CPSF recognition of an HIV-1 mRNA 3'-processing enhancer: multiple sequence contacts involved in poly(A) site definition

Gregory M. Gilmartin,¹ Elizabeth S. Fleming, Joyce Oetjen, and Brenton R. Graveley

Department of Microbiology and Molecular Genetics, The Markey Center for Molecular Genetics, Vermont Cancer Center, University of Vermont, Burlington, Vermont 05405 USA

The endonucleolytic cleavage and polyadenylation of a pre-mRNA in mammalian cells requires two *cis*-acting elements, a highly conserved AAUAAA hexamer and an amorphous U- or GU-rich downstream element, that together constitute the "core" poly(A) site. The terminal redundancy of the HIV-1 pre-mRNA requires that the processing machinery disregard a core poly(A) site at the 5' end of the transcript, and efficiently utilize an identical signal that resides near the 3' end. Efficient processing at the 3' core poly(A) site, both *in vivo* and *in vitro*, has been shown to require sequences 76 nucleotides upstream of the AAUAAA hexamer. In this report we demonstrate that this HIV-1 upstream element interacts directly with the 160-kD subunit of CPSF (cleavage polyadenylation specificity factor), the factor responsible for the recognition of the AAUAAA hexamer. The presence of the upstream element in the context of the AAUAAA hexamer directs the stable binding of CPSF to the pre-mRNA and enhances the efficiency of poly(A) addition in reactions reconstituted with purified CPSF and recombinant poly(A) polymerase. Our results indicate that the dependence of HIV-1 3' processing on upstream sequences is a consequence of the suboptimal sequence context of the AAUAAA hexamer. We suggest that poly(A) site definition involves the recognition of multiple heterogeneous sequence elements in the context of the AAUAAA hexamer.

[Key Words: Polyadenylation; 3' processing; poly(A) site; CPSF; RNA recognition]

Received September 8, 1994; revised version accepted November 23, 1994.

The biosynthesis of all mRNA in higher eukaryotes, with the exception of nonpolyadenylated histone message, requires the endonucleolytic cleavage and polyadenylation of the primary transcript. Despite the ubiquitous nature of mRNA polyadenylation, the sequences that direct pre-mRNA 3' processing are surprisingly diverse (for review, see Wahle and Keller 1992). In mammalian cells, only the AAUAAA hexamer, which resides 10–30 nucleotides upstream of the cleavage site, exhibits a high degree of sequence conservation among poly(A) sites. Immediately downstream of the cleavage site is an amorphous U- or GU-rich element that possesses neither an identifiable consensus sequence nor a conserved secondary structure. Together, these two elements appear to be sufficient, at least in some cases, to direct 3' processing *in vivo* (Levitt et al. 1989).

The factors responsible for mRNA 3' processing in mammalian cells have been characterized in considerable detail (for review, see Wahle and Keller 1992). The assembly of a 3'-processing complex appears to be initi-

ated by the binding of CPSF (cleavage polyadenylation specificity factor) to the pre-mRNA by virtue of its recognition of the AAUAAA hexamer (Bienroth et al. 1991; Murthy and Manley 1992). The interaction of a second factor, CstF (cleavage stimulation factor), which recognizes the downstream element in the context of the CPSF/AAUAAA complex (Takagaki et al. 1990; MacDonald et al. 1994), results in the establishment of a stable ternary complex (Gilmartin and Nevins 1989). The endonucleolytic cleavage of the pre-mRNA requires CPSF, CstF, CFI, CFII and, in some cases, poly(A) polymerase (Christofori and Keller 1989; Gilmartin and Nevins 1989; Takagaki et al. 1989). Subsequent poly(A) addition, however, requires only CPSF and poly(A) polymerase.

As an obligatory step in mRNA biosynthesis, 3' processing may contribute to the regulation of gene expression through the modulation of both the level and coding capacity of the mature message. Modulation of mRNA protein-coding capacity by alternative polyadenylation is illustrated by the regulation of IgM heavy chain mRNA synthesis (for review, see Guise et al. 1989). Poly(A) site choice contributes to the production of the secreted ver-

¹Corresponding author.

sus the membrane form of the μ heavy chain protein during the course of B-cell development. The efficiency of processing has been shown to correlate with the stability of the CPSF/CstF/RNA ternary complex (Weiss et al. 1991). The nature of the highly variable downstream element of the poly(A) site has a considerable impact on ternary complex stability and may therefore contribute to poly(A) site selection. Work by Mann et al. (1993) suggests that the modulation of the activity of CstF, the factor responsible for recognition of the downstream element, may play a key role in the regulation of adenovirus poly(A) site selection.

It has become increasingly clear that sequences outside of the AAUAAA hexamer and downstream element of the "core" poly(A) site may contribute to the efficiency of mRNA 3' processing. Sequences that enhance 3' processing have been identified upstream of the SV40 late (Carswell and Alwine 1989; Schek et al. 1992), adenovirus L1 (DeZazzo and Imperiale 1989; DeZazzo et al. 1991a), L3 (Prescott and Falck-Pedersen 1992, 1994), and L4 (Sittler et al. 1994), hepatitis B virus (Rusnak and Ganem 1990; Rusnak 1991; Cherrington et al. 1992), and HIV-1 (Brown et al. 1991; DeZazzo et al. 1991b; Valsamakis et al. 1991; Cherrington and Ganem 1992) core poly(A) sites. These elements function in an orientation- and position-dependent manner, and several appear to be functionally analogous (Rusnak and Ganem 1990; Valsamakis et al. 1991). Although generally U-rich, these elements exhibit little sequence or structural similarity.

Sequences upstream of the HIV-1 poly(A) site have been of considerable interest with respect to their role in poly(A) site selection. The terminal redundancy of the HIV-1 pre-mRNA requires that the processing machinery disregard a core poly(A) site at the 5' end of the transcript and efficiently utilize an identical signal that resides near the 3' end (for review, see Coffin and Moore 1990). An upstream element unique to the 3' end of the message has been found to be necessary for efficient processing at the HIV-1 poly(A) site both in vivo (Valsamakis et al. 1991) and in vitro (Gilmartin et al. 1992; Valsamakis et al. 1992). Further analysis suggested that this upstream element may enhance processing through the stabilization of the processing complex at the core poly(A) site. The enhancement of both processing efficiency and complex stability was shown to require the spatial juxtaposition of the upstream element and the core poly(A) site by the RNA stem-loop structure of the TAR (*trans*-activation region) (Gilmartin et al. 1992).

To address the mechanism by which the HIV-1 upstream element serves as an enhancer of 3' processing, we first examined the function of each of the sequence elements that comprise the poly(A) site. Our results indicate that the requirement for an upstream element is a consequence of the suboptimal sequences that flank the HIV-1 AAUAAA hexamer. We show that the direct interaction of the 160-kD subunit of CPSF with the upstream element enhances the stability of CPSF binding and the efficiency of processing. This study suggests that the recognition of the AAUAAA hexamer in the context of multiple heterogeneous sequence elements may be a

central component of the definition of a mammalian poly(A) site.

