MOLECULAR AND CELLULAR BIOLOGY, Apr. 1989, p. 1759–1771
0270-7306/89/041759-13$02.00/0

Vol. 9, No. 4

Sequences Downstream of AAUAAA Signals Affect Pre-mRNA Cleavage and Polyadenylation In Vitro Both Directly and Indirectly

LISA C. RYNER, YOSHIO TAKAGAKI, AND JAMES L. MANLEY*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 31 October 1988/Accepted 5 January 1989

To investigate the role of sequences lying downstream of the conserved AAUAAA hexanucleotide in pre-mRNA cleavage and polyadenylation, deletions or substitutions were constructed in polyadenylation signals from simian virus 40 and adenovirus, and their effects were assayed in both crude and fractionated HeLa cell nuclear extracts. As expected, these sequences influenced the efficiency of both cleavage and polyadenylation as well as the accuracy of the cleavage reaction. Sequences near or upstream of the actual site of poly(A) addition appeared to specify a unique cleavage site, since their deletion resulted, in some cases, in heterogeneous cleavage. Furthermore, the sequences that allowed the simian virus 40 late pre-RNA to be cleaved preferentially by partially purified cleavage activity were also those at the cleavage site itself. Interestingly, sequences downstream of the cleavage site interacted with factors not directly involved in catalyzing cleavage and polyadenylation, since the effects of deletions were substantially diminished when partially purified components were used in assays. In addition, these sequences contained elements that could affect 3′-end formation both positively and negatively.

Formation of the 3′ ends of polyadenylated mRNAs in higher eucaryotes is a complex process that appears to require multiple *cis*- and *trans*-acting factors (1, 15). Polyadenylation is minimally a two-step reaction in which a precursor RNA is first specifically cleaved, probably at the precise site of poly(A) addition (26, 35), and then polyadenylated by synthesis of a 200- to 300-nucleotide (nt) poly(A) tract. These two reactions appear to be coupled, since cleaved but not polyadenylated RNAs are normally not detected. However, the coupling is not obligatory, because cleavage and polyadenylation can each occur independently of the other under the appropriate conditions in vitro (14, 25).

The most important *cis*-acting sequence required for pre-mRNA cleavage-polyadenylation in vivo and in vitro is the highly conserved AAUAAA hexanucleotide, located 10 to 30 bases upstream of the poly(A) addition sites in the vast majority of mRNAs (7, 11, 23, 30, 45, 48). Interestingly, this sequence is required for both cleavage and polyadenylation when the two reactions are uncoupled and assayed independently of each other in vitro (16, 37, 48). Additional sequences lying downstream of the AAUAAA also play a role in the 3′-end formation reaction. Specifically, sequences located within about 50 bases 3′ to the RNA cleavage site have been shown to be required for efficient and accurate polyadenylation of a number of pre-mRNAs in vivo (3, 9, 10, 20, 33, 36, 47, 53) and in at least three cases in vitro (37, 51, 52). However, these sequences appear not to be as crucial as the AAUAAA hexanucleotide. For example, there is no precise, conserved sequence in this region. Required downstream elements appear to consist simply of U-rich or G+U-rich sequences, and a number of genes do not contain any recognizable downstream sequence motif (21, 41). In addition, although deletion and substitution mutations in these regions can have dramatic effects on the efficiency of 3′-end formation, they rarely inhibit accurate 3′-end formation completely.

The identities of the *trans*-acting factors that carry out the cleavage-polyadenylation reaction are not known. A number of studies have indicated that complexes form between factors in nuclear extracts and exogenously added pre-RNAs during in vitro incubations and that the formation of such complexes requires an intact AAUAAA sequence (27, 37, 38, 49, 50, 52). A downstream element may also be necessary for efficient complex formation (27, 51). We have shown recently that the components required to catalyze the cleavage-polyadenylation reaction can be separated into two fractions, one containing a nonspecific poly(A) polymerase and the other containing specific cleavage activity (40). Mixing of these fractions effectively reconstituted accurate and specific cleavage-polyadenylation. Interestingly, the partially purified cleavage enzyme was functional with only one of four pre-RNAs tested, although all four were utilized efficiently after addition of the poly(A) polymerase-containing fraction.

In this study, we used both crude and fractionated nuclear extracts to investigate the role of sequences that lie 3′ to the AAUAAA hexanucleotide on cleavage-polyadenylation of several pre-RNAs in vitro. Our results indicate that both the accuracy and the efficiency of cleavage and, to a lesser extent, the efficiency of polyadenylation can be influenced by such sequences. However, sequences that lie downstream of the cleavage site appear to interact primarily with factors that play an indirect role in 3′-end formation.

MATERIALS AND METHODS

Plasmid constructions. Plasmids pG3SVL-A and pG3E2-A have been described previously (40). pG3SVL+32 and pG3SVL+3 were constructed by excising *Bam*HI-to-*Bgl*II fragments from pB28L and pB35L (33), which contain simian virus 40 (SV40) sequences from nt 2533 to 2707 and 2533 to 2678, respectively, and inserting them into the *Bam*HI site of pGEM3 (Promega Biotech, Madison, Wis.). The G4 counterparts were constructed by inserting the same fragments into the *Bam*HI site of pGEM4. pG3SVL-7 was formed by excising the *Hin*dIII (pGEM3 nt 7)-to-*Hpa*I (SV40 nt 2668) fragment from pG3SVL+32 and inserting it into the *Hin*dIII and *Hin*cII site of pGEM3. pG4SVL-7 was similarly made by inserting an *Eco*RI-to-*Hpa*I (SV40 nt 2668) frag-

* Corresponding author.

ment from pG4SVL+32 into the *Eco*RI and *Hin*cII sites of pGEM4.

Plasmid pG4L3-A, in which the adenovirus type 2 (Ad2) L3 polyadenylation site was inserted into pGEM4, was described previously (40). To prepare 3′ deletion mutants, pG4L3-A DNA was digested with *Dra*I at a site 48 base pairs (bp) downstream of the poly(A) addition site and then treated with BAL 31 nuclease. After digestion with *Eco*RI at a site located in the 5′ polylinker sequence, *Eco*RI-*Bal* 31 fragments were isolated and inserted into pGEM4 digested with *Eco*RI and *Hin*cII. Relevant sequences of all mutants were determined by the method of Maxam and Gilbert (19). One of these mutants, pL3-18, in which the Ad2 L3-derived sequence ends 3 bp downstream of the AATAAA hexanucleotide, was used in construction of the chimeric genes described below.

To construct the plasmid containing the SV40 large-T-antigen-coding region (pLT), a 276-bp *Hin*dIII-*Pst*I fragment (SV40 nt 3480 to 3205; 43) containing an inactive AATAAA located at nt 3275 to 3270 was isolated and inserted into pGEM3 digested with *Hin*dIII and *Pst*I. To prepare 3′ deletion mutants, pLT was digested with *Pst*I and treated with *Bal* 31. After digestion with *Hin*dIII, *Hin*dIII-*Bal* 31 fragments were isolated and inserted into pGEM3 that had been digested with *Hin*dIII and *Sma*I.

For construction of chimeric genes, a 5′ deletion mutant of the Ad2 L3 poly(A) site was first generated. A 47-bp *Rsa*I (16 bp upstream of AATAAA)-to-*Ava*I [5 bp downstream of the poly(A) addition site] fragment and a 338-bp *Ava*I-to-*Sau*3AI [343 bp downstream of the poly(A) site] fragment were isolated from pG4L3-A and inserted into pGEM3 digested with *Hin*cII and *Bam*HI to create pRS. To delete the 5′ sequence of the Ad2 L3 poly(A) site, pRS was digested at the *Pst*I site in the polylinker sequence and treated with *Bal* 31. After digestion at the *Sau*3AI site at the 3′ end of Ad2 L3 sequence, *Bal* 31-*Sau*3AI fragments containing the poly(A) addition site as well as downstream sequences were isolated and inserted into pGEM3 digested with *Hin*cII and *Bam*HI. One of these mutants, pRS/14, in which Ad2 L3 sequences start 8 bp upstream of the poly(A) site, was used to construct chimeric genes. To construct the chimeric gene containing the AATAAA sequence of the SV40 large-T-antigen-coding region and the Ad2 L3 downstream sequences (pLT-L3), a 222-bp *Hin*dIII (205 bp upstream of the AATAAA sequence)-to-*Ssp*I (11 bp downstream of the AATAAA sequence) fragment isolated from pLT and a 355-bp *Pst*I (4 bp upstream of the 5′ end of Ad2 L3 sequences in the 5′ polylinker sequence, blunt ended with T4 DNA polymerase)-to-*Sau*3AI fragment isolated from pRS/14 were inserted into pGEM3 digested with *Hin*dIII and *Bam*HI.

To construct the chimeric gene containing the SV40 late AATAAA sequence and the poly(A) site of Ad2 L3 (pSVL-L3), a 165-bp *Hin*dIII (in the 5′ polylinker sequence)-to-*Hpa*I (6 bp downstream of AATAAA) fragment isolated from pG3SVL-A and the *Pst*I-*Sau*3AI fragment isolated from pRS/14 were inserted into pGEM3 digested with *Hin*dIII and *Bam*HI. To construct the chimeric gene containing the AATAAA sequence of Ad2 L3 and the SV40 late poly(A) site (pL3-SVL), a 209-bp *Eco*RI (in 5′ polylinker sequence)-to-*Pst*I (4 bp downstream of the 3′ end of Ad2 L3-derived sequences in the 3′ polylinker, blunt ended with T4 DNA polymerase) fragment isolated from pL3-18 and an 80-bp *Hin*cII [7 bp upstream of the poly(A) addition site]-to-*Sau*3AI (in 3′ polylinker) fragment isolated from pG3SVL-A were inserted into pGEM4 digested with *Eco*RI and *Bam*HI. A plasmid containing the Ad2 L3 AATAAA sequence and poly(A) site as well as 8 bp of the pGEM polylinker (pL3-L3) was prepared by inserting the *Eco*RI-*Pst*I fragment isolated from pL3-18 and the *Pst*I-*Sau*3AI fragment isolated from pRS/14 into pGEM4 digested with *Eco*RI and *Bam*HI.

