Electrophoretic separation of polyadenylation-specific complexes

Hagit Skolnik-David, Claire L. Moore,[1] and Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

A polyadenylation-specific complex composed of precursor RNA containing the adenovirus-2 L3 site and HeLa cellular components was detected by electrophoresis on a native, low-percentage polyacrylamide gel. Upon incubation in a reaction containing ATP and nuclear extract, precursor RNA was rapidly assembled into this complex. This assembly did not require poly(A) synthesis, as it occurred efficiently in the presence of ATP analogs that inhibited this reaction. Mutation of the hexanucleotide AAUAAA 20 nucleotides upstream of the L3 site to AAGAAA or deletion of sequence between +5 and +48 nucleotides downstream of the L3 site inactivates polyadenylation. The specific complex did not effectively form on substrate RNA with either the AAGAAA mutation or the downstream deletion mutation. Kinetic experiments showed that the assembly of this complex preceded processing of precursor RNA. We propose that formation of this complex represents an intermediate step in polyadenylation.

[*Key Words*: Eukaryotes; RNA processing; polyadenylation; ribonucleoprotein]

Received May 6, 1987; revised version accepted July 7, 1987.

Polyadenylation is an obligate step in the maturation of most eukaryotic mRNAs. The recent development of accurate and efficient in vitro systems has begun to clarify the biochemical mechanism of this process. Formation of a polyadenylated 3′ terminus occurs in two major steps; RNA sequences downstream of the polyadenylation site are released by an endonucleolytic cleavage, which is followed by the rapid polymerization of a poly(A) tract onto the newly generated 3′ terminus (Moore and Sharp 1985; Moore et al. 1986; Zarkower et al. 1986). The efficiency of the in vitro reaction is governed by the same sequence elements that have been implicated in vivo (Zarkower et al. 1986; Sperry and Berget 1986; Hart et al. 1985b). These include the highly conserved AAUAAA upstream of the polyadenylation site, as well as a much less well conserved downstream sequence, often in the form of a GU- or U-rich stretch (for review, see Birnstiel et al. 1985; Gil and Proudfoot 1987).

The nature of the factors responsible for sequence recognition and enzymatic processing in the polyadenylation reaction remains poorly understood. In addition to the common factors that must be present in all cell types, other factors must also exist that specify the alternative site selection that occurs in different cell types for some transcriptional units. Small nuclear ribonucleoprotein particles (snRNPs) of the U class are probably involved in polyadenylation because antisera to these snRNPs will inhibit the in vitro reaction (Moore and Sharp 1985; Sperry and Berget 1986). These antisera will also immunoprecipitate a component in nuclear extracts that binds to the AAUAAA sequence upstream of the polyadenylation site (Hashimoto and Steitz 1986). Although suggestive, these data are not conclusive that an snRNP particle is intimately involved in the polyadenylation process.

In this paper we show that in vitro cleavage and poly(A) synthesis at the adenovirus-2 L3 site is dependent on the AAUAAA sequence and a downstream sequence element. In addition, the interactions between components in HeLa nuclear extracts and polyadenylation-specific sequences are investigated by direct electrophoresis of complexes on nondenaturing, low-percentage polyacrylamide gels.

Results

Synthesis of substrate RNAs

The formation of polyadenylation complexes in nuclear extracts was studied using the different RNA substrates depicted in Figure 1. All RNA substrates were radioactively labeled by transcription with T3 RNA polymerase in the presence of [α-^{32}P]UTP. The standard precursor, L3(102), was 102 nucleotides in length and contained 54 nucleotides of sequence upstream of the adenovirus-2 L3 polyadenylation site. This RNA was prepared from a DNA template cleaved at the *Dra*I site. In some cases a longer precursor, L3(281), was used, which contained 227 nucleotides of sequence downstream of the poly-

[1]Present address: Department of Molecular Biology and Microbiology, Tufts University Medical School, Boston, Massachusetts 02111 USA.