Results

The sequence context of the HIV-1 AAUAAA hexamer necessitates the presence of an upstream element for efficient processing

The HIV-1 upstream element is unique to the 3' end of the viral pre-mRNA, where it may serve a key role in the selection of the 3' core poly(A) site in preference to an identical core poly(A) site at the 5' end of the message. To understand the mechanism by which the upstream element enhances 3' processing, we have examined the function of each of the sequence elements that comprise the HIV-1 poly(A) site. Our approach was to generate a series of chimeric RNAs in which sequences of the HIV-1 poly(A) site were systematically replaced with the corresponding sequences from the efficient adenovirus L3 poly(A) site. The HIV-1 poly(A) site was partitioned into four segments, depicted in Figure 1, A and B. Segment A included the HIV-1 wild type (*use*) or mutant (Δ *use*) upstream element and TAR. The Δ *use* RNAs contained the linker substitution mutation within the upstream element (NXS 9571–9588) shown to reduce processing efficiency both in vivo (Valsamakis et al. 1991) and in vitro (Gilmartin et al. 1992; Valsamakis et al. 1992). Segment B comprised 14 nucleotides between TAR and the AAUAAA hexamer; segment C contained the sequences from the AAUAAA to the cleavage site; segment D contained the downstream element. Each segment was replaced, individually or in combination, with the corresponding segment of the L3 poly(A) site. The cleavage efficiency of each poly(A) site was then assayed in a DEAE–Sepharose-fractionated HeLa cell nuclear extract (see Materials and methods) that contained all of the known processing factors (Gilmartin and Nevins 1989).

The 5' and 3' cleavage products of each of the chimeric RNAs are displayed in Figure 1, A and B. Figure 1A shows the cleavage products of RNAs that possess an HIV-1 downstream element, and Figure 1B contains RNAs that possess an L3 downstream element. The input RNA substrates were uniformly labeled with [α - 32 P]UTP during transcription, and the molarity of substrate RNA in each reaction was identical. Because each of the input substrates and 5' products differ in length and sequence, each will have a different specific activity. Consequently, visual inspection of the 5' products is relatively uninformative. In contrast, the 3' products are of identical length and sequence and therefore of identical specific activity—allowing for direct visual comparison in Figure 1, A and B. In addition, the comparison of identical 3' products eliminates the possibility of differential RNA stabilities among the processed products of the different RNA substrates. The RNAs were processed in a DEAE-fractionated HeLa cell nuclear extract to avoid the instability of 3' processing products inherent in the use of crude nuclear extract.

Figure 1. Processing efficiency of chimeric HIV-1/adenovirus L3 poly(A) sites. The HIV-1 poly(A) site was divided into four segments, A–D, which were systematically substituted with the corresponding sequences from the adenovirus L3 poly(A) site. The 5' and 3' cleavage products generated upon incubation of equimolar amounts of each RNA substrate in a DEAE–Sephacrose-fractionated HeLa cell nuclear extract and electrophoresis on a denaturing 10% polyacrylamide gel are illustrated. Note that the input substrates and 5' products differ in both length and sequence and, therefore, in specific activity. In contrast, the 3' products are of identical length, sequence, and specific activity within each experiment, thus allowing for direct visual comparison. (A) Cleavage products of RNAs that contain an HIV-1 downstream element (segment D). The cleavage efficiencies are relative to that of the wild-type HIV-1 poly(A) site (*use/CPS*), which was arbitrarily set to 100%. (B) Cleavage products of RNAs that contain an L3 downstream element (segment D). The cleavage efficiencies are relative to that of the wild-type L3 poly(A) site (L3), which was arbitrarily set to 100%. (*use*) Upstream element; (*dse*) downstream element. (Δ) The upstream element substitution mutation (NXS 9571–9588; Valsamakis et al. 1991).

Taken together, the data support two main conclusions. First, the sequences 5' to the AAUAAA hexamer of both the HIV-1 and L3 poly(A) sites are functionally equivalent. Replacement of both the HIV-1 A and B segments with L3 sequence had little effect compared with wild-type HIV-1 (Fig. 1A, cf. ΔAB , lane 10, and *use/CPS*, lane 1). Replacement of both the L3 A and B segments with HIV-1 sequence also had little effect compared with wild-type L3 (Fig. 1B, cf. *use/ ΔCD* , lane 2, and L3, lane 12). In addition, the downstream element had little impact on cleavage efficiency (data not shown).

The second conclusion to be derived from this experiment is that the inefficiency of the HIV-1 core poly(A) site is a consequence of sequences adjacent to the AAUAAA hexamer. The replacement of the C segment of the HIV-1 poly(A) site had the most dramatic effect on cleavage efficiency—in the context of both wild-type and mutant upstream elements. In the context of the upstream element mutation, the replacement of the HIV-1 C region with that of L3 restored processing efficiency to near wild-type HIV-1 levels (Fig. 1A, cf. *$\Delta use/CPS$* , lane 5, and *$\Delta use/\Delta C$* , lane 7). In contrast, replacement of the HIV-1 B segment alone had no effect in the context of the wild-type upstream element (Fig. 1A, cf. *use/CPS*, lane 1, and *use/ ΔB* , lane 2; Fig. 1B, cf. *use/ ΔD* , lane 1, and *use/ ΔBD* , lane 3), and a small effect in the context of the upstream element mutation (Fig. 1A, cf. *$\Delta use/CPS$* , lane 5, and *$\Delta use/\Delta B$* , lane 6; Fig. 1B, cf. *$\Delta use/\Delta D$* , lane 5, and *$\Delta use/\Delta BD$* , lane 7). In the context of the upstream element mutation, replacement of both the B and C segments of the HIV-1 poly(A) site with L3 sequence resulted in a further enhancement of processing efficiency (Fig. 1A, cf. *$\Delta use/\Delta BC$* , lane 8, and *$\Delta use/\Delta B$* , lane 6, and *$\Delta use/\Delta C$* , lane 7; Fig. 1B, cf. *$\Delta use/\Delta BCD$* , lane 8, and *$\Delta use/\Delta BD$* , lane 7, and *$\Delta use/\Delta CD$* , lane 6). In the complementary set of experiments, the replacement of the B and C segments of the L3 poly(A) site with HIV-1 sequence reduced processing to one-third of the wild-type L3 level (Fig. 1B, cf. ΔAD , lane 9, and L3, lane 12). In the context of the HIV-1 upstream element mutation, replacement of the L3 B and C segments with HIV-1 sequence resulted in a 10-fold reduction in processing efficiency (Fig. 1B, cf. *$\Delta use/\Delta BCD$* , lane 8, and *$\Delta use/\Delta D$* , lane 5). These results indicate that the sequences adjacent to the AAUAAA hexamer have a significant impact on processing efficiency and are primarily responsible for the inefficiency of the HIV-1 core poly(A) site.

