here because it resides downstream of -53 in sequences shown not to be involved in poly(A) length regulation (Fig. 1C). In addition, were albumin pre-mRNA to be processed by such a mechanism, we would have expected to see at least some population of molecules with poly(A) >200 nt on albumin pre-mRNA. We favor the idea that the PLE interacts with the 3' processing complex to affect poly(A) length regulation. One can speculate about a number of possible mechanisms by which length regulation might occur. As noted earlier, nuclear polyadenylation proceeds through a two-step process, the slow addition of 10+ residues, followed by the rapid addition of 200–250 (Wahle, 1991). PAB II acts to stimulate the second processive step of poly(A) addition and as a molecular ruler that limits the growing chain to 200–250 residues (Wahle, 1995). The PLE

In summary, we report here the first direct evidence for a *cis*-acting sequence, the poly(A)-limiting element or PLE, that regulates the length of poly(A) on nuclear pre-mRNA. The existence of the PLE raises new questions about the regulation of pre-mRNA 3' processing and its ultimate effects on gene expression. Experiments are in progress to map the precise nucleotides that constitute the PLE, to reconstitute regulated polyadenylation in vitro and to identify the protein(s) that bind to the PLE. The ability to manipulate the length of poly(A) added within the cell will provide a useful tool to address questions about the role of poly(A) in mRNA transport, translation, and mRNA stability.

RT-PCR assay for poly(A)

First-strand synthesis was performed in a 30- μ L reaction. Ten micrograms of total RNA was heat-denatured at 85 °C for 5 min in a 16.5- μ L volume containing 300 ng of oligo (dT) anchor XG2 (5'-GGGGATCCGCGGTTTTTTTTTT). This was cooled on ice and an additional 13.5 μ L of a mixture containing 6 μ L of 5 $\times$ first-strand buffer, 3 μ L of 0.1 M dithiothreitol, 2 μ L of 10 mM dNTPs, 0.5 μ L of RNasin, and 2 μ L MMLV reverse transcriptase (200 units/ μ L, Life Technologies, Inc.)

was added. This was incubated at 42 °C for 1.5 h followed by heating for 10 min at 85 °C for 10 min to inactivate the polymerase. One-tenth of this reaction mixture (for samples from transfected cells) or 1/30th of this reaction mixture (for frog liver RNA) was used for subsequent PCR amplification. In the experiments in Figure 6A and C, the primer was removed on a Sephadex G-50 spin column and PCR was performed using a primer of the same sequence lacking oligo(dT) (XG4). Hot-start PCR was performed in a 25- μ L reaction containing the first-strand cDNA synthesis product separated by wax from the second portion containing 40 ng of 5' [³²P]-labeled specific primer, 0.16 mM dATP, dCTP, dGTP, and dTTP, 6 mM MgCl₂, and 2.5 units of Taq DNA polymerase in 1 $\times$ PCR buffer provided with the enzyme (Life Technologies, Inc.). Amplification conditions were 94 °C/2 min followed by 25 cycles of 94 °C/1 min, 60 °C/1 min (for the albumin exon primer), 66 °C/30 s (for the globin intron primer) or 70 °C/30 s (for the globin exon 3 primer), and 57 °C/2 min (for the albumin intron primer). The products of the reaction were denatured by heating at 90 °C for 5 min in formamide loading and electrophoresed on a 6% acrylamide/urea gel. The albumin exon 15 primer used was 5'-CAAACCTACTGAGGGAACACC and the intron 14 primer used was 5'-CATCATACATTTGTCTTAG. The globin exon 3 primer was 5'-GGCAACGTGCTGGTCTGTGT and the intron 2 primer was 5'-CTGAGTCCAAGCTAGGCCCT. The data in Figures 4 and 5 were analyzed without baseline correction at a 169- μ pixel resolution using a Protein Data-bases model 420oe densitometer.

Cell transfection

L cells were maintained in Dulbecco's Minimal Essential Medium plus 5% fetal calf serum. They were transfected in 60-mm dishes with 5 μ g plasmid/dish by calcium phosphate precipitation using a kit and protocols from Life Technologies, Inc. (Figs. 1, 4, 5) or with Superfect reagent from Qiagen (Fig. 6). After 48 h, the medium was removed and the cells were rinsed with PBS. Nuclear RNA extraction followed published protocols (Sambrook et al., 1989). The extracted RNA was treated with RNase-free DNase prior to its use for cDNA synthesis.