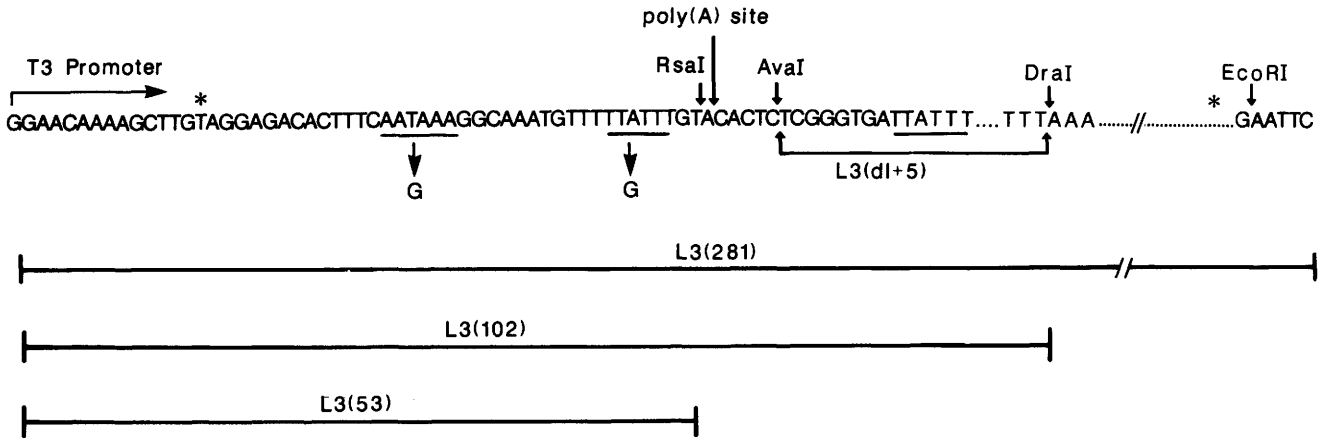


Figure 1. Sequence representation of the adenovirus-2 L3 polyadenylation site and description of substrate RNAs. Plasmid DNA containing the L3 site was digested with the indicated restriction endonucleases to create DNA templates for in vitro transcription of substrate RNA by T3 RNA polymerase. DNA cleaved with *Dra*I or *Eco*RI yielded L3(102) or L3(281), respectively. These transcripts were 102 or 281 nucleotides long and contained sequences flanking the L3 poly(A) site. DNA digested with *Rsa*I gave L3(53), a transcript 53 bases in length, which terminated 1 nucleotide upstream of the L3 site. Recombinant DNAs containing three different mutations were constructed by site-specific mutagenesis as follows: The AATAAA 20 nucleotides upstream of the L3 site was changed to AAGAAA; the TTATTT 3 nucleotides upstream of the L3 site was changed to TTGTTT; and in the third, L3(dl + 5), the sequences from +5 to +48 downstream of the L3 site were deleted. (*) Boundaries of adenovirus sequence.

adenylation site. Truncated RNA with 3′ termini 1 nucleotide upstream of the polyadenylation site, L3(53), was transcribed from a DNA template digested with *Rsa*I.

Three mutations of the L3 site were also studied (Fig. 1). In the first mutant, pT3L3G, the polyadenylation signal hexanucleotide AATAAA, 20 bases upstream of the polyadenylation site, was changed to AAGAAA. This mutation was selected to produce a totally inactive site. Although the conservation of the AATAAA sequence in genes of higher eukaryotes is remarkable, natural variants do occur. However, the T residue is present as the third base in all known sites (Wickens and Stephenson 1984), and when it is mutated to either a G, C, or A residue, cleavage efficiency is reduced greatly in vivo (Montell et al. 1983; Wickens and Stephenson 1984) and in vitro (Sperry and Berget 1986; Zarkower et al. 1986). There is a sequence complementary to AATAAA, 4 nucleotides upstream of the L3 site, and it has been suggested that such sequences could affect the efficiency of the reaction. In the second mutant, pL3UUGUU, this sequence (TTTATT) was changed to TTTGTT to destroy the complementarity. In the third mutant pL3(dl + 5), sequences beyond the L3 site, from +5 to +48, were removed. This deleted segment contained a TTATTT motif that is similar to downstream sequences important in the activity of other sites (Hart et al. 1985a). All mutations were confirmed by sequence analysis (data not shown).

Sequence requirements for cleavage and poly(A) synthesis in vitro at the L3 site

Polyadenylation of long substrate RNAs occurs by two steps: cleavage and poly(A) synthesis. Under conditions optimal for in vitro polyadenylation, these two steps appear coupled, as there is no accumulation of cleaved but not polyadenylated RNA (Moore and Sharp 1985). We have shown previously that a truncated form of the L3 substrate RNA, which has a 3′ terminus, 1 base upstream of the L3 site, is polyadenylated when incubated in a standard reaction (Moore et al. 1986). The kinetics of polyadenylation of this substrate RNA L3(53) was as rapid as that of a longer substrate RNA, L3(102), which was only polyadenylated following cleavage (Fig. 2, lanes 1–12). In both cases, the poly(A)$^+$ product possessed a tract of adenosine residues approximately 350 nucleotides in length and migrated slower in the 10% polyacrylamide/8.3 M urea gel than the substrate RNA. With the L3(53) substrate RNA, a series of product RNAs that increased in length by 1 adenosine residue at a time were observed (Fig. 2, lanes 7 and 8). The presence of this ladder suggests that during synthesis of the first few residues of the poly(A) tract, the poly(A) polymerase might dissociate frequently from the substrate RNA but become more processive and efficient as the length of the tract increases. A similar series of short-product RNAs was not observed during polyadenylation of the longer RNA substrate. Whether this reflects a more processive poly(A) synthesis reaction is unclear. Although rapid polyadenylation in vitro of the L3(53) RNA did not require the presence of sequences downstream of the poly(A) site, it was dependent on an intact AAUAAA sequence. Substrate from the mutant L3(53)AAGAAA, which was identical to L3(53) RNA, except for the point mutation in the hexanucleotide sequence, was completely inactive for polyadenylation under identical conditions (Fig. 2, lanes 13–18). This agrees with previous results from the study of the late SV40 polyadenylation site (Zarkower et al. 1986) and suggests that the observed polyadenylation of the truncated L3 substrate is related specifically to the in vivo process.