Sequences 5' to the AAUAAA hexamer influence the stability of the interaction of CPSF with the pre-mRNA

The observation that the HIV-1 AAUAAA hexamer resides within a suboptimal sequence context suggested that AAUAAA recognition by CPSF may be impaired. We have therefore examined the binding of CPSF to the pre-mRNA. It had been shown previously that the HIV-1 upstream element stabilized the interaction of processing factors at the poly(A) site in a HeLa cell nuclear extract (Gilmartin et al. 1992). We have now extended this

observation to examine the interaction of purified CPSF at the HIV-1 poly(A) site.

CPSF was purified to apparent homogeneity, essentially as described by Bienroth et al. (1991) (see Materials and methods). The three polypeptides of 160, 105, and 73 kD that comprise CPSF (Bienroth et al. 1991; Murthy and Manley 1992) are illustrated in Figure 2A (lane 1). The presence of the 30-kD polypeptide observed by Bienroth et al. (1991) was not evident on this silver-stained gel. Figure 2B demonstrates that the purified CPSF is capable of supporting AAUAAA-dependent polyadenylation of a pre-cleaved HIV-1 pre-mRNA substrate when combined with recombinant calf thymus poly(A) polymerase (Wahle et al. 1991). CPSF is also capable of forming an AAUAAA-dependent complex with a full-length HIV-1 pre-mRNA as assayed by gel mobility shift (Fig. 2C, lanes 1 and 2). CPSF also interacts with CstF (Fig. 2A, lane 2) to form a slower-migrating AAUAAA-dependent complex (Fig. 2C, lanes 7,8). Consistent with earlier observations in crude HeLa cell nuclear extract (Gilmartin et al. 1992), the amount of both the CPSF/RNA and CPSF/CstF/RNA complexes formed on the RNA possessing a mutation within the upstream element (*$\Delta use/CPS$*) was dramatically reduced relative to the wild-type RNA (*use/CPS*) (Fig. 2C, cf. lanes 1 and 3 and lanes 7 and 9). The CPSF and CstF fractions illustrated in Figure 2A were used in these and subsequent experiments, unless otherwise noted.

CPSF was incubated with a series of precleaved RNAs, and the stability of the CPSF/RNA complex was assayed by gel mobility shift (Fig. 3). Experiments done in parallel using full-length pre-mRNAs yielded essentially identical results (data not shown). The precleaved RNAs are shown, because they are used in subsequent polyadenylation assays. Following a short incubation of the uniformly ^{32}P -labeled pre-mRNA and CPSF at 30°C to allow for complex formation, an ~100-fold molar excess of unlabeled *use/CPS* pre-mRNA was added to the reaction to sequester CPSF that dissociated from the initial complexes formed on the labeled pre-mRNA. Aliquots were taken at various times, heparin was added, and the samples were electrophoresed on a native 3% polyacrylamide gel.

As illustrated in Figure 3A, the wild-type L3 (lanes 1–5) and HIV-1 (lanes 6–10) pre-mRNAs formed more stable complexes with CPSF than did the *$\Delta use/CPS$* RNA (lanes 11–15), for which complexes were barely detectable at the initial time point. The ability of L3 sequences to increase the stability of CPSF binding to the *$\Delta use/CPS$* RNA was then examined. The replacement of sequences 5' to the AAUAAA hexamer (segments A and B) with L3 sequences increased the stability of the CPSF/RNA complex over that of *$\Delta use/CPS$* (Fig. 3C, cf. ΔAB , lanes 31–35, and Fig. 3A, *$\Delta use/CPS$* , lanes 11–15). In contrast, replacement of HIV-1 sequences 3' to the AAUAAA hexamer (segment C) did not increase complex stability relative to *$\Delta use/CPS$* (cf. *$\Delta use/\Delta C$* , Fig. 3B, lanes 26–30, and *$\Delta use/CPS$* , Fig. 3A, lanes 11–15).

The comparison of *use/CPS* (Fig. 3A, lanes 6–10) and *$\Delta use/CPS$* (Fig. 3A, lanes 11–15) indicates that contacts

Figure 2. Reconstitution of polyadenylation and processing complex formation with purified CPSF. (A) Silver-stained 10% polyacrylamide-SDS gels of HeLa cell glycerol gradient-purified CPSF (lane 1) and CstF (lane 2). (B) Polyadenylation of precleaved HIV-1 pre-mRNA substrates in the presence of glycerol gradient-purified CPSF and recombinant calf thymus poly(A) polymerase. (Lanes 1,3,5) Wild-type HIV-1 pre-mRNA substrate (use/CPS); (lanes 2,4,6) the HIV-1 pre-mRNA substrate with the single point mutation (U → C) in the hexamer (use/CPS/AACAAA). (C) Gel mobility shift analysis of glycerol gradient-purified CPSF and CstF. Full-length pre-mRNAs (containing both the hexamer and downstream element): (use/CPS) Wild-type HIV-1 pre-mRNA; (use/CPS/AACAAA) HIV-1 pre-mRNA with a single point mutation (U → C) in the hexamer; (Δuse/CPS) HIV-1 pre-mRNA containing the upstream element substitution mutation [NXS 9571–9588; Valsamakis et al. 1991].

within the HIV-1 upstream element serve to stabilize the binding of CPSF to the pre-mRNA. In addition, L3 sequences 5' to the AAUAAA hexamer that were shown to be functionally equivalent to HIV-1 upstream sequences by cleavage analysis (Fig. 1A,B) were also found to stabilize the CPSF/RNA complex (cf. ΔAB, Fig. 3C, lanes 31–35, and Δuse/CPS, Fig. 3A, lanes 11–15). Interestingly, the C segment, which significantly increased cleavage efficiency (Fig. 1A,B), did not enhance the stability of the CPSF/RNA complex (see Discussion). In two cases, the L3 C segment had a negative impact on CPSF binding (cf. ΔA, Fig. 3B, lanes 16–20, and ΔAC, Fig. 3C, lanes 36–40, and cf. Δuse/ΔB, Fig. 3B, lanes 21–25 and Δuse/ΔBC Fig. 3C, lanes 41–45). However, in the presence of a fully intact L3 upstream sequence the C segment had a more modest impact (cf. ΔAB, Fig. 3C, lanes 31–35, and L3, Fig. 3A, lanes 1–5).

The stability of the CPSF/RNA complex is reflected in the efficiency of polyadenylation

To address the functional significance of the stability of the CPSF/RNA complex, we examined the polyadenylation efficiency of the precleaved chimeric RNAs. Specific polyadenylation of a precleaved RNA requires only two factors, CPSF and poly(A) polymerase (for review, see Wahle and Keller 1992). CPSF recruits poly(A) polymerase to the pre-mRNA, dependent on the presence of the AAUAAA hexamer (Murthy and Manley 1992; Bienroth et al. 1993). On its own, poly(A) polymerase lacks a specific polyadenylation activity (see Fig. 2B, lanes 3 and 4). An increase in the stability of the CPSF/RNA complex, mediated by sequences upstream of the AAUAAA hexamer, should therefore be reflected in an increase in polyadenylation efficiency.