Expression plasmids for poly(A) analysis

The plasmids CMV-glo-SPA and CMV-glo-alb(-59/+138) were based on the original expression plasmid CMVGLOB, which was a gift from Dr. Jeff Ross. *EcoR* I digestion of CMVGLOB yielded a 2.3-kb fragment containing the CMV immediate early promoter and 112 bp of CMV exon 1 joined to exon 1 of the human β -globin gene. The first AUG codon is that of β -globin. *EcoR* I digestion cleaved the β -globin gene at position 1497 (relative to the first nucleotide of the β -globin gene), within exon 3. This was ligated into the RI site of the plasmid pUC18-TAA, which contained a stop codon engineered into an oligonucleotide bearing *EcoR* I and *Kpn* I complementary ends so that the insertion of the 2.3-kb CMV-globin fragment would place the stop codon in frame within globin exon 3. The product of this was called CMV-glo-TAA. To make CMV-glo-SPA, oligonucleotides for both strands of a synthetic poly(A) site (SPA, Levitt et al., 1989) were synthe-

sized with overhanging *Sal* I and *Sph* I cohesive ends at the 5' and 3' termini, respectively (the upper strand sequence was 5'-TCGACAATAAAAGATCTTTATTTTCATTAGATCTGTG TGTTGGTTTTTGTGTGGCATG). The annealed oligos were ligated into CMV-glo-TAA that was first linearized with *Sal* I, then partially digested with *Sph* I and gel purified. The latter step was necessitated by the presence of a second *Sph* I site within the plasmid. A fragment spanning position -59/+138 relative to the site of poly(A) addition was prepared by PCR from a cloned fragment of the 68-kDa albumin gene. The primers used (5'-GCGCGCGTCTGACTTGCTTAGAGCATTC AACTG and GCGCGCATGCTGTACCTTCTTATCTAACTG) were designed to introduce *Sal* I and *Sph* I sites into the final product for cloning into CMV-glo-TAA as above to yield CMV-glo-alb(-59/+138).

The albumin minigene was constructed by ligating a 1,736-bp *EcoR* I fragment of the 68-kDa albumin cDNA to a 2.3-kb *EcoR* I fragment of the 68-kDa albumin gene bearing a portion of exon 13, exons 14 and 15, introns 13 and 14, and 939 bp of 3' flanking sequence (EMBL accession no. X13604) cloned in pGEM3Z. The CMV promoter from CMV-glo-alb(-59/+138) was prepared by PCR using primers bearing *Sal* I ends. This was digested with *Sal* I and ligated to the *Sal* I site of the above plasmid. Finally, a 19-bp tag from exon 2 of human β -globin (5'-ACGTGGATCCTGA GAATA) bearing terminal Sex AI sequences (5'-CCAGGT) was cloned into the single Sex AI site at the 3' end of exon 14. The latter serves as a tag for differentiating albumin mRNA expressed from this plasmid from authentic albumin mRNA in transfected hepatocytes.

To make the albumin minigene-SPA, the plasmid CMV-Glo-Alb/SPA was made, which contained a hybrid 3' UTR with albumin -53 sequence up to poly(A) addition site and SPA flanking sequence. Synthetic primers were designed with Albumin/SPA junctions: EC26 5'-CATTTAAAAAATATCTGT GTGTTG and EC27 5'-CAACACACAGATATTTTTTAAATTG. Plasmid CMV-glo-alb(-59/+138) was PCR amplified using globin-specific primer SEQ4 (5'-ATCAGGGTAATTTTGCA) and EC27, and CMV-glo-SPA was amplified using vector-specific primer SEQ5 (GTAAGTGGAGTGCACCA) and EC26. Because PCR products obtained with EC26 and EC27 had complementary ends, they were gel purified and reamplified using SEQ4 and SEQ5 primers and each other for a template. The resultant PCR product was restricted with *Xba* I and *Nar* I and inserted into restricted CMV-glo-alb(-59/+138) plasmid replacing the albumin 3'UTR with synthetic Alb/SPA UTR. To obtain the minigene-SPA, the two halves of the plasmids CMV-glo-alb/SPA and the albumin minigene were cut with *Spe* I and *Sal* I and ligated together. Because the last plasmid did not contain *Sal* I site between the albumin coding sequence and -53 position in 3' UTR, this site was incorporated by PCR with EC56 (5'-GACGACAGTCGACATGCTC TTTTCAGCTG) and a CMV promoter-specific primer EC41 (5'-TCGATCCGGGCAACGTTG). The resulting 4-kb PCR fragment was cut with *Spa* I and *Sal* I and ligated with CMV-glo-alb/SPA. To make CMV-glo-alb(intron14/+212), a portion of the albumin minigene spanning the 20 bp upstream of the 5' splice site of intron 14 to 212 nt downstream of the poly(A) addition site was amplified from the minigene using the primers 5'-CATAGGTACCAGGTACGTG (XG7) and 5'-CCGTCT AGAAATACACTGTACTG (XG8). The primers incorporated a *Kpn* I site at the 5' end and an *Xba* I site at the 3' end.