The effect of mutations on polyadenylation of longer

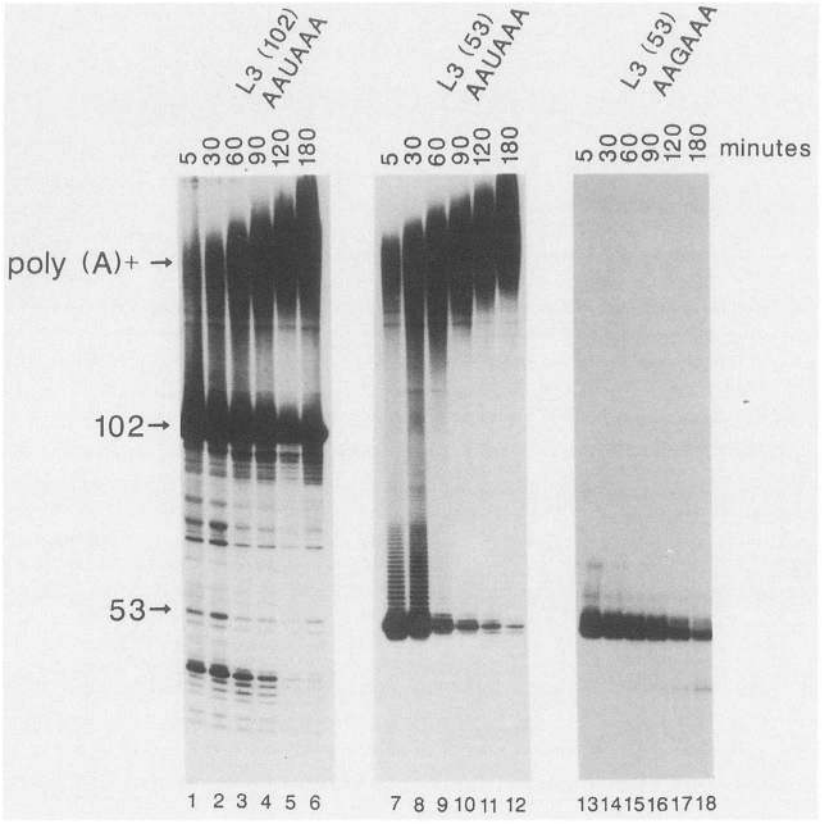


Figure 2. Activity of truncated RNA substrates. ^{32}P-Labeled substrate RNA was incubated with HeLa nuclear extract under conditions optimal for polyadenylation. The reactions were stopped at the times indicated by freezing, and purified RNAs were analyzed by electrophoresis on a 15% polyacrylamide/8.3 M urea gel. The unreacted substrate RNAs, 53 and 102 nucleotides long, and the polyadenylated RNA product, poly(A)$^+$, are indicated by arrows. Reactions contained the following substrates: L3(102) (lanes 1–6); L3(53) RNA (lanes 7–12); L3(53) RNA, with the AAGAAA mutation described in the legend to Fig. 1 (lanes 13–18).

precursor RNAs was also examined. To quantitate the level of accurate polyadenylation, RNAs were purified from reaction mixtures after different periods of incubation and hybridized to a single-stranded M13 cDNA recombinant (Fig. 3A). The hybrids were treated with ribonuclease T2 to degrade single-stranded RNA sequences. Accurately cleaved and polyadenylated RNA yielded a protected fragment of 76 nucleotides, whereas untreated precursor yielded the 43- and 60-nucleotide fragments. (The 60-nucleotide fragment is generated from the unprocessed substrate RNA; see legend to Fig. 3B.) Substrate RNA from the mutant L3(102)UUGUU, in which the sequence UUUAUU complementary to hexanucleotide had been changed to UUUGUU, was processed accurately and efficiently (Fig. 3B, lanes 9–12). In contrast with these results and in agreement with the results using the truncated substrate (Fig. 2), full-length substrate RNA containing the AAGAAA mutation in the conserved hexanucleotide was completely inactive by this analysis (Fig. 3B, lanes 1–4). In comparison with the activity of the normal precursor L3(102)AAUAAA (Fig. 3B, lanes 5–8), deletion of sequences from +5 to +48 nucleotides downstream of the poly(A) site in substrate RNA L3(dl + 5) inactivated in vitro polyadenylation activity (Fig. 3B, lanes 13–16). When total RNA from reactions containing the two inactive mutants was examined directly by gel electrophoresis (as in Fig. 2), very little ^{32}P-labeled RNA shifted in mobility to the position of poly(A)$^+$ RNA (data not shown). This suggests that cryptic cleavage sites in these substrates had not been utilized and that very little poly(A) synthesis occurred at the preexisting ends of the mutant substrate RNAs.

RNA substrate specificity and reaction conditions for the formation of polyadenylation-specific complexes

Complexes were formed on precursor RNA by incubation with nuclear extracts of HeLa cells under conditions optimal for polyadenylation. Heparin was added, and the reaction was further incubated on ice for 10 min. Products of the reaction were analyzed by electrophoresis on a nondenaturing 4% polyacrylamide gel. The addition of heparin blocks further polyadenylation (data not shown) and reduces the nonspecific binding of components in the extract to the negatively charged RNA. Under these gel conditions, heparin treatment is necessary to get a good yield of stable complexes to migrate homogeneously in the gel (Konarska and Sharp 1986).