The Ad2 E2a cleavage site substitution mutant, pG4E2-S, was made by the method of Kunkel et al. (13). Single-stranded DNA was prepared from M13mp19L3-A (40). Second-strand synthesis (DNA polymerase I large fragment; New England BioLabs, Beverly, Mass.) was primed with a 47-base oligonucleotide (synthesized by using an Applied Biosystems model 380A DNA synthesizer) that has the sequence 5′-GTGTACAAATAAAAACATTAACAACAA CAAAAAGTGTCTCCTAGTAC-3′. A *Hin*dIII-to-*Eco*RI fragment containing the substitution mutation was excised from M13mp19L3-A and inserted into pGEM4, resulting in pG4E2-S.

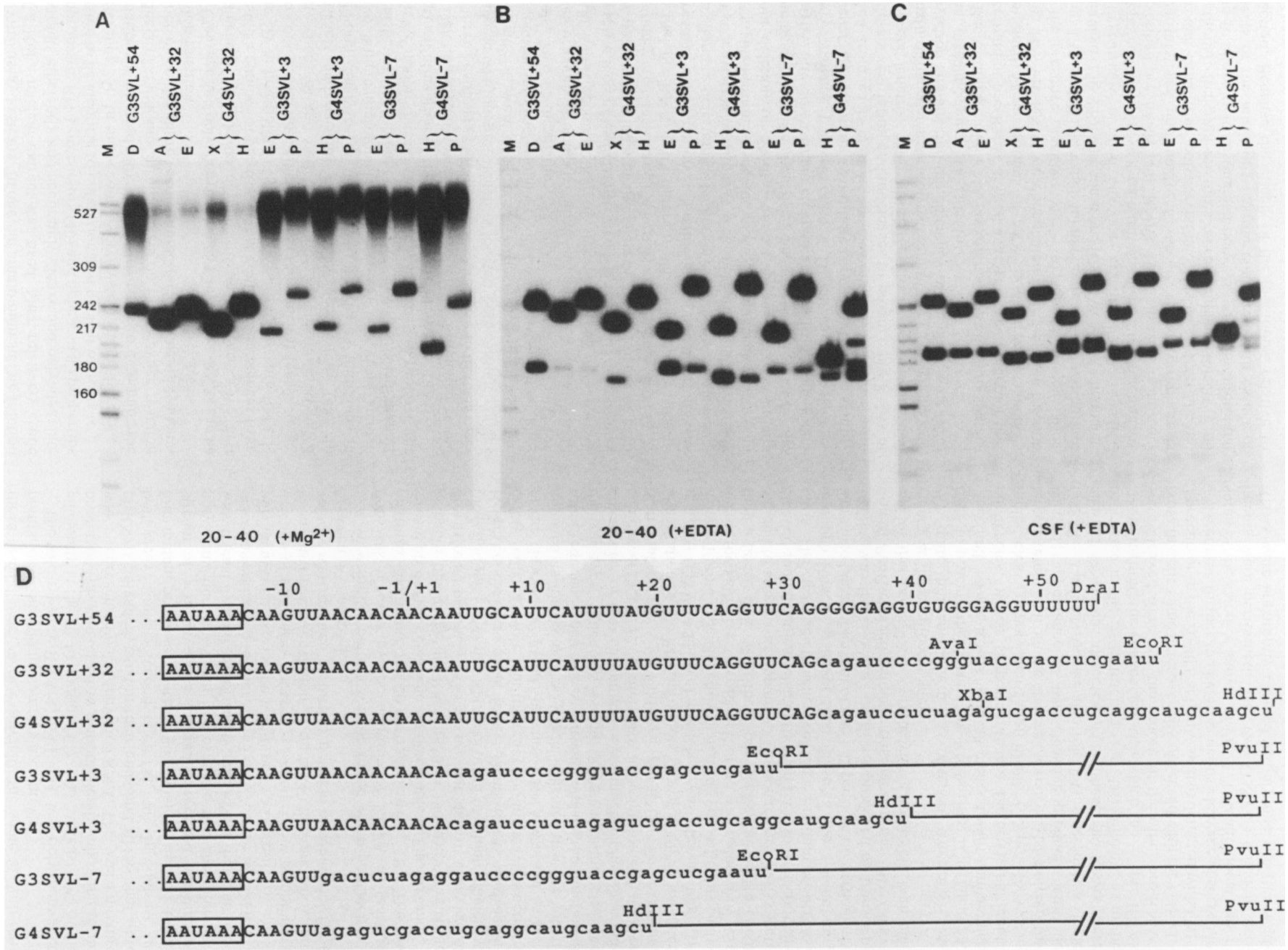
Preparation of fractionated nuclear extract. The 20 to 40% ammonium sulfate fraction (20-40 fraction) and the Superose 6 cleavage specificity factor (CSF) and poly(A) polymerase fractions were prepared from HeLa cell nuclear extracts as described previously (40).

RNA processing and analysis. To prepare runoff templates, plasmid DNA was digested with the appropriate restriction enzyme (as indicated in the figure legends), extracted with phenol-chloroform (1:1), and precipitated with ethanol. Capped pre-RNAs were synthesized with SP6 RNA polymerase essentially as previously described (12, 22). RNA processing reactions were performed and products were analyzed as described previously (32, 40). RNase T1 analysis of the cleaved RNA was performed as described by Moore et al. (26). Cleaved RNA was isolated as described previously (40).

RESULTS

To investigate the role in the cleavage-polyadenylation reaction of pre-mRNA sequences lying downstream of the conserved AAUAAA hexanucleotide, we synthesized a number of different wild-type and mutant pre-RNAs and analyzed their relative processing efficiencies in two types of

FIG. 1. Cleavage and polyadenylation of SV40 late pre-RNAs containing deletions of downstream sequences. Wild-type and mutant SV40 late pre-RNAs were incubated in the 20-40 fraction (A and B) or in a Superose 6 gel filtration fraction containing partially purified CSF (C) for 90 min in the presence of Mg^{2+} (A) or EDTA (B and C). (D) Structures of the wild-type and mutant pre-RNAs. Upper- and lowercase letters represent wild-type SV40 sequence and vector sequences, respectively. Vector sequences beyond the pGEM polylinkers are indicated by straight lines. Numbering of nucleotides and positions of deletion endpoints are relative to the SV40 late wild-type cleavage site, which lies between the nucleotides labeled −1 and +1. The 3′ end of each pre-RNA is indicated by the name of the restriction endonuclease that was used to linearize the DNA templates for runoff transcription by SP6 RNA polymerase. The first letters of these restriction endonuclease sites were used to label the lanes in panels A, B, and C (and indicate which pre-RNA was used in each processing reaction). RNA isolated from reactions was subjected to electrophoresis on 5% polyacrylamide–8.3 M urea gels. Numbers at the left of panel A indicate the sizes (in nucleotides) of pBR322 *Hpa*II size markers (lane M), which were used in all experiments. The RNA products migrating close to the 527-nt size marker represent polyadenylation RNA (A). Upstream cleavage products migrated between the 180- and 190-nt size markers (B and C). The downstream cleavage products (<85 nt) are not shown.


A
B
C
M
D
G3SVL+54
G3SVL+32
G4SVL+32
G3SVL+3
G4SVL+3
G3SVL-7
G4SVL-7
527
309
242
217
180
160
20-40 (+Mg2+)
20-40 (+EDTA)
CSF (+EDTA)
D
-10
-1/+1
+10
+20
+30
+40
+50
DraI
G3SVL+54 ...AAUAAACAAGUUAACAACAACAAUUGCAUUCAUUUUAUGUUUCAGGUUCAGGGGGAGGUGUGGGAGGUUUUUU
AvaI
EcoRI
G3SVL+32 ...AAUAAACAAGUUAACAACAACAAUUGCAUUCAUUUUAUGUUUCAGGUUCAGcagaucccccgggguaccgagcucgaauu
XbaI
HdIII
G4SVL+32 ...AAUAAACAAGUUAACAACAACAAUUGCAUUCAUUUUAUGUUUCAGGUUCAGcagauccucuagagucgaccugcaggcaugcaagcu
EcoRI
PvuII
G3SVL+3 ...AAUAAACAAGUUAACAACAACAcagaucccccgggguaccgagcucgauu
HdIII
PvuII
G4SVL+3 ...AAUAAACAAGUUAACAACAACAcagauccucuagagucgaccugcaggcaugcaagcu
EcoRI
PvuII
G3SVL-7 ...AAUAAACAAGUUgacucuagaggaucccccgggguaccgagcucgaaauu
HdIII
PvuII
G4SVL-7 ...AAUAAACAAGUUagagucgaccugcaggcaugcaagcu

in vitro assay systems. In one, an ammonium sulfate fraction of a HeLa nuclear extract was used. Such relatively crude extracts efficiently and accurately cleave and polyadenylate appropriate RNA precursors (40) and appear in all properties tested to be identical to unfractionated nuclear extracts (e.g., reference 25). In the second assay system, partially purified fractions were used (40). These allowed us to determine the effects of mutations in pre-RNAs on the cleavage reaction alone or on the entire reconstituted the cleavage-polyadenylation reaction.

Cleavage and polyadenylation of SV40 late pre-RNAs can be influenced by deletion of downstream sequences. We first tested a number of pre-RNAs that contained sequences from the SV40 late poly(A) addition site. SV40 late pre-RNA is processed with high efficiency in crude nuclear lysates (e.g., reference 48) and both by ammonium sulfate-fractionated extracts and partially purified cleavage activity (40). In addition, analyses of the effects of deletion mutations in vivo, both in microinjected *Xenopus laevis* oocytes (3) and in transfected COS monkey cells (33), have provided support for the notion that downstream sequences are required for optimal utilization of this poly(A) site in vivo.