Following restriction digestion with these two enzymes, the DNA was ligated into the CMV-glo-TAA that had been digested with the same enzymes. For CMV-glo-intron14-SPA, albumin intron 14 and 20 nt of sequence flanking both the 5' and 3' splice sites was amplified from the minigene using the 5' primer XG7 and EC73 (5'-CAGATCTAGAAGGTGTTCC TCAGTG). The product was digested with *Kpn* I and *Xba* I as above and inserted into similarly digested CMV-glo-SPA. The exon 15 deletion set was prepared by amplifying the albumin gene using the 5' primer XG3 (5'-CAAACCTCACTGAGGA ACACC) and the 3' primer XG8 (−124/+212), JOY20 (5'-TCCATCTCTCAAACACAAGAAA) and XG8 (−103/+212), and JOY21 (5'-AAAAAGTTCCTTCAGCTGAA) and XG8 (−80/+212). All of the products were cleaved with *Xba* I and cloned into CMV-glo-TAA digested with *Sma* I and *Xba* I. CMV-glo(−124/−104)-SPA was prepared by first annealing primers XG3 and EC73, then filling in the ends with Klenow fragment of DNA polymerase. The product was digested with *Xba* I and cloned into CMV-glo-SPA digested with *Sma* I and *Xba* I.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants GM38277 and GM55407 from the National Institutes of Health to D.R.S., and center grant P30 CA16058 to the Ohio State University Comprehensive Cancer Center. We thank Kirsi Multanen for her expert technical assistance and Dr. Greg Gilmartin for helpful comments on this manuscript.

Received November 25, 1997; returned for revision January 9, 1998; revised manuscript received April 14, 1998