Incubation of ^{32}P-labeled precursor RNA L3(102) in reactions containing nuclear extract resulted in the rapid formation of two distinct complexes (called complex A

and complex B; Fig. 4A). All of the substrate RNA was incorporated into these two complexes. Complex A migrated faster than complex B and appeared as a smear in the lower part of the gel. Complex A formed very rapidly on precursor RNA, even in the presence of heparin and during incubation on ice (data not shown). A complex with similar mobility has been described previously in the analysis of complexes involved in splicing (Konarska and Sharp 1986). Formation of A-type complexes occurs on all added RNA substrates, independent of the presence or absence of sequences related to either polyadenylation (Fig. 4A) or splicing (Konarska and Sharp 1986). In addition, formation of complex A does not require inclusion of ATP in the reaction (Fig. 4A, lanes 6–11). This complex migrates more slowly than free RNA in the gel and thus must represent the binding of one or more sequence-nonspecific factors to the RNA. In studies of splicing complexes, it has been shown that complexes equivalent to A contain one or more C proteins of the heterogeneous nuclear ribonucleoprotein (hnRNP) particle (Choi et al. 1986; M. Konarska, unpubl.).

Complex B forms within a few minutes of incubation at 30°C, in the presence of ATP (Fig. 4A, lanes 1–6). The low level of complex B observed at the start of the reaction is the result of formation of this complex during incubation on ice for 10 min in the presence of heparin. In fact, inclusion of 5 mg/ml of heparin in a polyadenylation reaction does not block the formation of complex B. This is unusual in that addition of heparin to splicing reactions blocks the formation of all specific complexes involving the precursor RNA (Konarska and Sharp 1986).

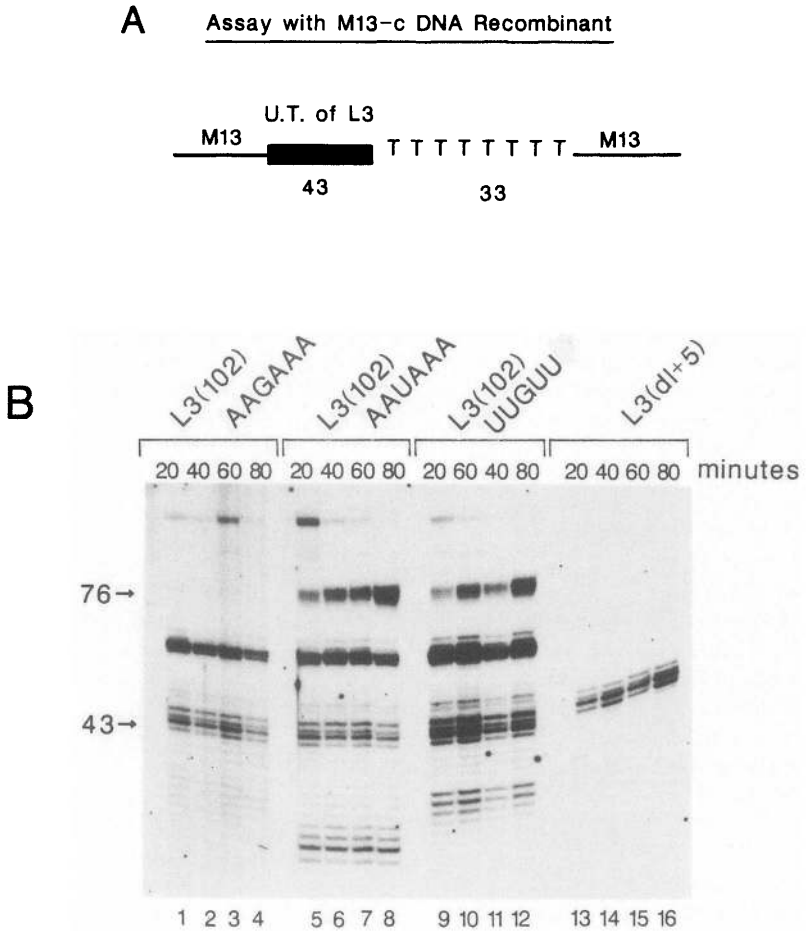


Figure 3. Sequence requirements for accurate cleavage and polyadenylation at the L3 site. (*A*) The M13 cDNA recombinant used for hybridization analysis. This single-stranded DNA is complementary to 43 nucleotides of the 3′ untranslated region (U.T.) of L3 and a tract of 33 adenosine residues (Moore and Sharp 1985). (*B*) Activity of mutated substrate RNAs. L3(102)AAGAAA (lanes 1–4), L3(102)AAUAAA (lanes 5–8), L3(102)UUGUU (lanes 9–12), and L3(dl + 5) (lanes 13–16) substrate RNAs were incubated for the times indicated with HeLa nuclear extract and ATP at 30°C, as described previously (Moore and Sharp 1985). ^{32}P-Labeled RNA from these reactions was hybridized to the M13 cDNA clone. Hybrids were digested with ribonuclease T2, and protected fragments run on a 15% polyacrylamide/8.3 M urea gel. Correctly polyadenylated RNA gives a 76-nucleotide-long fragment, and precursor yields bands at 43 and 60 nucleotides. The 60-nucleotide fragment is an artifact of the hybridization assay that results in protection from the unprocessed substrate RNA of the 17 nucleotides of RNA immediately downstream at the L3 polyadenylation site. Because most of this region is missing from substrate RNA L3(dl + 5), the 60-nucleotide fragment was produced in lanes 13–16.