Deletion mutants with SV40 sequences terminated at one of three positions, +32, +3, and −7, relative to the cleavage-polyadenylation site, which is located between nucleotides +1 and −1, were constructed by inserting the appropriate sequences into bacteriophage SP6 promoter-containing plasmids. Two of these, the +32 and +3 mutations, had been used previously for in vivo analyses, and each was shown to reduce the accumulation of specific mRNA by approximately a factor of 4 in transfected COS cells (33). To control for potential artifacts in the analysis, two variants of each mutant were constructed (Materials and Methods and Fig. 1) so that pre-RNAs with different plasmid sequences downstream of each deletion endpoint could be synthesized. This was done to determine whether any effects on the cleavage-polyadenylation reaction, or lack thereof, might have arisen from substitution of specific plasmid sequences rather than from deletion of SV40 sequences. This is a potentially significant concern because, as mentioned above, the poly(A) site downstream element appears to be rather simple, consisting perhaps only of U-rich or G+U-rich sequences. In addition, two different pre-RNAs were synthesized from each plasmid by cleaving the templates with different restriction endonucleases to produce pre-RNAs of different lengths. This allowed us to assess whether the length of the precursor downstream of the cleavage-polyadenylation site affected the efficiency of the processing reaction, as suggested previously (16).

Bacteriophage SP6 RNA polymerase was used to synthesize mutant pre-RNAs as well as a wild-type precursor that contained 54 nucleotides of SV40 downstream sequences (Fig. 1D). Pre-RNAs were incubated with the 20-40 fraction in the presence of either Mg^{2+}, to allow polyadenylation (Fig. 1A), or EDTA, which inhibits polyadenylation but not cleavage (Fig. 1B). The same pre-RNAs were also incubated with partially purified cleavage activity (CSF; Fig. 1C) prepared by Superose 6 gel filtration of the 20-40 fraction, as described previously (40). RNAs were extracted from reaction mixtures and analyzed directly by electrophoresis through denaturing polyacrylamide gels (see Materials and Methods). Reaction mixtures were incubated for 10 or 90 min to allow us to distinguish between possible differential effects of the mutations on rate compared with extent of reaction. Because the effects of the mutations, relative to those of wild type, were essentially the same at both time points, only the 90-min samples are shown.

The wild-type pre-RNA, G3SVL+54, was polyadenylated efficiently in the 20-40 fraction (Fig. 1A). At least 80% of the precursor was converted to a more slowly migrating form, which represented polyadenylated RNA. This was a mixture of RNAs polyadenylated at their ends and RNAs that were both cleaved and polyadenylated (32, 40). In contrast, all four mutant precursors with SV40 sequences deleted downstream of +32 gave rise to greatly reduced amounts of polyadenylated RNA, and we estimate that the inhibition was at least fivefold. Surprisingly, pre-RNAs that contained deletions extending further, to +3 or −7, were as efficiently polyadenylated as was the wild-type (+54) pre-RNA. Thus, deletion of sequences between +54 and +32 was inhibitory to in vitro polyadenylation of an SV40 late pre-RNA, but activity was restored by further deletion of sequences lying between +32 and +3.

A very similar pattern was observed when polyadenylation was blocked by the addition of EDTA to reaction mixtures containing the 20-40 fraction (Fig. 1B). The upstream products of cleavage are detectable in Fig. 1B (and C) as bands below the pre-RNAs (180 to 190 nt); whereas the downstream products (<85 nt) are not shown. Cleavage of the wild-type (+54) pre-RNA occurred at a high efficiency. However, cleavage of all four +32 pre-RNAs was significantly reduced. Note that the reduction in cleavage activity of the G4SVL+32 (*Xba*I) precursor RNA was not as pronounced as for the others, which indicates that the substituted sequences may at least partially replace the function of the wild-type sequences. As with polyadenylation, cleavage efficiencies of the +3 mutants were much higher than efficiencies of their +32 counterparts, although somewhat lower than that of the +54 wild-type pre-RNA. For example, quantitation of the RNA products shown in Fig. 1 indicates that the G3SVL+32 *Ava*I pre-RNA was cleaved approximately 10-fold less efficiently than wild type, whereas cleavage of the G3SVL+3 *Eco*RI precursor was reduced by less than a factor of 2. With these pre-RNAs, the length effect alluded to above was apparent. The longer pre-RNAs were cleaved somewhat less efficiently than their shorter counterparts. The cleavage efficiencies of the −7 mutant pre-RNAs, although somewhat lower than those of the +3 mutants, were also higher than that of the +32 pre-RNAs. Particularly striking was the G4SVL-7 pre-RNA produced by transcription of *Pvu*II-cleaved template (last lane in Fig. 1B). Cleavage of this pre-RNA occurred with a relatively high efficiency at three different sites.

A quite different pattern was observed when CSF was used to process these same pre-RNAs (Fig. 1C). The +54, +32, and +3 pre-RNAs were all cleaved with similar high efficiencies, which suggested that a factor(s) that could inhibit cleavage-polyadenylation of the +32 pre-RNAs was largely removed from the cleavage activity. Also noteworthy was the appearance of two cleavage products with the G3SVL+3 pre-RNAs (see also Fig. 2A), since only a single cleaved RNA was detected when these pre-RNAs were processed in the 20-40 fraction (Fig. 1B). Finally, as in the 20-40 fraction, the efficiencies with which the −7 pre-RNAs were cleaved were less than those of the other pre-RNAs. Note that the same heterogeneity observed with the G4SVL-7 precursor in the 20-40 fraction was also detected with CSF.

To pursue further the heterogeneous cleavage observed in some of the samples shown in Fig. 1, two additional experiments were performed. The two G3SVL+3 mutant pre-RNAs were apparently cleaved at only one site (the wild-

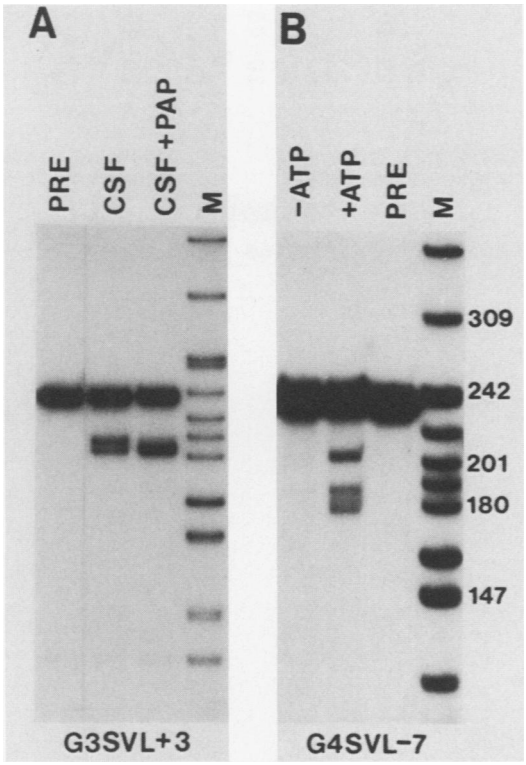


FIG. 2. Deletion of sequences near the cleavage site resulting in heterogeneous cleavage. (A) G3SVL+3 (*Eco*RI) pre-RNA was incubated in a standard processing reaction mixture with either 5 μl of the cleavage-containing Superose 6 fraction (CSF) or 5 μl of CSF plus 1 μl of the Superose 6 fraction containing poly(A) polymerase activity (CSF + PAP). (B) G4SVL-7 (*Pvu*II) pre-RNA was processed in a standard processing reaction mixture with and without creatine phosphate and ATP (+ATP and −ATP, respectively). EDTA (1 mM) was included in all reaction mixtures. Other lanes: PRE, unprocessed pre-RNA; M, size markers (indicated in nucleotides).

type site) in 20-40 fraction but at two sites by CSF. One possibility is that this difference was due to the presence or absence of poly(A) polymerase in these fractions. To test this possibility, we compared the patterns of cleavage produced by CSF with a G3SVL+3 pre-RNA in the presence and absence of partially purified poly(A) polymerase. The poly(A) polymerase-containing fraction was indeed able to suppress the heterogeneity observed with the cleavage activity alone (Fig. 2), which possibly reflected a functional interaction between cleavage and polyadenylation activities.

A perhaps different sort of heterogeneity was observed when cleavage of a −7 mutant pre-RNA was examined (Fig. 1B and C). In this case, cleavage appeared to have occurred at sites (all in vector sequences) separated by as much as 25 bases, and the pattern of cleavage products was not qualitatively different when the 20-40 fraction and CSF were compared. To provide support for the idea that these RNAs were all indeed produced by the enzymatic machinery that is responsible for the normal cleavage and polyadenylation of pre-RNAs, we investigated whether formation of these heterogeneous RNAs required ATP. The pre-RNA cleavage reaction may be unique in its apparent requirement for ATP but lack of requirement for a divalent cation (e.g., references 25 and 48). In the presence of EDTA, formation of heterogeneous cleavage products from the G4SVL-7 pre-RNA did in fact require ATP, creatine phosphate, or both; ATP was also added to the sample shown in the lane marked +ATP in Fig. 2B but not the lane marked −ATP (40). Further support for the idea that these RNAs were authentic cleavage products comes from the observation that the sizes of the products indicated that each of the cleavages had occurred downstream of AAUAAA and fell within the range in which naturally occurring cleavage sites are found (13 to 38 (±2) nt downstream of AAUAAA).