REFERENCES

- Bauren G, Wieslander L. 1994. Splicing of balbiani ring-1 gene pre-messenger RNA occurs simultaneously with transcription. *Cell* 76:183–192.
- Bienroth S, Keller W, Wahle E. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J* 12:585–594.
- Caponigro GR, Parker R. 1995. Multiple functions for poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes & Dev* 9:2421–2432.
- Colgan DF, Murthy KGK, Prives C, Manley JL. 1996. Cell-cycle related regulation poly(A) polymerase by phosphorylation. *Nature* 384:282–285.
- Cooke C, Alwine JC. 1996. The cap and the 3' splice site similarly affect polyadenylation efficiency. *Mol Cell Biol* 16:2579–2584.
- Dompenciel RE, Garnepudi VR, Schoenberg DR. 1995. Purification and characterization of an estrogen-regulated *Xenopus* liver poly-somal nuclease involved in the selective destabilization of albumin mRNA. *J Biol Chem* 270:6108–6118.
- 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.
- Gunderson SI, Beyer K, Martin G, Keller W, Boelens WC, Mattaj JW. 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* 76:531–541.
- Keller W. 1995. 3' End cleavage and polyadenylation of nuclear messenger RNA precursors. In: Lamond AI, ed. *Pre-mRNA processing*. Austin, Texas: R.G. Landes. pp 113–134.
- Lagnado CA, Brown CY, Goodall GJ. 1994. AUUUA is not sufficient to promote poly(A) shortening and degradation of an mRNA: The functional sequence within AU-rich elements may be UUAUUU A(U/A)(U/A). *Mol Cell Biol* 14:7984–7995.
- Levitt N, Briggs D, Gil A, Proudfoot NJ. 1989. Definition of an efficient synthetic poly(A) site. *Genes & Dev* 3:1019–1025.
- Lutz CS, Alwine JC. 1994. Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes & Dev* 8:576–586.
- Lutz CS, Murthy KGK, Schek N, O'Connor JP, Manley JL, Alwine JC. 1996. Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes & Dev* 10:325–337.
- Macdonald CC, Wilusz J, Shenk T. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol Cell Biol* 14:6647–6654.
- Murthy KG, Manley JL. 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J Biol Chem* 267:14804–14811.
- Nesic D, Maquat LE. 1994. Upstream introns influence the efficiency of final intron removal and RNA 3'-end formation. *Genes & Dev* 8:363–375.
- Nesic D, Zhang J, Maquat LE. 1995. Lack of an effect of the efficiency of RNA 3'-end formation on the efficiency of removal of either the final or the penultimate intron in intact cells. *Mol Cell Biol* 15:488–496.
- Niwa M, Berget SM. 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes & Dev* 5:2086–2095.
- Niwa M, Rose SD, Berget SM. 1990. In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes & Dev* 4:1552–1559.
- Paillard L, Omilli R, Legagneaux 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. 1991a. Coordinate estrogen-regulated instability of serum protein-coding messenger RNAs in *Xenopus laevis*. *Mol Endocrinol* 5:461–468.
- Pastori RL, Moskaitis JE, Buzek SW, Schoenberg DR. 1992. Differential regulation and polyadenylation of transferrin mRNA in *Xenopus* liver and oviduct. *J Steroid Biochem Mol Biol* 42:649–657.
- Pastori RL, Moskaitis JE, Schoenberg DR. 1991b. Estrogen-induced ribonuclease activity in *Xenopus* liver. *Biochemistry* 30:10490–10498.
- Rao MN, Chernokalskaya E, Schoenberg DR. 1996. Regulated nuclear polyadenylation of *Xenopus* albumin pre-mRNA. *Nucleic Acids Res* 24:4078–4083.
- Riegel AT, Martin MB, Schoenberg DR. 1986. Transcriptional and post-transcriptional inhibition of albumin gene expression by estrogen in *Xenopus* liver. *Mol Cell Endocrinol* 44:201–209.
- 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.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Plainview, New York: Cold Spring Harbor Laboratory Press.
- Schiavi SC, Wellington CL, Shyu AB, Chen CYA, Greenberg ME, Belasco JG. 1994. Multiple elements in the *c-fos* protein-coding region facilitate messenger RNA deadenylation and decay by a mechanism coupled to translation. *J Biol Chem* 269:3441–3448.
- Schoenberg DR, Moskaitis JE, Smith JH Jr, Pastori RL. 1989. Extracellular estrogen-regulated destabilization of *Xenopus laevis* serum albumin mRNA. *Mol Endocrinol* 3:805–814.
- Shyu AB, Belasco JG, Greenberg ME. 1991. Two distinct destabilizing elements in the *c-fos* message trigger deadenylation as a first step in rapid mRNA decay. *Genes & Dev* 5:221–231.
- Stoeckle MY, Guan L. 1993. High-resolution analysis of gro-alpha messenger RNA poly(A) shortening—Regulation by interleukin-1β. *Nucleic Acids Res* 21:1613–1617.
- Takagaki Y, Macdonald CC, Shenk T, Manley JL. 1992. The human 64-kDa polyadenylation factor contains a ribonucleoprotein-type RNA binding domain and unusual auxiliary motifs. *Proc Natl Acad Sci USA* 89:1403–1407.

- van Gelder CWG, Gunderson SI, Jansen EJR, Boelens WC, Polycarpou-Schwarz M, Mattaj JW, Vanvenrooij WJ. 1993. A complex secondary structure in U1A pre-messenger RNA that binds two molecules of U1A protein is required for regulation of polyadenylation. *EMBO J* 12:5191–5200.
- Virtanen A. 1995. Strategies for regulating nuclear pre-mRNA polyadenylation. In: Lamond AI, ed. *Pre-mRNA processing*. Austin, Texas: R.G. Landes. pp 135–149.
- Wahle E. 1991. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 66:759–768.
- Wahle E. 1995. Poly(A) tail length control is caused by termination of processive synthesis. *J Biol Chem* 270:2800–2808.
- Zhao WQ, Manley JL. 1996. Complex alternative RNA processing generates an unexpected diversity of poly(A) polymerase isoforms. *Mol Cell Biol* 16:2378–2386.