After 5 min of incubation at 30°C, almost all of the L3(102) substrate had been converted into complex B, and the level of this complex remains fairly constant for 60 min. The efficient generation of complex B requires the addition of ATP. The small amount of complex B generated in the absence of this high-energy cofactor probably reflects a low level of endogenous ATP in the nuclear extract (Fig. 4A, lanes 7–12). Generation of complex B did not require poly(A) synthesis because a complex with similar mobility assembled in reactions containing either a nonhydrolyzable analog of ATP, α, β-methylene-adenosine 5′-triphosphate [AMP(CH_2)P_2],

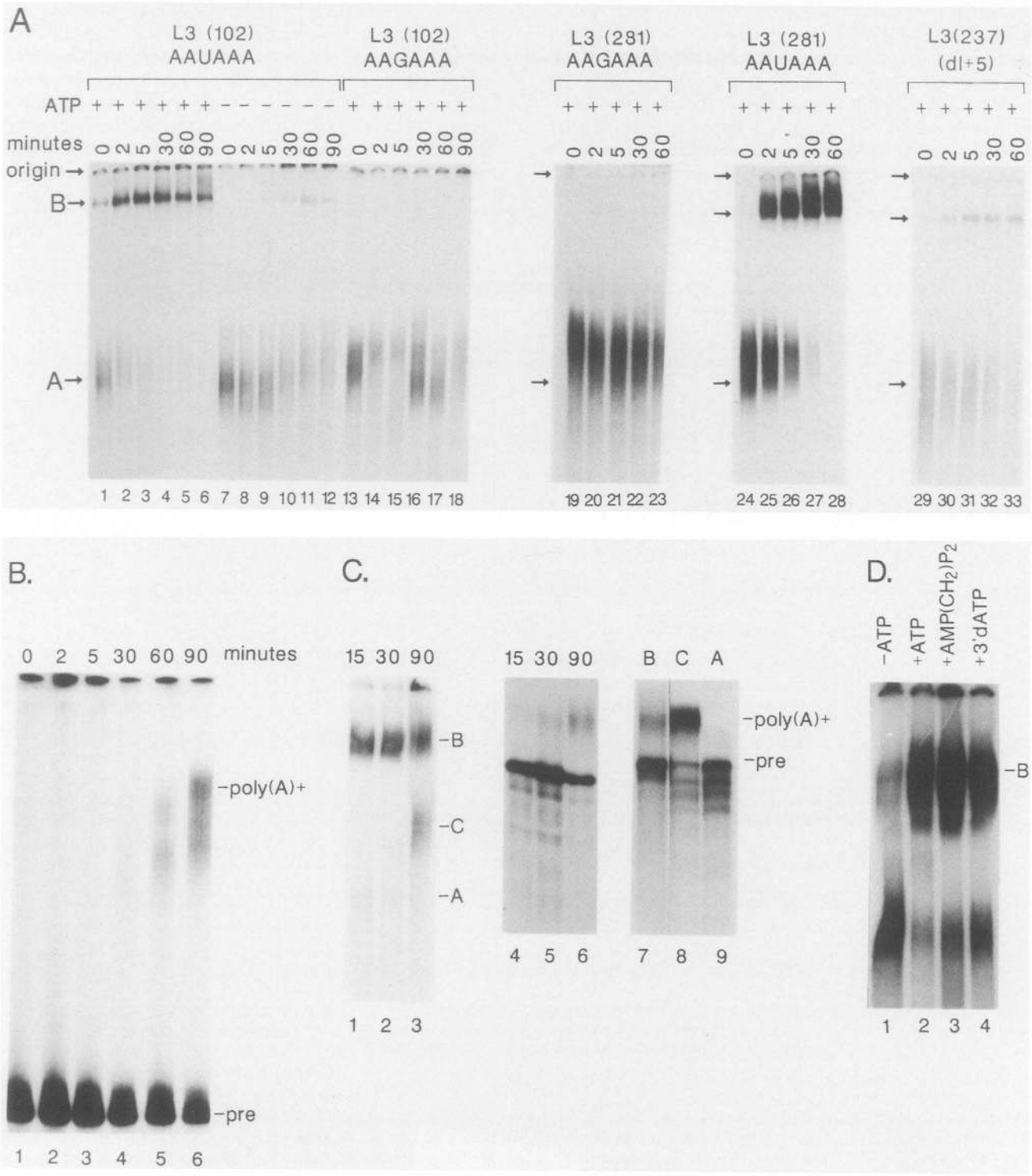


Figure 4. (*See facing page for legend.*)

or 3′dATP (cordycepin triphosphate; Fig. 4D). Both of these ATP analogs have been shown to inhibit poly(A) synthesis (Moore and Sharp 1985).