Optimal cleavage, but not polyadenylation, of an adenovirus late pre-RNA requires specific downstream sequences. To extend the type of analysis described above to an additional pre-RNA, we examined the efficiencies with which derivatives of an adenovirus late pre-RNA (L3 family) were processed. This poly(A) site has been shown previously to be used efficiently in vitro (25), and it has also been suggested that downstream sequences, between +5 and +48, are required for efficient in vitro processing (37). In contrast to the SV40 late pre-RNA, wild-type L3 precursor is not cleaved efficiently by isolated CSF, although it is processed efficiently when partially purified poly(A) polymerase and CSF are mixed (40).

To analyze the role of downstream sequences in cleavage-polyadenylation of L3 pre-mRNA in vitro, a series of 3′ deletion mutants was constructed and SP6-generated transcripts were processed under a variety of conditions. In the analysis shown in Fig. 3A, reaction mixtures lacked Mg^{2+} so that only products of the cleavage reaction were generated. Pre-RNAs were incubated with the 20-40 fraction or with CSF, either alone or mixed with Superose 6-purified poly(A) polymerase. Several points are apparent from this analysis. Most notably, all mutant pre-RNAs, even one that extended to within 3 nt of the AAUAAA (−18), gave rise to readily detectable levels of cleaved RNAs. The cleavage efficiencies of the pre-RNAs shown in Fig. 3A and B, as well as efficiencies of pre-RNAs terminated at +22 and +3, were estimated by densitometry of appropriate autoradiograms (Fig. 3C). Sequence downstream of +22 could be deleted without significantly affecting cleavage efficiency in both crude and fractionated extracts. (The slight decrease observed with the +25 and +22 pre-RNAs relative to wild type [+48] may reflect the longer lengths of the mutant pre-RNAs.) Deletions extending to +16 or +7 resulted in three- to fourfold reductions in cleavage efficiency. However, as observed with the SV40 late pre-RNA (Fig. 2), deletion of additional sequences appeared to restore cleavage efficiency. This result was most striking in reactions in which partially purified components were used. In fact, the pre-RNA deleted to −18 was cleaved by CSF plus poly(A) polymerase with nearly 90% the efficiency of the wild-type pre-RNA. As observed with certain of the SV40 late pre-RNAs that deleted the cleavage site itself, the cleavage products generated from the −18 mutant pre-RNA were somewhat heterogeneous. Finally, it is noteworthy that the separated cleavage activity was equally inactive on all L3 pre-RNAs tested, which ruled out the possibility that specific L3 downstream sequences were inhibitory to the activity of CSF.

To determine the effects of the L3 downstream mutations on polyadenylation, reactions were performed in the presence of Mg^{2+}. Analysis of the processed RNAs revealed that these mutations had little, if any, effect on the efficiency of polyadenylation. Results obtained with two of these pre-RNAs are shown in Fig. 4; other mutants gave similar results (data not shown).

An AAUAAA not normally used as a cleavage-polyadenyla-

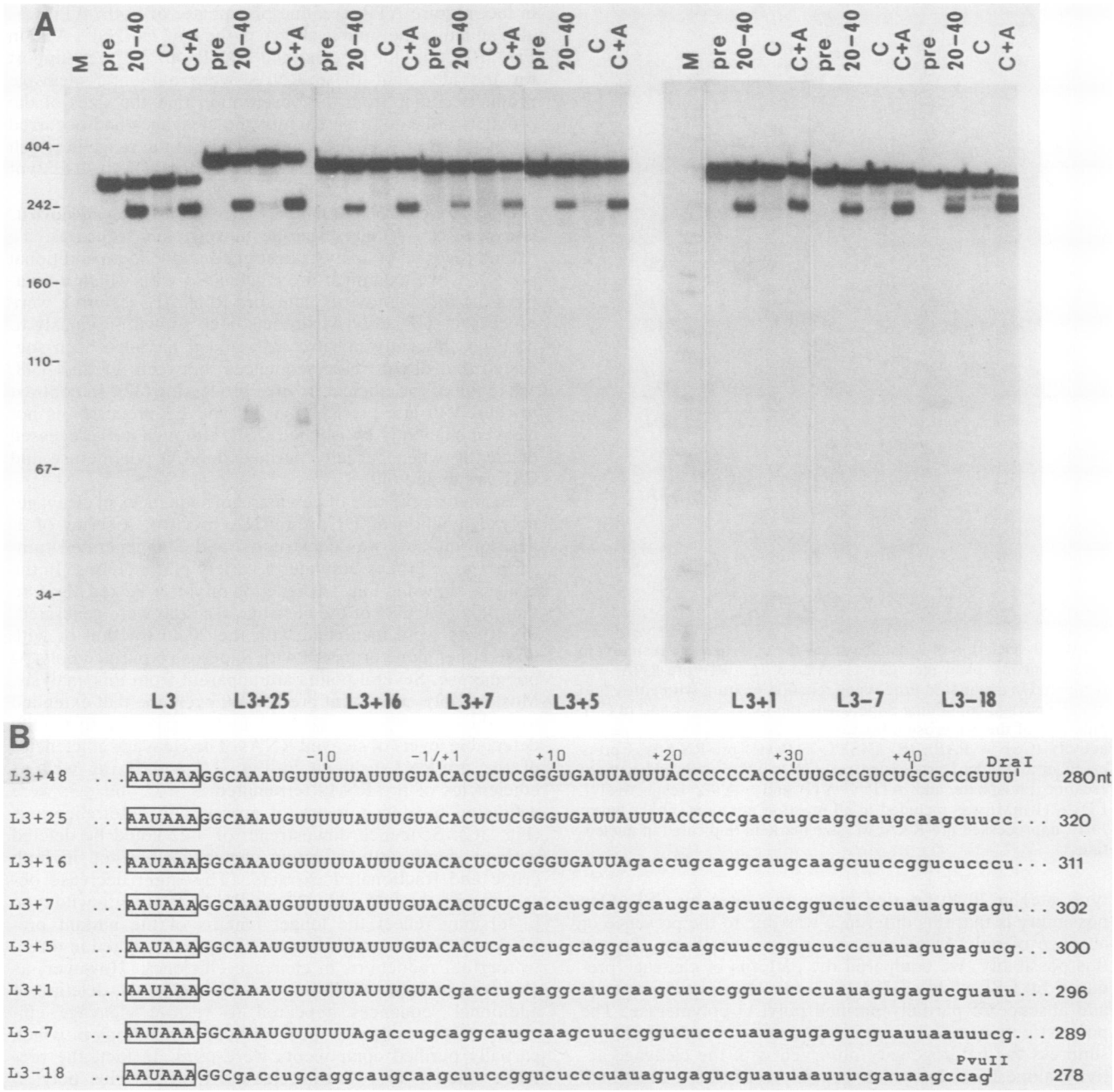


FIG. 3. Cleavage of Ad2 L3 3′ deletion mutant pre-RNAs. (A) Gels of a series of Ad2 L3 3′ deletion mutant pre-RNAs processed in standard reaction mixtures (12.5 μl) that contained 2.5 ng of pre-RNA and 3 μl of the 20-40 fraction (20-40) or 4 μl of CSF alone (C) or mixed with 1 μl of Mn^{2+}-dependent poly(A) polymerase activity (C+A) in the absence of divalent cation. Pre-RNAs (pre) were run for comparison. Sequences of 3′ portions of the wild-type and mutant pre-RNAs. Ad2 L3-derived sequences are in capital letters; the pGEM4 sequences that follow are in lowercase letters. The AAUAAA sequence is boxed. 3′ deletion mutants are designated to indicate the positions of the 3′ ends of Ad2 L3-derived sequence that are located upstream (−) or downstream (+) of the normal poly(A) site. Lengths (in nucleotides) of the pre-RNAs (lane M) are shown on the right. (C) Cleavage efficiencies of Ad2 L3 wild-type and 3′ deletion mutant pre-RNAs are shown as molar percentages of the upstream cleavage products and plotted against the positions of the 3′ ends of Ad2 L3-derived sequences. Shown are results of cleavage reactions that were obtained by using the 20-40 fraction (○) or CSF mixed with Mn^{2+}-dependent poly(A) polymerase (●). Cleavage efficiencies were determined by scanning appropriate autoradiograms with a model 1650 densitometer (Bio-Rad Laboratories, Richmond, Calif.).

tion signal can be activated by specific downstream sequences. The results present above indicate that deletion of sequences downstream of AAUAAA sequences can reduce the efficiency of pre-RNA 3′-end processing in vitro. An alternative approach to investigate the functional significance of such sequences is to determine whether they are capable of activating a nonfunctional AAUAAA. To test this, we analyzed processing of pre-RNAs containing an AAUAAA that lies within coding sequences for the SV40 large T antigen (43). This AAUAAA is not used to direct polyadenylation at