CPSF recognition of a 3'-processing enhancer

Figure 3. Sequences 5' to the AAUAAA hexamer contribute to the stability of the CPSF/RNA complex. 32 P-labeled pre-cleaved RNAs (schematically illustrated in Fig. 1) were incubated with glycerol gradient-purified CPSF for 10 min at 30°C. An aliquot was removed, denoted time 0, and treated with heparin at 0°C. An ~100-fold molar excess of unlabeled *use*/CPS pre-mRNA [containing the wild-type HIV-1 poly(A) site] was added to the remaining reaction and aliquots were removed after an additional incubation for 1, 2.5, 5, and 10 min at 30°C. Each aliquot was treated with heparin at 0°C. The RNA-protein complexes were then resolved on a nondenaturing 3% polyacrylamide gel.

Each of the RNAs analyzed in Figure 3 was incubated with recombinant calf thymus poly(A) polymerase and a limiting amount of CPSF (as judged by gel mobility shift; data not shown) in the presence of Mg^{2+} and ATP. Aliquots were taken at various times, and the extent of poly(A) addition was analyzed by denaturing gel electrophoresis (Fig. 4). A comparison of the data in Figures 3 and 4 indicates that the stability of the CPSF/RNA complex is reflected in the efficiency of polyadenylation. In particular, the Δuse /CPS and Δuse /ΔC RNAs that exhibited a weak interaction with CPSF were polyadenylated inefficiently (Fig. 4, Δuse /CPS, lanes 5–8, and Δuse /ΔC, lanes 13–16). Extensive polyadenylation of the Δuse /CPS and Δuse /ΔC RNAs was observed only at high concentrations of CPSF (data not shown).

Taken together, the data in Figures 3 and 4 indicate that sequences upstream of the AAUAAA hexamer that stabilize the binding of CPSF to the pre-mRNA enhance the efficiency of polyadenylation in vitro.

The 160-kD subunit of CPSF interacts with the HIV-1 upstream element

UV cross-linking was used to directly examine the interaction of CPSF with the HIV-1 upstream element. Uni-

formly [α - 32 P]UTP-labeled full-length HIV-1 pre-mRNA substrates were incubated with partially purified CPSF [poly(A)-Sepharose fraction; see Materials and methods]. Following UV-cross-linking and RNase A digestion, the labeled proteins were analyzed by SDS-PAGE. A protein of 160 kD was found to be cross-linked in an AAUAAA-dependent manner (Fig. 5A, lanes 3,4). The elimination of the 160-kD UV cross-linked product upon the introduction of a single point mutation within the hexamer (AAUAAA $\rightarrow$ AACAAA) suggests that the cross-linked product is the 160-kD component of CPSF (Bienroth et al. 1991; Murthy and Manley 1992). The single hexamer mutation has been shown previously to eliminate processing both in vivo and in vitro (for review, see Wahle and Keller 1992), and as illustrated in Figure 2C (lanes 1 and 2), eliminated CPSF binding to the pre-mRNA. Furthermore, the AAUAAA-dependent UV cross-linking of a ~160-kD product has been observed previously in CPSF-containing fractions (Moore et al. 1988; Gilmartin and Nevins 1989; Keller et al. 1991). The formation of cross-linked products of 70 and 30 kD was also found to be enhanced on the wild-type RNA. The 70-kD band can be distinguished from the cross-linking of the 64-kD polypeptide of CstF (Takagaki et al. 1990) present in the

Figure 4. The stability of the CPSF/RNA complex is reflected in the efficiency of polyadenylation. The ^{32}P -labeled precleaved RNAs analyzed in Fig. 3 were incubated with recombinant calf thymus poly(A) polymerase and a limiting amount of glycerol gradient-purified CPSF under polyadenylation conditions. Aliquots of the reaction were removed at 0, 2.5, 5, and 10 min, and the RNA was analyzed on a denaturing 5% polyacrylamide gel.

adjacent lanes (Fig. 5A, lanes 1,2). These cross-linked products may be derived from the 73-kD subunit of CPSF and the 30-kD protein observed by Bienroth et al. (1991), as they both respond to the point mutation in the hexamer; however, we cannot rule out the possibility that they may be derived from contaminating RNA-binding proteins or may represent degradation products of the 160-kD polypeptide.

To identify potential UV cross-links within the HIV-1 upstream element, we adapted an approach developed by Moore and Sharp (1992) to exclusively label upstream sequences. An RNA containing only upstream sequences (sequences -73 to -145 upstream of the AAUAAA hexamer, encompassing the upstream element) was produced by transcription in the presence of $[\alpha\text{-}^{32}\text{P}]\text{UTP}$. This ^{32}P -labeled RNA was then ligated to an unlabeled RNA containing the core poly(A) site (sequences -72 to $+56$) to yield an RNA that was identical in sequence to the full-length *use/CPS* or *use/CPS/AA-CAAA* RNAs used in Figure 5A. The ^{32}P label within the ligated RNAs was therefore restricted to sequences upstream of position -73 (relative to the AAUAAA hexamer). Following gel purification, the ligated RNAs were incubated with CPSF [poly(A)-Sepharose fraction] and subjected to UV irradiation. As in Figure 5A, the complexes were then extensively digested with RNase A and electrophoresed on an SDS-polyacrylamide gel (Fig. 5B). The 160-kD protein cross-linked only to the RNA containing the wild-type AAUAAA hexamer. The UV cross-

linking of the 70-kD and 30-kD polypeptides was also strongly enhanced on the wild-type versus the mutant RNA. (The gel in Fig. 5B was electrophoresed slightly longer than the gel in Fig. 5A). The experiment was repeated using glycerol gradient-purified CPSF (Fig. 2A, lane 1), and again, a 160-kD polypeptide was observed that was specifically UV cross-linked in an AAUAAA-dependent manner (Fig. 5C). In the glycerol gradient-purified fraction, the 30-kD cross-linked product could not be detected, and the 70-kD product, though still present, was markedly reduced relative to the 160-kD product. As yet, we are unable to determine whether the 70-kD product represents the 73-kD subunit of CPSF, a contaminating RNA-binding protein, or whether it is a degradation product of the 160-kD polypeptide. These results indicate that the 160-kD subunit of CPSF interacts directly with the HIV-1 upstream element in the context of the AAUAAA hexamer.