In addition, a comparison of the kinetics of formation of complex B and accurately polyadenylated RNA shows that complex B forms before the synthesis of poly(A)$^+$ RNA (Fig. 4B). Because the nuclear extract was preincubated at 30°C before initiating the reactions shown in Figure 4, A and B, the efficiency of the in vitro reactions was reduced compared to polyadenylation with fresh extracts (Fig. 2). In this case, polyadenylated RNA was not detected until after 30 min of incubation.

As polyadenylated RNA accumulated, a new complex, C, appeared that migrated in the region between complexes A and B. This complex was more distinct when a longer precursor RNA, containing an additional 165 nucleotides of sequences upstream of the L3 site, was used as substrate (Fig. 4C, lanes 1–3). This RNA, L3(267), has been described previously (Moore and Sharp 1985) and was assembled into complex B and polyadenylated with kinetics similar to that for the L3(102) precursor (Fig. 4C, lanes 1–6). The composition of RNAs in the different complexes, formed after 90 min of incubation, was determined by eluting RNA from the appropriate regions of the gel shown in Figure 4C and analyzing it by electrophoresis on a 15% polyacrylamide/8.3 M urea gel. As expected from the properties of complex A, it contained only untreated precursor (Fig. 4C, lane 9). Complex B contained both precursor and poly(A)$^+$ RNA in approximately the same ratio as that observed with RNA from the total reaction (Fig. 4C, lanes 6 and 7). Thus, complex B probably is an intermediate in the cleavage and polyadenylation reaction. On the other hand, complex C contained primarily poly(A)$^+$ RNA. The faster mobility of complex C compared with that of complex B suggests that some components may dissociate from the latter complex following cleavage and polyadenylation.

The sequence specificity for formation of complex B is similar to that required for accurate polyadenylation in the in vitro reaction. Mutation of the conserved hexanucleotide AAUAAA to AAGAAA resulted in a substrate that was inactive for both complex B formation (Fig. 4A, lanes 13–18) and polyadenylation. Interestingly, formation of complex B was also dependent upon the presence of specific sequences downstream of the poly(A) site. RNA substrate from the L3(dl + 5) plasmid, which lacked sequences from +5 to +48, was much less efficient for formation of complex B than the normal precursor (Fig. 4A; cf. lanes 24–28 with lanes 29–33). The small amount of complex B formed with this RNA substrate was roughly comparable to the small amount of correctly cleaved and polyadenylated product RNA detected in this reaction (data not shown). These observations suggest that the cellular components involved in complex B formation recognize both the conserved hexanucleotide sequence upstream of the poly(A) site and the important downstream sequences.

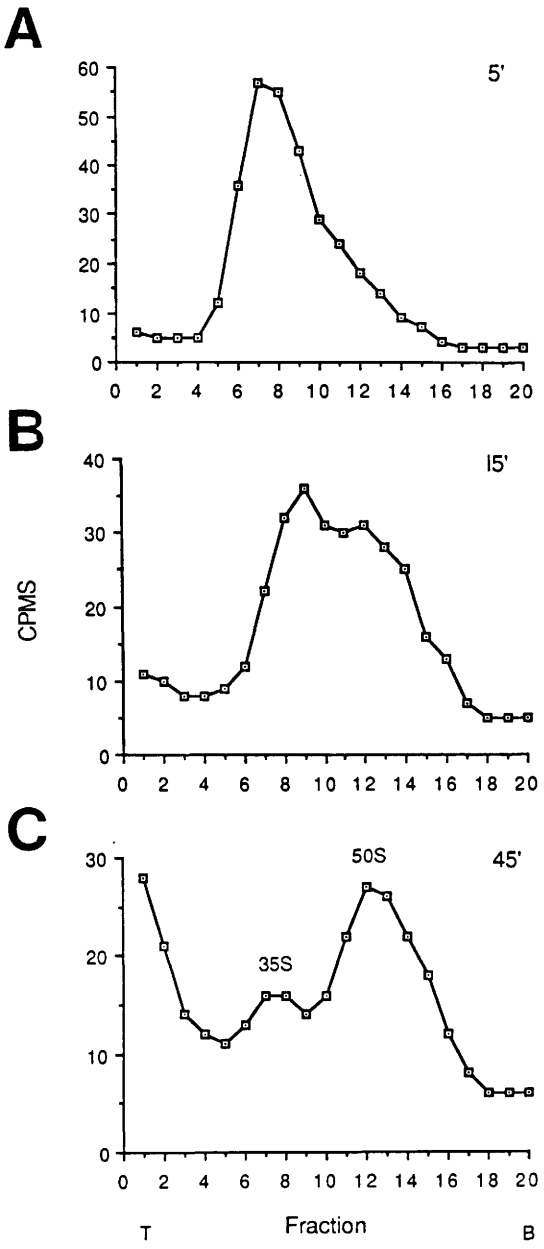
Comparison of complexes resolved by sedimentation and gel electrophoresis