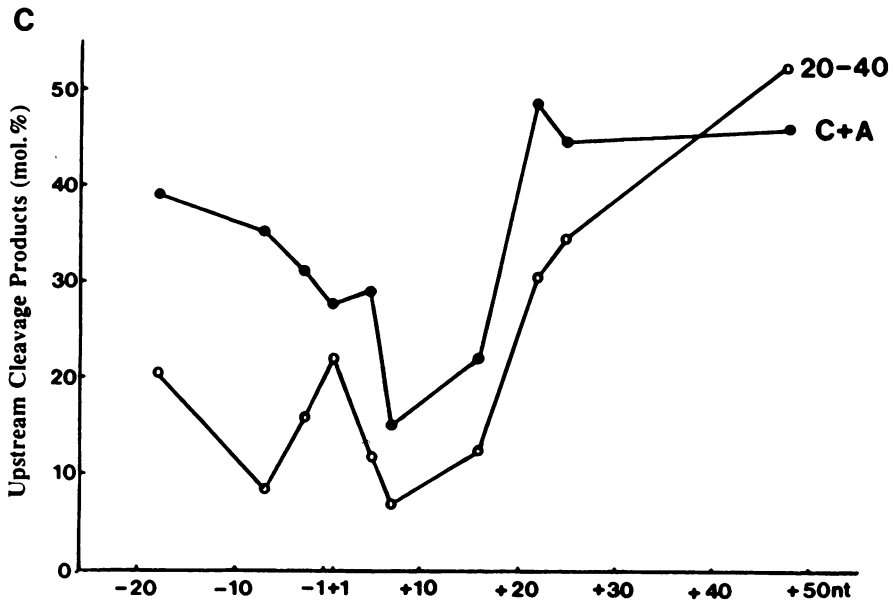


FIG. 3—*Continued*

detectable levels in vivo. An appropriate fragment of SV40 DNA was cloned into an SP6 promoter-containing plasmid to produce transcripts for processing reactions (see Materials and Methods). To determine whether pre-RNAs containing this AAUAAA are substrates for in vitro cleavage-polyadenylation, two runoff transcripts that differed only in length (LT/*Hin*cII [287 nt] and LT/*Rsa*I [261 nt]; Fig. 5) were synthesized. These RNAs were processed in fractions lacking Mg^{2+}, and the products were analyzed by gel electrophoresis as for Fig. 3. Although both pre-RNAs could be cleaved at detectable levels, the efficiency of processing was very low (Fig. 5A). The shorter pre-RNA was cleaved more efficiently than the longer one, and in each case the partially purified components cleaved these RNAs considerably more efficiently than did the 20-40 fraction. That the products detected were indeed authentic cleavage products is supported by the fact that longer exposure of the gel shown revealed RNAs of the size expected for the downstream product of the reaction (data not shown).

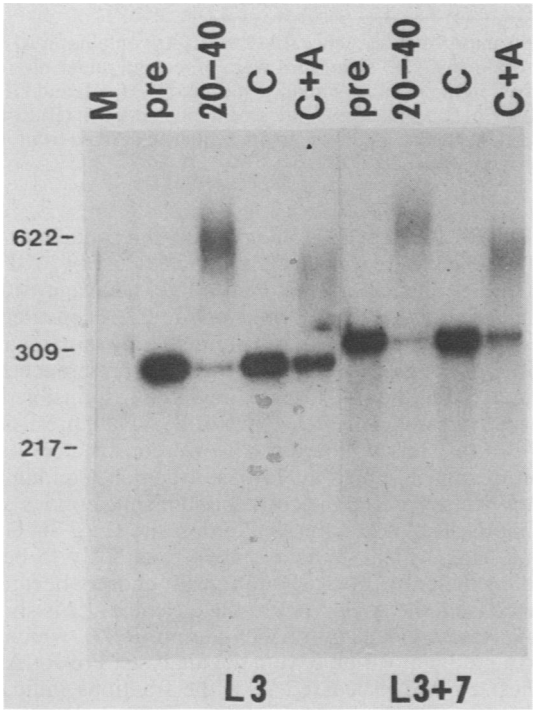


FIG. 4. Polyadenylation of an Ad2 L3 3′ deletion mutant pre-RNA. Wild-type pre-RNA (L3) and the 3′ deletion mutant pre-RNA (L3+7) that was least efficiently cleaved were processed under the reaction conditions described in the legend to Fig. 3 except that 1 mM Mg^{2+} was included in reaction mixtures. Lane M, Size markers (indicated in nucleotides).

Two possible explanations for the relative inactivity of the AAUAAA sequence in these RNAs are that sequences downstream of the hexanucleotide are in some manner inhibitory or, alternatively, that these sequences are not inhibitory but lack specific sequences that are positively required for the cleavage-polyadenylation reaction. To differentiate between these two hypotheses, we constructed three recombinant plasmids in which the sequences downstream of AAUAAA were replaced with foreign sequences (Fig. 5B and Materials and Methods). In two cases (LTΔ11 and LTΔ6), these were nonspecific plasmid DNA sequences; in the third, sequences from downstream of the Ad2 L3 AAUAAA were used (LT-L3). Transcripts from these plasmids were processed and analyzed as before. Replacement of the SV40 coding sequences with plasmid sequences had no effect on the processing efficiency of the transcripts (Fig. 5A). In contrast, the L3 sequences gave rise to transcripts that were processed efficiently both in the crude extract and by purified components, even though the LT-L3 pre-RNA was significantly longer than the other precursors. In addition, a construct analogous to LT-L3 in which the downstream sequence of SV40 late was used instead also activated the nonfunctional AAUAAA (data not shown). These results provide further support for the idea that poly(A) site downstream sequences play a positive role in the pre-mRNA cleavage reaction.

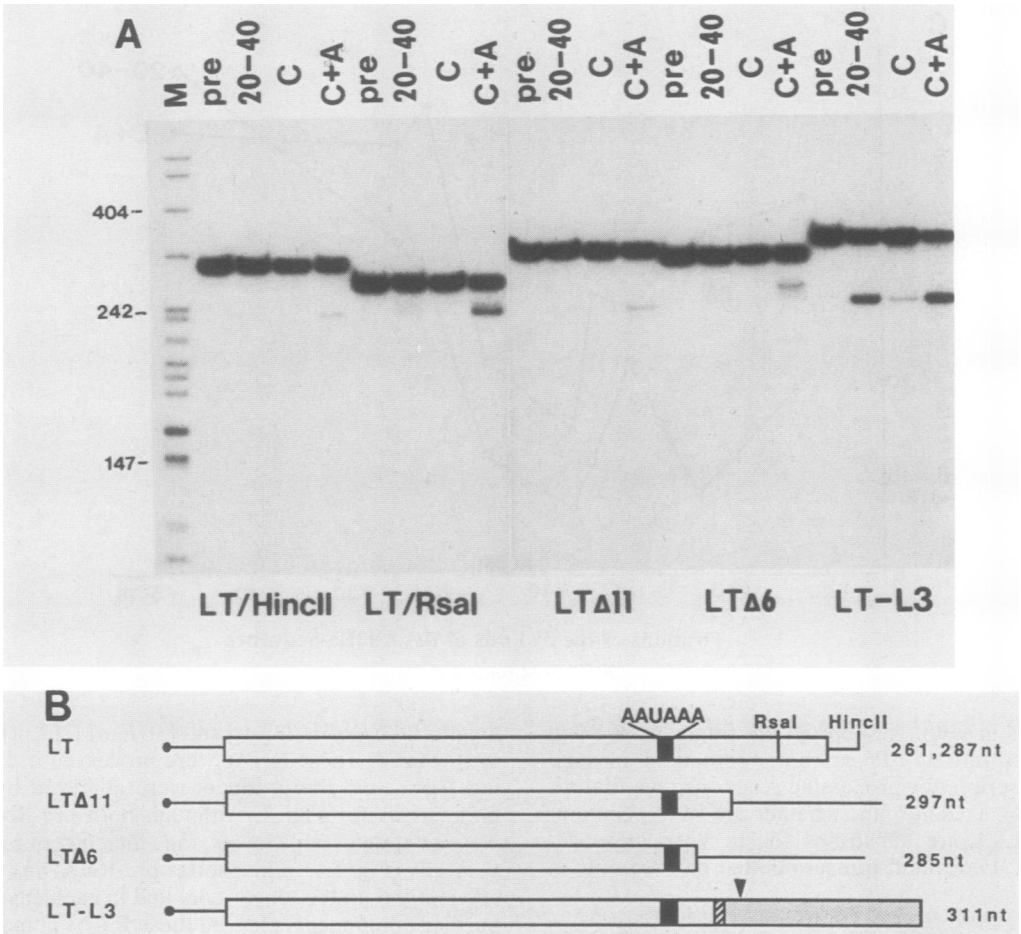


FIG. 5. Cleavage of pre-RNAs that contain a normally inactive AAUAAA hexanucleotide sequence. (A) Pre-RNAs contains an AAUAAA sequence derived from the SV40 large-T-antigen-coding region and a variety of downstream sequences were processed under the reaction conditions specified in the legend to Fig. 3. The downstream sequences are derived from the intact coding sequence (LT/*Hin*cII and LT/*Rsa*I), pGEM3 sequences (LT/Δ11 and LT/Δ6), or the Ad2 L3 poly(A) site (LT-L3). Lane M, Size markers (indicated in nucleotides). (B) Structures of the pre-RNAs. Symbols: □, SV40 large-T-antigen-coding region; ■, AAUAAA sequence; □, downstream sequence derived from Ad2 L3; ▼, authentic cleavage site; ———, pGEM3-derived sequence; ▨, pGEM4-derived sequence.

Sequences at the SV40 late cleavage site are required for cleavage by CSF. The SV40 pre-RNA was the only one of several tested that could be efficiently cleaved by partially purified CSF (40; see also above). To determine what sequences in addition to AAUAAA are required for this substrate preference, we carried out two experiments. In the first, plasmids containing chimeric poly(A) sites were constructed (Fig. 6B and Materials and Methods). In SVL-L3, SV40 late sequences downstream of position −7 (relative to the cleavage site) were replaced with the corresponding sequences from the Ad2 L3 poly(A) site; in L3-SVL, essentially the opposite configuration was generated. Transcripts from these plasmids and the corresponding wild-type templates were synthesized, processed, and analyzed as described above. SV40 late sequences downstream of −7 were sufficient to convert the inactive L3 pre-RNA into a precursor that could be efficiently cleaved by CSF (Fig. 6). The L3-L3 pre-RNA was analyzed to eliminate the possibility that activation of the L3-SVL pre-RNA was due to the inserted linker instead of the SV40 sequences.