Discussion

An increasing number of *cis*-acting elements distinct from the AAUAAA hexamer and downstream element that constitute the core poly(A) site have been found to modulate poly(A) site efficiency. In this report we demonstrate that an element upstream of the HIV-1 core poly(A) site interacts directly with the 160-kD subunit of CPSF, the factor responsible for the recognition of the AAUAAA hexamer. The presence of the upstream ele-

Figure 5. The 160-kD subunit of CPSF interacts directly with the upstream element of the HIV-1 poly(A) site. Full-length HIV-1 pre-mRNAs were incubated with CPSF, and the reactions were treated with heparin at 0°C, irradiated with UV light, and digested with RNase A, as described in Materials and methods. The resulting labeled proteins were then analyzed on a 10% polyacrylamide-SDS gel. (Lanes 1,3,5,7) The wild-type HIV-1 pre-mRNA substrate (*use/CPS*); (lanes 2,4,6,8) the HIV-1 pre-mRNA substrate with the single point mutation (U → C) in the hexamer (*use/CPS/AACAAA*). (A) UV cross-linking of partially purified CPSF [poly(A)-Sephacrose fraction, see Materials and methods] to uniformly [α - 32 P]UTP-labeled HIV-1 pre-mRNAs. (B) UV cross-linking of partially purified CPSF [poly(A)-Sephacrose fraction] to RNAs specifically 32 P-labeled within the upstream sequence (sequences -73 to -145 upstream of the AAUAAA hexamer; see Results). (C) UV cross-linking of glycerol gradient-purified CPSF to RNAs specifically labeled within the upstream sequence.

ment in the context of the AAUAAA hexamer directs the stable binding of CPSF to the pre-mRNA and enhances the efficiency of poly(A) addition in reactions reconstituted with purified CPSF and recombinant poly(A) polymerase. We also show that the HIV-1 upstream element is necessitated by the suboptimal sequence context of the AAUAAA hexamer. The replacement of sequences adjacent to the hexamer with the analogous sequences from the efficient adenovirus L3 poly(A) site was found to eliminate the requirement for an upstream element. Taken together, our results suggest that poly(A) site selection in HIV-1 may be a consequence of the inability of the processing machinery to recognize the inefficient core poly(A) site at the 5' end of the message, coupled with the enhancement of CPSF binding by an upstream element unique to the 3' end of the message. Promoter occlusion of the 5' core poly(A) site, which might also contribute to poly(A) site selection in HIV-1 (Weichs an der Glon et al. 1991; Cherrington and Ganem 1992), may also be facilitated by the inefficient core poly(A) site.

Although the HIV-1 upstream element may serve to enhance processing by stabilizing the interaction of CPSF with the pre-mRNA, our results clearly indicate that processing complex stability is not the sole determinant of processing efficiency. The replacement of HIV-1 sequences between the hexamer and the cleavage site with those of L3 increased the efficiency of cleavage

but did not enhance the efficiency of poly(A) addition or the stability of the CPSF/RNA complex. The impact of sequences 3' to the hexamer may be indicative of the requirements for endonuclease recognition or access to the cleavage site. The upstream element may serve to complement the inability of these sequences to efficiently promote cleavage at the HIV-1 poly(A) site through its interaction with CPSF. These results suggest that a mammalian poly(A) site is not defined by the sum of a modular set of *cis*-acting elements but by a set of interdependent sequences in the context of the AAUAAA hexamer that act in concert to direct 3' processing.

The interaction of CPSF with the AAUAAA hexamer has been analyzed in considerable detail (for review, see Wahle and Keller 1992). Purification of CPSF to near homogeneity has revealed that it is composed of at least three polypeptides of 160, 105, and 73 kD (Bienroth et al. 1991; Murthy and Manley 1992). A fourth polypeptide of 30 kD may also be a subunit of CPSF (Bienroth et al. 1991). Mutagenesis and RNA modification-interference experiments have indicated that the nucleotides essential for CPSF binding are confined to the AAUAAA hexamer (Barwell et al. 1991; Keller et al. 1991). RNAs as short as 10 nucleotides formed specific CPSF/RNA complexes (Wigley et al. 1990), and polypeptides of 160 and 30 kD were shown to be UV cross-linked to RNAs as short as 18 nucleotides in an AAUAAA-dependent man-

ner (Keller et al. 1991). Our work suggests that the recognition of both the AAUAAA hexamer and upstream sequences may enable CPSF to identify a hexamer within an authentic poly(A) site, to the exclusion of those hexamers that reside elsewhere in the message. [AAUAAA is not an unfavorable choice for protein-coding sequence and is present within the coding region of ~16% of all mRNAs (Day 1992)]. CPSF contacts within the upstream sequences may be of particular importance in the recognition of rare processing sites that possess suboptimal variants of the AAUAAA hexamer (see Wahle and Keller 1992).

If the recognition of upstream sequences is part of a general strategy for poly(A) site definition, a sequence bias in this region might be expected. In a survey of 216 poly(A) sites, Nussinov (1986) observed a bias for pyrimidine-rich sequences, particularly C₄₋₅ and U₄₋₅ homopolymers, extending ~50 nucleotides upstream of the AAUAAA hexamer. Such a sequence bias is consistent with the general U-rich nature of upstream elements identified in SV40 (Schek et al. 1992), HBV (Rusnak and Ganem 1990), adenovirus (Prescott and Falck-Pedersen 1992, 1994), and HIV-1 (Valsamakis et al. 1991). The HIV-1 upstream element that we have examined in vitro contains a run of 5 uridine residues in a pyrimidine-rich (14/18) sequence. A general role for upstream sequences in polyadenylation is supported by the observation that the sequences 5' to the HIV-1 AAUAAA hexamer are functionally equivalent to those 5' to the adenovirus L3 hexamer. The mechanism by which the HIV-1 upstream element enhances 3' processing therefore does not appear to be unique but may be a component of the general mechanism of poly(A) site definition.

Recent work by Bilger et al. (1994), suggests that the U-rich upstream elements that serve to direct the developmentally regulated addition of poly(A) to mRNAs in the cytoplasm of oocytes and early embryos [CPEs (cytoplasmic polyadenylation elements)] may be functionally equivalent to upstream sequences involved in nuclear 3' processing. In parallel with our observations in HIV-1, Bilger et al. (1994) argue that CPEs enhance cytoplasmic polyadenylation in an AAUAAA-dependent manner by increasing the affinity of the mRNA for CPSF. CPSF recognition of upstream sequences in the context of the AAUAAA hexamer may therefore be exploited to direct both nuclear and cytoplasmic polyadenylation.

The expansion of the mammalian poly(A) site to encompass a set of heterogeneous sequence elements that extend 5' to the AAUAAA hexamer reveals a 3' processing signal with a greater degree of similarity to those of *Saccharomyces cerevisiae* and plants (Wahle and Keller 1992). In the absence of a simple highly conserved element equivalent to the AAUAAA hexamer, 3' processing in yeast and plants is apparently directed by a heterogeneous set of sequences that may extend a considerable distance upstream of the cleavage site. Extensive analysis of the signals that direct 3'-end formation of the yeast *CYC1* mRNA (Russo et al. 1993) revealed an arrangement of processing signals strikingly similar to

those of HIV-1. An element that appears to be functionally equivalent to the AAUAAA hexamer in positioning the site of cleavage downstream was found to function together with an upstream element that enhanced the efficiency of processing. A similar arrangement of processing signals has been proposed for plant poly(A) sites. A U-rich far upstream sequence (FUE) coupled with a near upstream element (NUE), which may be functionally equivalent to the AAUAAA hexamer, appear to be the primary determinants of poly(A) site recognition in plants (Hunt 1994). As in yeast, neither FUE nor the NUE sequence is strictly conserved. The importance of a diverse set of generally U-rich upstream sequences that act in concert with an element near the site of cleavage in the 3' processing of yeast, plant, and mammalian pre-mRNAs suggests that these eukaryotes may share a common mechanism of mRNA 3' end formation.