Polyadenylation complexes have also been analyzed by sedimentation through glycerol gradients (C. Moore, H. Skolnik-David, and P. Sharp, unpubl.). Two complexes were observed, which sedimented at 35S and 50S. The 35S complex formed upon addition of any substrate RNA, independent of the presence or absence of sequences specific for polyadenylation. This complex also formed in the absence of ATP and is probably the result of sequence-nonspecific binding of proteins to RNA. The second complex appeared only upon incubation with ATP and sedimented at 50S. Formation of this complex was specific for the presence of both upstream and downstream polyadenylation sequences in the substrate RNA. The sequence specificity of formation of these complexes suggested that the 35S and 50S complexes would correlate to complexes A and B, respectively, as resolved by native gel electrophoresis. To test this possibility, polyadenylation reactions were stopped after different periods of incubation and resolved by sedimenta-

Figure 4. (*A*) The formation of polyadenylation-specific complexes requires ATP and authentic polyadenylation signal sequences. [α-^{32}P]UTP-labeled RNA substrates were incubated for different periods of time with HeLa nuclear extract in the presence or absence of ATP. After addition of heparin (5 mg/ml) for 10 min on ice, aliquots of the reaction mixtures were loaded on a nondenaturing 4% polyacrylamide gel, and the complexes separated by electrophoresis. (Lanes *1–6*) L3(102)AAUAAA, plus ATP; (lanes *7–12*) L3(102)AAUAAA, minus ATP; (lanes *13–18*) L3(102)AAGAAA, plus ATP; (lanes *19–23*) L3(281)AAGAAA, plus ATP; (lanes *24–28*) L3(281)AAUAAA, plus ATP; (lanes *29–33*) L3(dl + 5), plus ATP. The origin refers to the top of the gel. Two complexes were detected and designated A and B. (*B*) The assembly of polyadenylation-specific complexes precedes processing of precursor RNA. RNA was purified from aliquots of the same reaction mixtures described in *A* (lanes *1–6*). These reactions contained the normal L3(102)AAUAAA precursor. RNA was analyzed by electrophoresis on 10% polyacrylamide/8.3 M urea gels. (*C*) Analysis of RNA species in polyadenylation-specific complexes. L3(267) precursor (Moore and Sharp 1985) was incubated with nuclear extract and ATP at 30°C for the times indicated. The reaction mixtures were treated with heparin (5 mg/ml) and electrophoresed on a native 4% polyacrylamide gel (lanes *1–3*). Purified RNA from each time point was analyzed by electrophoresis on a 15% polyacrylamide/8.3 M urea gel (lanes *4–6*). The RNAs in the different complexes of the 90-min reaction, shown in lane *3* (indicated as A, B, and C), were transferred from the gel to DEAE-membrane by electroblotting and eluted as described by Pikielny and Rosbash (1986). This RNA was then analyzed by electrophoresis on a 15% polyacrylamide/8.3 M urea gel. (Lane *7*) RNA from complex B; (lane *8*) RNA eluted from the region containing complex C; (lane *9*) RNA from complex A. (*D*) Polyadenylation-specific complexes form in the presence of ATP analogs that block polyadenylation. L3(281) precursor RNA was incubated with nuclear extract and ATP or ATP analogs, as indicated below for 30 min at 30°C. Aliquots were treated with heparin, and complexes resolved by electrophoresis on native 4% polyacrylamide gels. (Lane *1*) Reaction in the absence of ATP; (lane *2*) reaction containing 1 mM ATP; (lane *3*) reaction containing 5 mM α,β-methylene-adenosine 5′ triphosphate [AMP(CH_2)P_2]; (lane *4*) reaction containing 1 mM 3′dATP.

tion through glycerol gradients (Fig. 5A–C). An aliquot representing the total reaction, as well as aliquots of fractions from the 35S or 50S region of the gradient, was treated with heparin and analyzed by gel electrophoresis (Fig. 5D). After 45 min of incubation, most of the precursor RNA from the total reaction was present in form B (Fig. 5D, lane 3). At this time, most of the RNA, as analyzed by sedimentation, was present in the 50S complex (Fig. 5C). Furthermore, aliquots from the 50S region of the gradient yielded material that migrated on the gels with a mobility similar to complex B (Fig. 5D, lanes 8–10). In contrast, aliquots from the 35S peak of the gradients yielded material that migrated more rapidly than complex B and slightly slower than complex A from the total reaction (Fig. 5D, lanes 5–7). The 35S complex probably corresponds to a slightly altered form of complex A.

Thus, the complex sedimenting at 50S in glycerol gradients generated complex B when analyzed by native gel electrophoresis after treatment with heparin. To test the effect of heparin treatment on the sedimentation velocity of polyadenylation complexes, reactions were treated with 5 mg/ml of this polyanion prior to centrifugation in a standard glycerol gradient. An equivalent reaction was analyzed in parallel but without addition of heparin. Surprisingly, all the labeled precursor RNA sedimented between 15S and 20S after treatment with heparin (Fig. 6A). This suggests that either the conformation or the composition of both the 35S and 50S complexes was altered by addition of the polyanion. Complexes A and B, as resolved by gel electrophoresis, should have survived the heparin treatment. To determine the sedimentation velocity of these complexes, a heparin-treated reaction was centrifuged for a longer time (Fig. 6B), and fractions were directly analyzed by gel electrophoresis (Fig. 6C). Complex B sedimented slightly faster than complex A (22S, as compared to 12S, respectively).