Two reasons led us to suspect that sequences at or near the SV40 late cleavage site might in some manner facilitate the high reactivity of SV40 late pre-RNAs with CSF. First, the deletion analysis of the SV40 late downstream sequence required for cleavage (Fig. 1) revealed that sequences downstream of +3 were dispensable for cleavage by CSF. Second, the sequence at the SV40 late cleavage site was unusual, consisting of the trinucleotide CAA repeated three times. Because CA is found immediately upstream of a majority of pre-mRNA cleavage sites (e.g., reference 21), we speculated that the CAA repeat might be responsible for the observed high efficiency cleavage of SV40 late pre-RNAs by CSF. To test this notion directly, we constructed a substitution mutant in which an 11-base fragment containing the CAA repeats replaced the corresponding nucleotides encompassing the Ad2 E2a poly(A) addition site (Fig. 7). The E2a poly(A) site, which we have shown previously to be processed inefficiently by CSF (40), was chosen because the distance from the AAUAAA to the cleavage site is similar to the equivalent distance in SV40 late pre-RNA, which eliminated a potential variable from the analysis. Pre-RNAs were synthesized and processed with the fractions indicated in Fig. 7A. Analysis of the products revealed that the CAA repeats were indeed sufficient to allow efficient cleavage of the E2a pre-RNA. RNase T1 analysis of the cleaved RNA (data not shown) revealed that the majority of it was terminated precisely at the site corresponding to that used in SV40 late pre-RNA (Fig. 7B).

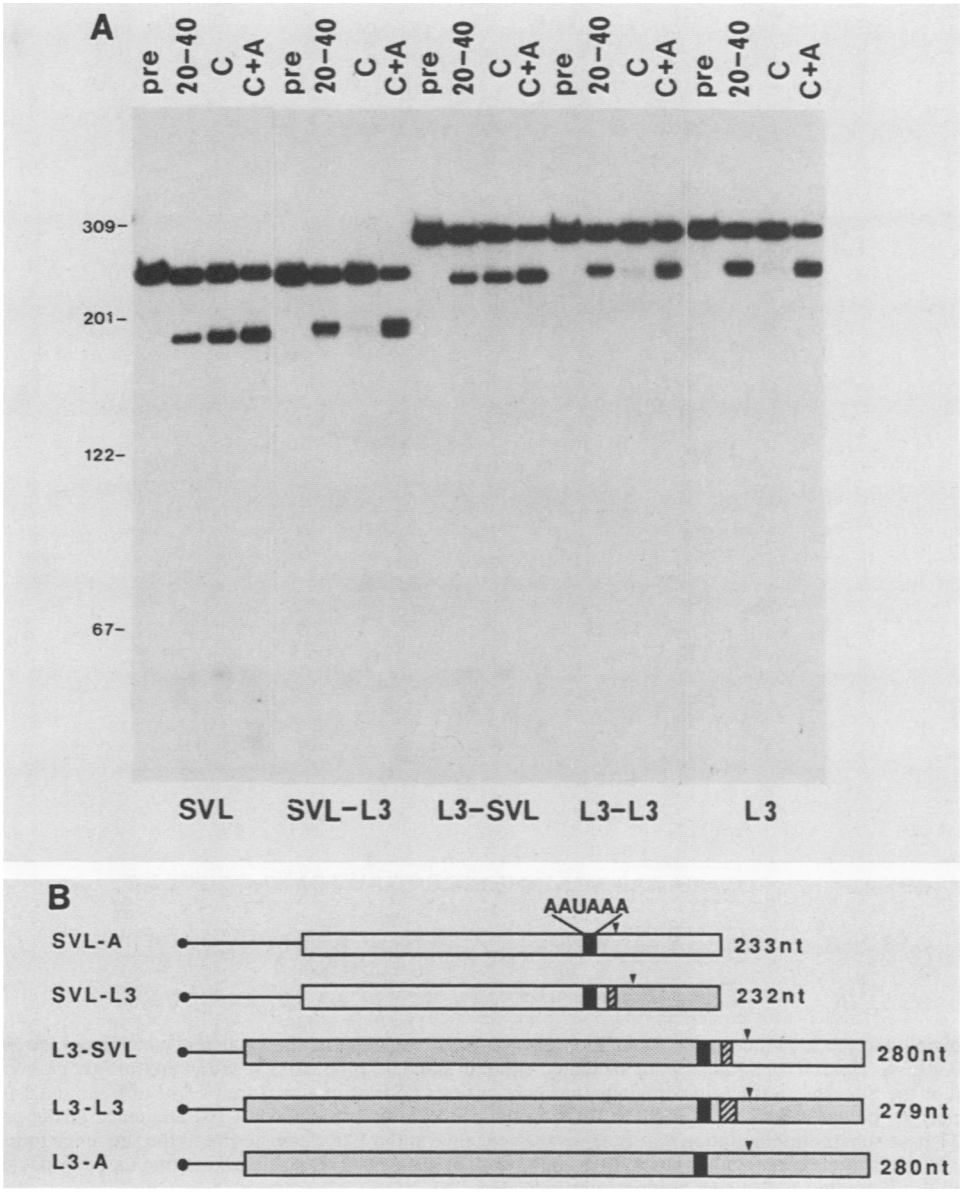


FIG. 6. Cleavage of chimeric pre-RNAs. (A) Wild-type SV40 late (SVL) and Ad2 L3 (L3) and chimeric pre-RNAs were processed under the reaction conditions specified in the legend to Fig. 3. The SV40 late upstream sequence was joined to the Ad2 L3 downstream sequence (SVL-L3), and the Ad2 L3 upstream sequence was joined to the SV40 late (L3-SVL) or Ad2 L3 (L3-L3) downstream sequence. Lane M, Size markers (indicated in nucleotides). (B) Structures of the pre-RNAs used. Symbols: □, SV40 late sequence; ⊡, Ad2 L3 sequence; ■, AAUAAA hexanucleotide; ▨, pGEM vector-derived polylinker sequence; ——, 5′ sequences derived from pGEM vectors; ▼, authentic poly(A) sites. To prepare pre-RNAs, plasmid DNAs were digested at *Dra*I sites located 53 bp downstream of the SV40 late poly(A) site or 48 bp downstream of the Ad2 L3 poly(A) site.

DISCUSSION

The results presented here suggest that sequences in pre-mRNAs located downstream of the AAUAAA hexanucleotides can influence the cleavage-polyadenylation reaction both qualitatively and quantitatively. The qualitative effects brought about by mutations consist solely of what appears to be heterogeneity in vitro in the cleavage reaction. This was detected by the appearance of more than one upstream RNA product when cleavage was assayed in the absence of polyadenylation. These species most likely result from endonucleotic cleavage of the pre-RNA at more than one specific site. Although cleavage site heterogeneity in vitro has not previously been reported, several apparent examples of this phenomenon have been observed in vivo (17, 29, 34, 36, 44, 47). The results of these studies, coupled with those presented here, strongly suggest that sequences surrounding the cleavage site itself play a role in directing cleavage to a particular site. However, precisely what distinguishes a preferred cleavage site from an inactive one, other than an apparently strict requirement for an adenosine immediately preceding the cleaved bond, is not clear. Although we determined the sequence of only one cleaved RNA precisely to the nucleotide, the apparent sizes of all of the products are consistent with cleavage having occurred adjacent to an A residue.

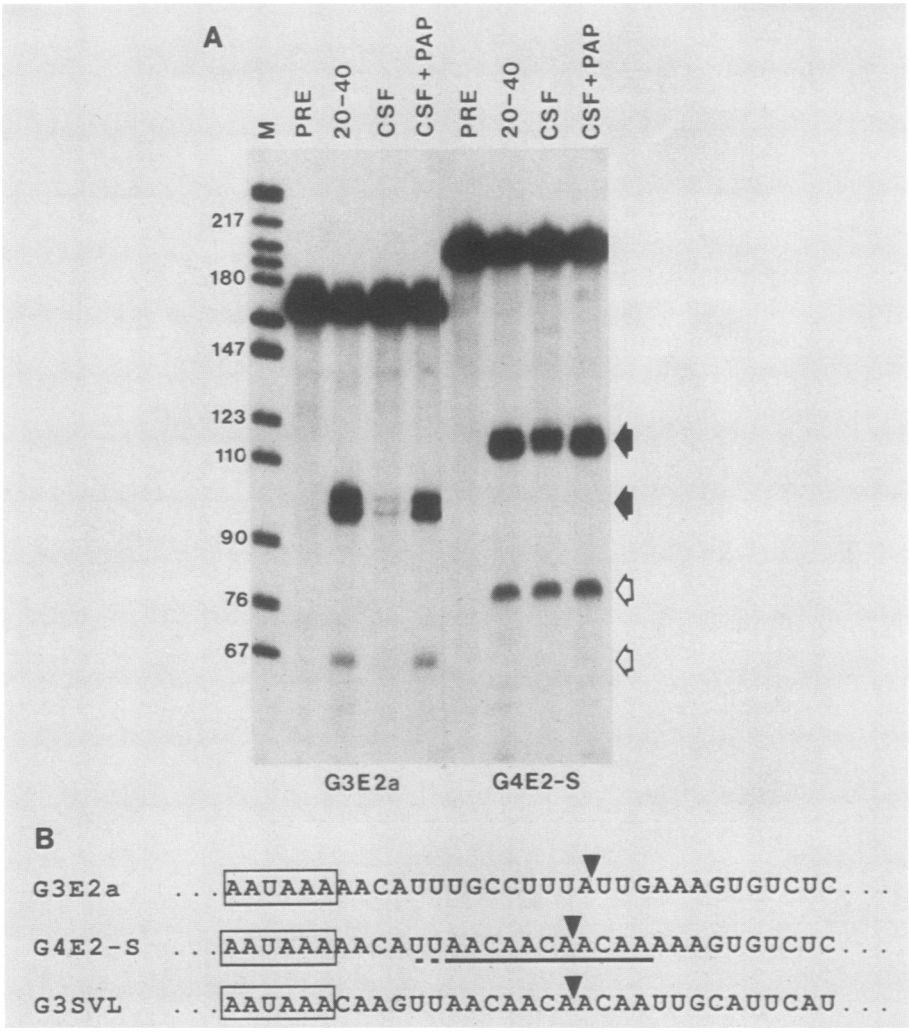


FIG. 7. Results showing that the sequence at the SV40 late cleavage site is sufficient to allow efficient cleavage of the Ad2 E2a pre-RNA. (A) G3E2a (*Xba*I) and G4E2-S (*Hin*dIII) pre-RNAs were incubated either in standard processing reaction mixtures containing 5 μl of the 20-40 fraction (20-40) or 5 μl of the Superose 6 fraction containing partially purified CSF or in a mixture of 5 μl of CSF and 1 μl of the Superose 6 fraction containing poly(A) polymerase activity (CSF + PAP) in the presence of 1 mM EDTA. (B) Sequence of the pre-RNAs. G4E2-S pre-RNA contains an 11-base substitution mutation that changes the sequence at the E2a cleavage site to the sequence found at the SV40 late cleavage site. Symbol: ▼, sites of cleavage. Other lanes: PRE, unprocessed pre-RNA; M, size markers (indicated in nucleotides). Positions of the upstream (⧫) and downstream (◊) cleavage products are indicated.