The recognition of multiple *cis*-acting elements by the basal polyadenylation machinery, as discussed in this report, is but one aspect of the definition of a poly(A) site. Clearly, the coupling of pre-mRNA splicing and polyadenylation may also play a central role in poly(A) site definition. Splicing may preclude the use of potential poly(A) sites within introns (Adami and Nevins 1988; Levitt et al. 1989), suggesting that the realm of poly(A) sites that are accessible to the 3' processing machinery may initially be defined by spliceosome assembly on the nascent pre-mRNA. Splicing also appears to enhance the use of poly(A) sites within 3' exons (Niwa et al. 1990; Nesic et al. 1993; Nesic and Maquat 1994). Recent evidence presented by Lutz and Alwine (1994) suggest that the U1 snRNP-A protein may serve to couple polyadenylation and splicing. The multiple sequence contacts between the polyadenylation machinery and the pre-mRNA coupled with the constraints imposed by splicing may therefore begin to explain how such a diverse array of sequences may serve to direct mRNA 3' processing.

Materials and methods

Preparation of RNA substrates

Pre-mRNA substrates were prepared by in vitro transcription with T7 or SP6 RNA polymerase (Epicenter Technologies) in the presence of 15 μ M GTP and 0.4 μ M [α -³²P]GTP and gel purified prior to use. The structure and synthesis of the CPS, *use*/CPS, Δ *use*/CPS, *use*/CPS/AACAAA, CPS/AACAAA, and L3 pre-mRNA substrates has been described previously (Gilmartin et al. 1992). For the construction of the chimeric HIV/L3 transcription templates, the HIV-1 poly(A) site (*puse*/CPS; Gilmartin et al. 1992) was divided into four segments (A–D) as illustrated in Figure 1 and described in Results. Each segment was then replaced, individually or in combination, with the corresponding sequence from the adenovirus L3 poly(A) site (McDevitt et al. 1988). Oligonucleotides containing the appropriate sequences were utilized along with unique restriction sites to generate exact HIV/L3 junctions. The C/D segment junctions were formed by digestion with *Xho*I (HIV-1) or *Acc*I (L3), which encompasses the cleavage sites of the HIV-1 (see Gilmartin et al. 1992) and L3 (see McDevitt et al. 1988) poly(A) sites, respectively, followed by treatment with Klenow polymerase and DNA ligase.

The DNAs that served as templates for the *in vitro* transcription of the precleaved pre-mRNAs *use*/CPS, ΔA , and ΔAB were digested with *XhoI* (Gilmartin et al. 1992). The DNAs that served as templates for the *in vitro* transcription of the precleaved pre-mRNAs L3, and ΔAC were digested with *AccI* (McDevitt et al. 1988). The analogous templates for Δuse /CPS, $\Delta use/\Delta B$, $\Delta use/\Delta C$, and $\Delta use/\Delta BC$ were produced by PCR, because of the presence of both an *XhoI* and an *AccI* site within the A segment of the Δuse linker substitution (NXS 9571–9588; Valsamakis et al. 1991).

The ligated RNAs used for UV cross-linking (Fig. 5B,C) were generated by a procedure developed by Moore and Sharp (1992). A ^{32}P -labeled RNA containing the HIV-1 upstream element (–73 to –145 upstream of the AAUAAA hexamer) was synthesized by SP6 RNA polymerase in the presence of [α - ^{32}P]UTP. This ^{32}P -labeled RNA was then ligated to an unlabeled, GMP-primed transcript of the pCPS or pCPS/AACAAA templates linearized with *EcoRI* [containing the core poly(A) site, sequences –72 to +56] (Gilmartin et al. 1992). The two RNAs to be ligated were annealed to a 66-nucleotide splint DNA oligonucleotide complementary to the junction of the two RNAs and a 50-nucleotide disruptor DNA oligonucleotide complementary to the sequences of the TAR stem-loop. The annealing reaction contained 1 mM EDTA, 10 mM Tris (pH 7.6), 1 mM spermidine, 2 pmoles of the [α - ^{32}P]UTP-labeled upstream sequence RNA, 20 pmoles of unlabeled CPS or CPS/AACAAA RNA, 20 pmoles of disruptor oligonucleotide, and 20 pmoles of splint oligonucleotide in a final volume of 5 μ l. The samples were incubated for 1 min at 95°C and slow cooled to ~30°C. After annealing, the reaction was brought to 0.5 mM EDTA, 71 mM Tris-HCl (pH 7.6), 6.6 mM MgCl₂, 10 mM DTT, 0.5 mM ATP, and 2.5 U/ μ l of T4 DNA ligase (U.S. Biochemical) in a final volume of 10 μ l. The reactions were incubated for 8 hr at room temperature, and the ligated products were gel purified. The ligated products were colinear with the full-length *use*/CPS and *use*/CPS/AACAAA transcripts, respectively.

Extract preparation and fractionation

HeLa cells were grown in suspension in Joklik's modified minimal essential medium containing 5% calf serum. Nuclear extracts were prepared by the method of Dignam et al. (1983) except that the nuclei were extracted with 350 mM KCl and all buffers contained 0.1 mM phenylmethylsulfonyl fluoride (PMSF). The extracts were dialyzed to 75 mM KCl, 20 mM HEPES (pH 7.9), 10% glycerol, 0.2 mM EDTA, 0.5 mM DTT, and 0.1 mM PMSF. Purification of CPSF was carried out essentially as described by Bienroth et al. (1991), followed by glycerol gradient centrifugation at 36,000 rpm at 4°C for 39 hr in an SW41 rotor as described by Murthy and Manley (1992). Purification of CstF through the glycerol gradient centrifugation step was carried out as described by Gilmartin and Nevins (1991). The hexahistidine-tagged calf thymus poly(A) polymerase cDNA clone was a gift of E. Wahle and W. Keller (Biozentrum, Basel, Switzerland). The hexahistidine-tagged calf thymus poly(A) polymerase was expressed in *Escherichia coli* BL21(DE3)pLysS and purified by Ni²⁺–NTA (Qiagen) chromatography.

In vitro 3' processing

Cleavage reactions were carried out in DEAE–Sephacrose-fractionated HeLa cell nuclear extract (Gilmartin and Nevins 1989). Twenty-nine fmoles of ^{32}P -labeled pre-mRNA was incubated in a total reaction volume of 25 μ l containing 8.5 μ l of DE150 fraction (2.9 mg/ml), 8.5 μ l of DE250 (5.0 mg/ml), 54 mM KCl, 14.4 mM HEPES (pH 7.9), 7.2% glycerol, 0.5 mM ATP, 2% poly-

vinylalcohol, 0.14 mM EDTA, 0.5 μ g of tRNA, 0.07 mM PMSF, and 0.36 mM DTT for 60 min at 30°C. The reactions were stopped by the addition of SDS to 0.5%, treated with proteinase K, phenolized, and ethanol precipitated, and the products were electrophoresed on a denaturing 10% polyacrylamide gel.

Polyadenylation reactions were carried out essentially as described by McDevitt et al. (1988). The reactions contained 23 fmoles of precleaved ^{32}P -labeled pre-mRNA incubated in a total reaction volume of 25 μ l containing 4 ng of glycerol gradient-purified CPSF, 25 ng recombinant poly(A) polymerase, 100 mM KCl, 12.4 mM HEPES (pH 7.9), 0.01% NP-40, 6.2% glycerol, 0.12 mM EDTA, 0.06 mM PMSF, 0.31 mM DTT, 2% polyvinyl alcohol, 1 mM ATP, 5 mM creatine phosphate, 0.7 mM MgCl₂, and 0.5 μ g of tRNA. The amount of CPSF in each reaction was limiting as judged by gel mobility shift (data not shown). The reactions depicted in Figure 4 were incubated at 30°C and 5- μ l aliquots were taken at 0, 2.5, 5, and 10 min. The reactions depicted in Figure 2B were performed in essentially the same manner except that they were incubated for 30 min at 30°C. The reactions were stopped with the addition of SDS to 0.5%, treated with proteinase K, and the products were electrophoresed on a denaturing 8% polyacrylamide gel.

Complex formation

Complex formation assays were carried out as described previously (Gilmartin et al. 1992). The complex formation assays depicted in Figure 2C were carried out in 25- μ l reactions containing 83 mM KCl, 15 fmoles of ^{32}P -labeled precleaved pre-mRNA, 0.4 μ g of tRNA, 1% polyvinyl alcohol, 0.016% NP-40, 16.5 mM HEPES (pH 7.9), 8.3% glycerol, 0.17 mM EDTA, 0.08 mM PMSF, and 0.41 mM DTT. Each reaction contained either 10 ng of glycerol gradient-purified CPSF or CstF, or 5 ng of both CPSF and CstF. The complex stability assays depicted in Figure 3 were carried out in 30- μ l reactions containing 15 ng of glycerol gradient-purified CPSF, 83 mM KCl, 23 fmoles of ^{32}P -labeled precleaved pre-mRNA, 0.6 μ g of tRNA, 1% polyvinyl alcohol, 0.016% NP-40, 16.5 mM HEPES (pH 7.9), 8.3% glycerol, 0.17 mM EDTA, 0.08 mM PMSF, and 0.41 mM DTT. The reactions were incubated for 5 min at 30°C, and a 5- μ l aliquot was taken of each sample for the zero time point. To the remaining reaction, ~3 pmoles (in 1 μ l) of unlabeled *use*/CPS pre-mRNA was added. The incubation was continued at 30°C and 5.2- μ l aliquots were taken of each sample at 1, 2.5, 5, and 10 min, placed on ice, and treated with 5 mg/ml of heparin. The reactions were electrophoresed on a nondenaturing 3% polyacrylamide (100:1) gel in 25 mM Tris, 25 mM boric acid, and 1 mM EDTA (pH 8.0) at 300 V for 1.5 hr at 4°C. The gel was pre-electrophoresed for 30 min prior to loading.

UV cross-linking

UV cross-linking reactions were performed in a total volume of 20 μ l containing either 150 ng of a partially purified CPSF fraction [poly(A)–Sephacrose fraction, Bienroth et al. 1991] or 15 ng of glycerol gradient-purified CPSF, 25 fmoles of ^{32}P -labeled precleaved pre-mRNA, 83 mM KCl, 0.5 μ g of tRNA, 0.016% NP-40, 16 mM HEPES (pH 7.9), 8% glycerol, 0.7 mM MgCl₂, 1 mM ATP, 0.16 mM EDTA, 0.08 mM PMSF, and 0.41 mM DTT. After incubation for 10 min at 30°C, the reactions were placed on ice, heparin was added to 5 mg/ml, and the reactions were irradiated with 254 nm light (UVP model UVG-54) at a distance of 5 cm for 10 min. RNase A was then added to 50 μ g/ml, and the reaction was incubated for 15 min at 37°C. Protein gel loading buffer was added, and the samples were heated to 100°C for 2 min and electrophoresed on a 10% polyacrylamide–SDS gel.