Our findings provide insights into the interaction of the factors that catalyze cleavage-polyadenylation with the pre-RNA cleavage site. First, with one exception, isolated CSF appears to cleave all pre-RNAs tested at the same site used in the presence of poly(A) polymerase. Although CSF cleaves only SV40 late pre-RNAs efficiently, low levels of cleavage of other precursors can be detected (e.g., Fig. 3), and, where examined, these products appear identical in size to their counterparts produced in the presence of poly(A) polymerase. Thus, in most instances, the accuracy of the cleavage reaction is not influenced by poly(A) polymerase. The one exception, the G3SVL+3 pre-RNA (Fig. 1D), suggests that subtle features of RNA sequence or structure can influence cleavage specificity, since the related G4SVL+3 precursor is cleaved at a single site (Fig. 1). The elimination of the second cleavage product after addition of the poly(A) polymerase-containing fraction suggests that the polymerase can, at least in this instance, influence the cleavage reaction. We believe that this results from a direct physical interaction between the polymerase and CSF, although our hypothesis remains to be proven.

Our results also show that the sequence in the SV40 late pre-RNA that allows it to be cleaved efficiently by CSF is that surrounding the cleavage site itself and the nucleotides lying further downstream are not important in this reaction. It is noteworthy that this sequence, a CAA repeat, was not found in any of the large number of poly(A) sites that we examined. Although other explanations remain viable, we believe these observations suggest that the ability of SV40 late pre-RNA to be cleaved by CSF may be fortuitous, reflecting a preferred interaction between the fractionated factor and pre-RNA that is not important in an unfractionated or reconstituted system.

The results of our analyses comparing processing of different pre-RNAs in crude and fractionated extracts leads to the suggestion that the quantitative effects on cleavage and polyadenylation that result from mutations downstream of the cleavage site reflect at least in part the interactions of

factors not directly involved in the enzymatic reaction with the pre-RNA. This view is supported by studies with all three pre-RNAs tested, i.e., SV40 late, Ad2 L3, and that derived from the SV40 large-T-antigen-coding region.

The cleavage pattern of the SV40 late pre-RNAs provides perhaps the strongest evidence supporting the above hypothesis. The facts that cleavage (and polyadenylation) of the SVL+32 pre-RNAs was significantly inhibited in crude extracts, but not when fractionated components were used, and that both fractionated and crude extracts processed the wild-type (+54) pre-RNA with equally high efficiencies suggest that a component that enhances cleavage and polyadenylation only in a crude extract interacts with sequences between +32 and +54. The striking observation that the requirement for these sequences is eliminated when sequences between +3 and +32 are deleted suggests that the role of the far downstream sequences may be to prevent an inhibitory effect of the sequences between +3 and +32. This again appears not to reflect a direct interaction with the 3′-end formation machinery, since the purified components are not affected by this deletion. A possible mechanism to explain these observations is discussed below.

The results obtained with the other two pre-RNAs tested also support the view that the role of downstream sequences is at least in part indirect. The RNA substrates containing the AAUAAA from within the large-T-antigen-coding region, although processed very inefficiently, are cleaved at easily detectable levels only by fractionated components. Similarly, all mutant, but not wild-type, L3 pre-RNAs were cleaved more efficiently by partially purified components than in the crude extract. This finding was particularly dramatic in the case of the pre-RNAs deleted to −7 and −18. The sequences required for optimal cleavage of the L3 precursor in crude extracts lie between +7 and +22, substantially closer to the cleavage site than in the SV40 late pre-RNA. However, as with SV40 late pre-RNA, cleavage efficiency is partially restored by further deletion, which suggests that sequences upstream of +7 may be inhibitory to cleavage in the absence of specific sequences situated further downstream. It is interesting that, unlike the case with SV40 late pre-RNA, polyadenylation was not detectably inhibited by any of the deletions we tested. Polyadenylation of an SV40 early pre-RNA in a HeLa whole-cell lysate, in which cleavage does not occur, was similarly shown to be insensitive to deletion of downstream sequences (16). Likewise, polyadenylation of SV40 late (48) and L3 (26) pre-RNAs totally lacking downstream sequences (precleaved RNAs) occurs efficiently in vitro. Together, these results suggest that AAUAAA-dependent polyadenylation is much less dependent on downstream sequences than is the cleavage reaction. One exception to this rule is the SVL+32 pre-RNA, in which polyadenylation is significantly inhibited. The basis for these differing results is unclear.

Do the downstream sequences implicated in affecting 3′ processing in vitro play a similar role in vivo? Two studies using different types of in vivo assays have examined the effects of deleting sequences downstream of the SV40 late poly(A) addition site. Sadofsky et al. (33) showed that the sequence AGGUUUUUU, which is located between 46 and 54 bases downstream of the cleavage site, is required for optimal accumulation of SV40-specific RNA in the cytoplasm of transfected COS cells. Deletion of this sequence resulted in about a fourfold reduction in mRNA levels, which agrees well with the effects observed when SV40 late pre-RNAs lacking sequences downstream of +32 were processed in crude nuclear extracts. However, when a similar analysis was carried out on SV40 RNAs synthesized from plasmids microinjected into *X. laevis* oocytes, deletion of the sequences defined by the former experiments appeared to have no effect on cleavage-polyadenylation (3). In contrast, a U-rich region between 9 and 26 bases 3′ to the poly(A) site was found to be required for optimal 3′-end formation. The basis for these apparently contradictory results is not clear but may reflect differences in the plasmid constructs or the very different assay systems used.

Recently, Zarkower and Wickens (51) presented results indicating that sequences between +7 and +23 in the SV40 late poly(A) site are required for optimal cleavage in HeLa nuclear extracts and that sequences further downstream are not. Although we do observe that the sequences between +3 and +32 influence cleavage efficiency in both the 20-40 fraction and CSF (Fig. 1 and data not shown), the effects observed were much smaller than those observed with the deletion of sequences downstream of +32. In support of the view presented here that bases downstream of +32 can influence cleavage-polyadenylation, Wilusz et al. (46) have observed that deletion of sequences downstream of +29 reduces the efficiency of cleavage in nuclear extracts. The apparent discrepancy between these results could again reflect differences in the precursors or assay conditions used; nevertheless, these findings suggest a complex but nonessential role for downstream sequences.

At the moment, we can only speculate on the identity of the factors that interact with downstream sequences. However, a model consistent with the data and involving both positive and negative interactions between *trans*-acting factors and the pre-RNA can be put forth. A potentially important factor when one considers the mechanism of nuclear pre-mRNA processing is the higher-order structure of the pre-RNA, e.g, its association with proteins to form heterogeneous nuclear ribonucleoproteins (hnRNPs) (reviewed in reference 4). hnRNP particles form rapidly on nascent RNA in vivo (5, 6) and in vitro (18), and at least one type of hnRNP (the C protein) has been suggested to play a role in pre-RNA splicing (2). Wilusz et al. (46) have observed, by using UV cross-linking, that C protein can bind to sequences lying downstream of several poly(A) addition sites. Most relevant, they found that C proteins interact specifically with sequences between +29 and +54 on an SV40 late pre-RNA. Similarly, Moore et al. (24) have found that C protein specifically interacts with the U-rich sequence just upstream of the Ad2 L3 cleavage site. Interestingly, both of these regions (+29 to +54 in SV40 late and −12 to −3 in Ad2 L3) correspond to deletion mutants that exhibit very low levels of cleavage activity in the 20-40 fraction (SVL+32, Ad2L3-7, and −18 mutants) but exhibit nearly wild-type levels of cleavage activity when assayed in the more purified Superose 6 fractions (Fig. 1 and 3). On the basis of these observations, we suggest that formation of a specific hnRNP structure involving this sequence properly positions the pre-RNA such that the AAUAAA and cleavage site are readily accessible to the polyadenylation machinery. In the absence of these sequences, hnRNPs, or perhaps other RNA-binding proteins, interact with the RNA in a manner that interferes with polyadenylation. Interestingly, the nucleotide sequence between +3 and +32 of the SV40 late pre-RNA contains an excellent match to the pre-mRNA 3′ splice site consensus, $Py_{(n > 11)}NPyAG/G$ (28). Several proteins, including hnRNPs, that are able to bind to 3′ splice sites have been identified (8, 31, 42; G. Dreyfuss, personal communication). We suggest that in the absence of properly positioned hnRNPs, one of these proteins can bind to this 3′

splice site, thereby blocking the cleavage-polyadenylation reaction. Partial purification of the polyadenylation machinery might be expected to remove these factors, thereby eliminating both the inhibitory effect of the sequences situated between +3 and +32 and the requirement for sequences downstream of +32. Although the situation with the Ad2 L3 RNA may not be completely analogous, we note that similar U-rich sequences are appropriately situated in this pre-RNA, and these may also play a role in C-protein binding (24, 39). A more detailed understanding of factors that interact with downstream sequences should result from reconstitution experiments with purified components.

ACKNOWLEDGMENTS

We thank M. Kapczynski and J. Voulgaris for excellent technical assistance, W. Weast for help in preparing the manuscript, and L. Chasin for the use of his densitometer. We also thank J. Wilusz, T. Shenk, C. Moore, and G. Dreyfuss for communicating their results to us before publication. L.C.R. and Y.T. contributed equally to this work.

This work was supported by Public Health Service grant GM 28983 from the National Institutes of Health.

LITERATURE CITED

1. **Birnstiel, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41**:349–359.
2. **Choi, Y. D., P. J. Grabowski, P. A. Sharp, and G. Dreyfuss.** 1986. Role for heterogeneous nuclear ribonucleoproteins in RNA splicing. Science **231**:1534–1539.
3. **Conway, L., and M. Wickens.** 1985. A sequence downstream of AAUAAA is required for formation of simian virus 40 late mRNA 3′ termini in frog oocytes. Proc. Natl. Acad. Sci. USA **82**:3949–3953.
4. **Dreyfuss, G.** 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Annu. Rev. Cell Biol. **2**: 459–498.
5. **Dreyfuss, G., Y. D. Choi, and S. A. Adam.** 1984. Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. Mol. Cell. Biol. **4**:1104–1114.
6. **Economidis, I. V., and T. Pederson.** 1983. Structure of nuclear ribonucleoprotein: heterogeneous nuclear RNA is complexed with a major sextet of proteins *in vivo*. Proc. Natl. Acad. Sci. USA **80**:1599–1602.
7. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5′-AAUAAA-3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
8. **Gerke, V., and J. A. Steitz.** 1986. A protein associated with small nuclear ribonucleoprotein particles recognizes the 3′ splice site of premessenger RNA. Cell **47**:973–984.
9. **Gil, A., and N. J. Proudfoot.** 1984. A sequence downstream of AAUAAA is required for rabbit β-globin mRNA 3′-end formation. Nature (London) **312**:473–474.
10. **Gil, A., and N. J. Proudfoot.** 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA 3′ end formation. Cell **49**:399–406.
11. **Higgs, D. R., S. E. Y. Goodbourn, J. Lamb, J. B. Clegg, D. J. Weatherall, and N. J. Proudfoot.** 1983. α-Thalassemia caused by a polyadenylation signal mutation. Nature (London) **306**:398–400.
12. **Konarska, M. M., R. A. Padgett, and P. A. Sharp.** 1984. Recognition of cap structure in splicing in vitro of mRNA precursors. Cell **38**:731–736.
13. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
14. **Manley, J. L.** 1983. Accurate and specific polyadenylation of mRNA precursors in a soluble whole-cell lysate. Cell **33**: 595–605.
15. **Manley, J. L.** 1988. Polyadenylation of mRNA precursors. Biochim. Biophys. Acta **950**:1–12.
16. **Manley, J. L., H. Yu, and L. C. Ryner.** 1985. RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation in vitro. Mol. Cell. Biol. **5**:373–379.
17. **Mason, P. J., J. A. Elkington, M. M. Lloyd, M. B. Jones, and J. G. Williams.** 1986. Mutations downstream of the polyadenylation site of a Xenopus β-globin mRNA affect the position but not the efficiency of 3′ processing. Cell **46**:263–270.
18. **Mayrand, S. H., N. Pedersen, and T. Pederson.** 1986. Identification of proteins that bind tightly to pre-mRNA during *in vitro* splicing. Proc. Natl. Acad. Sci. USA **83**:3718–3722.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
20. **McDevitt, M. A., M. J. Imperiale, H. Ali, and J. R. Nevins.** 1984. Requirement of a downstream sequence for generation of a poly(A) addition site. Cell **37**:993–999.
21. **McLaughlan, J., D. Gaffney, J. L. Whiton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from AATAAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. **13**:1347–1368.
22. **Melton, D., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
23. **Montell, C., E. F. Fisher, M. H. Caruthers, and A. J. Berk.** 1983. Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3′ consensus sequence AAUAAA. Nature (London) **305**:600–605.
24. **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.
25. **Moore, C. L., and P. A. Sharp.** 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell **41**:845–855.
26. **Moore, C. L., H. Skolnik-David, and P. A. Sharp.** 1986. Analysis of RNA cleavage at the adenovirus-2 L3 polyadenylation site. EMBO J. **5**:1929–1938.
27. **Moore, C. L., H. Skolnik-David, and P. A. Sharp.** 1988. Assembly of a 50S ribonucleoprotein complex requires ATP and intact polyadenylation signals. Mol. Cell. Biol. **8**:226–233.
28. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
29. **Norbury, C. J., and M. Fried.** 1987. Polyomavirus early region alternative poly(A) site: 3′-end heterogeneity and altered splicing pattern. J. Virol. **61**:3754–3758.
30. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′ non-coding region sequences in eukaryotic messenger RNA. Nature (London) **263**:211–214.
31. **Ruskin, B., P. D. Zamore, and M. R. Green.** 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell **52**:207–219.
32. **Ryner, L. C., and J. L. Manley.** 1987. Requirement for accurate and efficient mRNA 3′ end cleavage and polyadenylation of a simian virus 40 early pre-RNA in vitro. Mol. Cell. Biol. **7**: 495–503.
33. **Sadofsky, M., S. Connelly, J. L. Manley, and J. C. Alwine.** 1985. Identification of a sequence element on the 3′ sides of AAUAAA which is necessary for simian virus 40 late mRNA 3′-end processing. Mol. Cell. Biol. **5**:2713–2719.
34. **Sasavage, N. L., M. Smith, S. Gillam, R. P. Woychik, and F. M. Rottman.** 1982. Variation in the polyadenylation site of bovine prolactin mRNA. Proc. Natl. Acad. Sci. USA **79**:223–227.
35. **Sheets, M. D., P. Stephenson, and M. Wickens.** 1987. Products of in vitro cleavage and polyadenylation of simian virus 40 late pre-mRNAs. Mol. Cell. Biol. **7**:1518–1529.
36. **Simonsen, C. C., and A. D. Levinson.** 1983. Analysis of processing and polyadenylation signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. Mol. Cell. Biol. **3**:2250–2258.
37. **Skolnik-David, H., C. L. Moore, and P. A. Sharp.** 1987. Electrophoretic separation of polyadenylation-specific complexes. Genes Dev. **1**:672–682.

38. **Stefano, J. E., and D. E. Adams.** 1988. Assembly of a polyadenylation-specific 25S ribonucleoprotein complex in vitro. Mol. Cell. Biol. **8**:2052–2062.
39. **Swanson, M. S., and G. Dreyfuss.** 1988. Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. Mol. Cell. Biol. **8**:2237–2241.
40. **Takagaki, Y., L. C. Ryner, and J. L. Manley.** 1988. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. Cell **52**:731–742.
41. **Taya, Y., R. Devos, J. Tavernier, H. Cheroutre, G. Engler, and W. Fiers.** 1982. Cloning and structure of the human interferon-γ chromosomal gene. EMBO J. **1**:953–958.
42. **Tazi, J., C. Alibert, J. Temsamani, I. Reveillaud, G. Cathala, C. Burnel, and P. Jeanteur.** 1986. A protein that specifically recognizes the 3′ splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein. Cell **47**: 755–766.
43. **Tooze, J. (ed.).** 1981. DNA tumor viruses, 2nd ed., revised. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
44. **Weidemann, L. M., and R. Perry.** 1984. Characterization of the expressed gene and several processed pseudogenes for the mouse ribosomal protein L30 gene family. Mol. Cell. Biol. **4**:2518–2525.
45. **Wickens, M., and P. Stephenson.** 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3′ end formation. Science **206**:1045–1051.
46. **Wilusz, J., D. I. Feig, and T. Shenk.** 1988. The C proteins of heterogeneous nuclear ribonucleoprotein complexes interact with RNA sequences downstream of polyadenylation cleavage sites. Mol. Cell. Biol. **8**:4477–4483.
47. **Woychik, R. P., R. H. Lyons, L. Post, and F. M. Rottman.** 1984. Requirements for the 3′ flanking region of the bovine growth hormone gene for accurate polyadenylation. Proc. Natl. Acad. Sci. USA **81**:3944–3948.
48. **Zarkower, D., P. Stephenson, M. Sheets, and M. Wickens.** 1986. The AAUAAA sequence is required for both cleavage and polyadenylation of simian virus 40 pre-mRNA in vitro. Mol. Cell. Biol. **6**:2317–2323.
49. **Zarkower, D., and M. Wickens.** 1987. Formation of mRNA 3′ termini: stability and dissociation of a complex involving the AAUAAA. EMBO J. **6**:177–186.
50. **Zarkower, D., and M. Wickens.** 1987. Specific pre-cleavage and post-cleavage complexes involved in the formation of SV40 late mRNA 3′ termini in vitro. EMBO J. **6**:4185–4192.
51. **Zarkower, D., and M. Wickens.** 1988. A functionally redundant downstream sequence in SV40 late pre-mRNA is required for mRNA 3′ end formation and for assembly of a precleavage complex in vitro. J. Biol. Chem. **263**:5780–5788.
52. **Zhang, F., and C. N. Cole.** 1987. Identification of a complex associated with processing and polyadenylation in vitro of herpes simplex virus type 1 thymidine kinase precursor RNA. Mol. Cell. Biol. **7**:3277–3286.
53. **Zhang, F., R. M. Denome, and C. N. Cole.** 1986. Fine-structure analysis of the processing and polyadenylation region of the herpes simplex virus type 1 thymidine kinase gene by using linker scanning, internal deletion, and insertion mutations. Mol. Cell. Biol. **6**:4611–4623.