Acknowledgments

We thank Sean Flaherty and Richard Fishel for many stimulating discussions and critical reading of the manuscript. This work was supported by grants from The Lucille P. Markey Charitable Trust [NSF-VT EPSCoR (RII-8610679)] and from the National Institutes of Health (GM46624). G.M.G. was supported by an American Cancer Society Junior Faculty Award (JRFA-367).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Adami, G. and J.R. Nevins. 1988. Splice site selection dominates over poly(A) site choice in RNA production from complex adenovirus transcription units. *EMBO J.* **7**: 2107–2116.
- Bardwell, V.J., M. Wickens, S. Bienroth, W. Keller, B. Sproat, and A.I. Lamond. 1991. Site-directed ribose methylation identifies 2'-OH groups in polyadenylation substrates critical for AAUAAA recognition and poly(A) addition. *Cell* **65**: 125–133.
- Bienroth, S., E. Wahle, C. Suter-Crazzolara, and W. Keller. 1991. Purification of the cleavage and polyadenylation factor involved in the 3' processing of messenger RNA precursors. *J. Biol. Chem.* **266**: 19768–19776.
- Bienroth, S., W. Keller, and E. Wahle. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* **12**: 585–594.
- Biliger, A., C.A. Fox, E. Wahle, and M. Wickens. 1994. Nuclear polyadenylation factors recognize cytoplasmic polyadenylation elements. *Genes & Dev.* **8**: 1106–1116.
- Brown, P.H., L.S. Tiley, and B.R. Cullen. 1991. Efficient polyadenylation within the human immunodeficiency virus type 1 long terminal repeat requires flanking U3-specific sequences. *J. Virol.* **65**: 3340–3343.
- Carswell, S. and J.C. Alwine. 1989. Efficiency of utilization of the simian virus 40 late polyadenylation site: Effects of upstream sequences. *Mol. Cell. Biol.* **9**: 4248–4258.
- Cherrington, J. and D. Ganem. 1992. Regulation of polyadenylation in human immunodeficiency virus (HIV): Contributions of promoter proximity and upstream sequences. *EMBO J.* **11**: 1513–1524.
- Cherrington, J., R. Russnak, and D. Ganem. 1992a. Upstream sequences and cap proximity in the regulation of polyadenylation in ground squirrel hepatitis virus. *J. Virol.* **66**: 7589–7596.
- Christofori, G. and W. Keller. 1989. Poly(A) polymerase purified from HeLa cell nuclear extract is required for both cleavage and polyadenylation of pre-mRNA *in vitro*. *Mol. Cell. Biol.* **9**: 193–203.
- Coffin, J.M. and C. Moore. 1990. Determination of 3' end processing in retroelements. *Trends Genet.* **6**: 267–277.
- Day, I.N.M. 1992. Analysis of the 5'-AAUAAA motif and its flanking sequence in human RNA: Relevance to cDNA library sorting. *Gene* **110**: 245–249.
- DeZazzo, J.D. and M.J. Imperiale. 1989. Sequences upstream of AAUAAA influence poly(A) site selection in a complex transcription unit. *Mol. Cell. Biol.* **9**: 4951–4961.
- DeZazzo, J.D., E. Falck-Pedersen, and M.J. Imperiale. 1991a. Sequences regulating temporal poly(A) site switching in the adenovirus major late transcription unit. *Mol. Cell. Biol.* **11**: 5977–5984.
- DeZazzo, J.D., J.E. Kilpatrick, and M.J. Imperiale. 1991b. Involvement of LTR U3 sequences overlapping the transcription control region in human immunodeficiency virus type 1 mRNA 3' end formation. *Mol. Cell. Biol.* **11**: 1624–1630.
- 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.
- Gilmartin, G.M. and J.R. Nevins. 1989. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. *Genes & Dev.* **3**: 2180–2189.
- . 1991. Molecular analyses of two poly(A) site processing factors that determine the recognition and efficiency of cleavage of the pre-mRNA. *Mol. Cell. Biol.* **11**: 2432–2438.
- Gilmartin, G.M., E.S. Fleming, and J. Oetjen. 1992. Activation of HIV-1 3' processing *in vitro* requires both an upstream element and TAR. *EMBO J.* **11**: 4419–4428.
- Guise, J.W., G. Galli, J.R. Nevins, and P.W. Tucker. 1989. Developmental regulation of secreted and membrane forms of immunoglobulin mu chain. In *Immunoglobulin genes* (ed. T. Honjo et al.), pp. 275–301. Academic Press, New York.
- Hunt, A.G. 1994. Messenger RNA 3' end formation in plants. *Annu. Rev. Plant Physiol. Mol. Biol.* **45**: 47–60.
- Keller, W., S. Bienroth, K.M. Lang, and G. Christofori. 1991. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J.* **10**: 4241–4249.
- Levitt, N., D. Briggs, A. Gil, and N.J. Proudfoot. 1989. Definition of an efficient synthetic poly(A) site. *Genes & Dev.* **3**: 1019–1025.
- Lutz, C.S. and J. C. Alwine. 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.
- MacDonald, C.C., J. Wilusz, and T. Shenk. 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.
- Mann, K.P., E.A. Weiss, and J.R. Nevins. 1993. Alternative poly(A) site utilization during adenovirus infection coincides with a decrease in the activity of a poly(A) site processing factor. *Mol. Cell. Biol.* **13**: 2411–2419.
- McDevitt, M.A., G.M. Gilmartin, W.H. Reeves, and J.R. Nevins. 1988. Multiple factors are required for poly(A) addition to a mRNA 3' end. *Genes & Dev.* **2**: 588–597.
- Moore, C.L., J. Chen, and J. Whoriskey. 1988. Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding. *EMBO J.* **7**: 3159–3169.
- Moore, M.J. and P.A. Sharp. 1992. Site-specific modification of pre-mRNA: The 2'-hydroxyl groups at the splice sites. *Science* **256**: 992–997.
- Murthy, K.G. and J.L. Manley. 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.* **267**: 14804–14811.
- Nesic, D. and L.E. Maquat. 1994. Upstream introns influence the efficiency of final intron removal and RNA 3'-end formation. *Genes & Dev.* **8**: 363–375.
- Nesic, D., J. Cheng, and L.E. Maquat. 1993. Sequences within the last intron function in RNA 3'-end formation in cultured cells. *Mol. Cell. Biol.* **13**: 3359–3369.
- 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.
- Nussinov, R. 1986. Sequence signals which may be required for efficient formation of mRNA 3' termini. *Nucleic Acids Res.* **14**: 3557–3571.
- Prescott, J.C. and E. Falck-Pedersen. 1992. Varied poly(A) site

- efficiency in the adenovirus major late transcription unit. *J. Biol. Chem.* **267**: 8175–8181.
- . Sequence elements upstream of the 3' cleavage site confer substrate strength to the adenovirus L1 and L3 polyadenylation sites. *Mol. Cell. Biol.* **14**: 4682–4693.
- Russnak, R. 1991. Regulation of polyadenylation in hepatitis B viruses: Stimulation by the upstream activating signal PS1 is orientation-dependent, distance-dependent, and additive. *Nucleic Acids Res.* **19**: 6449–6456.
- Russnak, R. and D. Ganem. 1990. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes & Dev.* **4**: 764–776.
- Russo, P., W.-Z. Li, Z. Guo, and F. Sherman. 1993. Signals that produce 3' termini in CYC1 mRNA of the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13**: 7836–7849.
- Schek, N., C. Cooke, and J. Alwine. 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.
- Sittler, A., H. Gallinaro, and M. Jacob. 1994. Upstream and downstream cis-acting elements for cleavage at the L4 polyadenylation site of adenovirus-2. *Nucleic Acids Res.* **22**: 222–231.
- Takagaki, Y., L.C. Ryner, and J.L. Manley. 1989. Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes & Dev.* **3**: 1711–1724.
- Takagaki, Y., J.L. Manley, C.C. MacDonald, J. Wilusz, and T. Shenk. 1990. A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. *Genes & Dev.* **4**: 2112–2120.
- Valsamakis, A., S. Zeichner, S. Carswell, and J.C. Alwine. 1991. The human immunodeficiency virus type 1 polyadenylation signal: A 3' long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. *Proc. Natl. Acad. Sci.* **88**: 2108–2112.
- Valsamakis, A., N. Schek, and J.C. Alwine. 1992. Elements upstream of the AAUAAA within the human immunodeficiency virus polyadenylation signal are required for efficient polyadenylation *in vitro*. *Mol. Cell. Biol.* **12**: 3699–3705.
- Wahle, E. and W. Keller. 1992. The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. *Annu. Rev. Biochem.* **61**: 419–440.
- Wahle, E., G. Martin, E. Schiltz, and W. Keller. 1991. Isolation and expression of cDNA clones encoding mammalian poly(A) polymerase. *EMBO J.* **10**: 4251–4257.
- Weichs an der Glon, C., J. Monks, and N.J. Proudfoot. 1991. Occlusion of the HIV poly(A) site. *Genes & Dev.* **5**: 244–253.
- Weiss, E.A., G.M. Gilmartin, and J.R. Nevins. 1991. Poly(A) site efficiency reflects the stability of complex formation involving the downstream element. *EMBO J.* **10**: 215–219.
- Wigley, P.L., M.D. Sheets, D.A. Zarkower, M.E. Whitmer, and M. Wickens. 1990. Polyadenylation of mRNA: Minimal substrates and a requirement for the 2' hydroxyl of the U in AAUAAA. *Mol. Cell. Biol.* **10**: 1705–1713.

CPSF recognition of an HIV-1 mRNA 3'-processing enhancer: multiple sequence contacts involved in poly(A) site definition.

G M Gilmartin, E S Fleming, J Oetjen, et al.

Genes Dev. 1995, **9**:

Access the most recent version at doi:[10.1101/gad.9.1.72](https://doi.org/10.1101/gad.9.1.72)

References

This article cites 48 articles, 31 of which can be accessed free at:
<http://genesdev.cshlp.org/content/9/1/72.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).