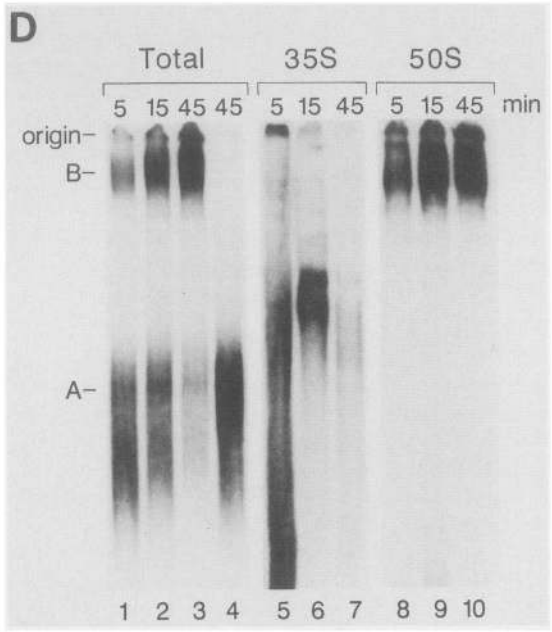


Figure 5. Polyadenylation complexes compared by sedimentation analysis through a glycerol gradient and electrophoretic separation on a nondenaturing gel. ^{32}P-Labeled RNA L3(281) was incubated with nuclear extract and ATP for the indicated times. An aliquot was removed from each reaction to an ice bath and designated "total." The remaining mixture was separated on a 10–30% glycerol gradient. The "total" aliquot, as well as equal-volume aliquots of gradient fractions from the 35S and 50S regions (C. Moore, H. Skolnik-David, and P. Sharp, unpubl.) were treated with heparin at 5 mg/ml for 10 min at 0°C and electrophoresed on a nondenaturing 4% polyacrylamide gel. (*A–C*) Total profiles (cpm) of gradients. The time of incubation is given in the upper right corner of each graph. (cpm values should be multiplied by 1000.) (*D*) Autoradiograph of nondenaturing gel showing the types of complex at different times in the total reaction (lanes *1–3*); fractions 7, representing the 35S peak (lanes *5–7*), and fractions 12, corresponding to the 50S peak (lanes *8–10*). (Lane *4*) Gel complex formed on substrate RNA L3(281)G in which AAUAAA has been changed to AAGAAA.

However, under these conditions, the two complexes were not resolved into two distinct peaks. For comparison, free L3(267) RNA substrate would only sediment into the first few fractions of the gradient.

Discussion

Polyadenylation of precursor RNA containing the L3 site of adenovirus required both an intact AAUAAA element and a sequence located between 5 and 48 nucleotides downstream of the L3 site. Mutation of the conserved hexanucleotide sequence 20 bp upstream of the L3 site to AAGAAA completely inactivated polyadenylation in vitro. Sequences downstream of polyadenylation sites have previously been shown to be important for activity (Birnsteil et al. 1985). Many, but not all, poly(A) sites have a GU-rich sequence in this position, and specific deletion of this sequence reduces the efficiency of processing (McLauchlan et al. 1985). Other types of sequence elements have also been identified by mutational analysis. In particular, a U-rich tract (UA/GUUUU) has been shown to be important for polyadenylation of the E2A site of adenovirus (Hart et al. 1985a). Recently, a quantitative analysis has suggested that both GU- and U-rich tracts can contribute to the efficiency of polyadenylation of the rabbit β-globin site (Gil and Proudfoot 1987). Precursor RNA that lacked sequences from 5 to 48 nucleotides downstream of the L3 site was neither cleaved nor polyadenylated in vitro. Assuming that this effect is due to removal of an essential element in this region, the most likely candidate is the sequence UUAUUU, located 13 nucleotides downstream of the L3 site. A similar sequence $(U)_n(A)_p(U)_q$ $(n,p,q \geqslant 1)$ is common to many of the adenovirus polyadenylation sites (LeMoullec et al. 1983).

The conserved hexanucleotide sequence AAUAAA is essential for both poly(A) synthesis and cleavage at the L3 site. These two reactions can be separated conveniently in the L3 case by the use of a truncated substrate RNA that terminates 1 nucleotide upstream of the normal polyadenylation site (Moore et al. 1986). Addition of the truncated RNA substrate to a standard in vitro reaction resulted in efficient and rapid polyadenylation. This reaction was totally dependent upon the conserved AAUAAA sequence, as it did not occur when this sequence was altered to AAGAAA. A similar observation has been made by Zarkower et al. (1986) for the late polyadenylation site of SV40 and suggests that the relevant poly(A) polymerase is part of a complex entity that recognizes a sequence 10–30 nucleotides upstream of the site of polymerization. The permissible length between the AAUAAA sequence and the 3′ terminus of

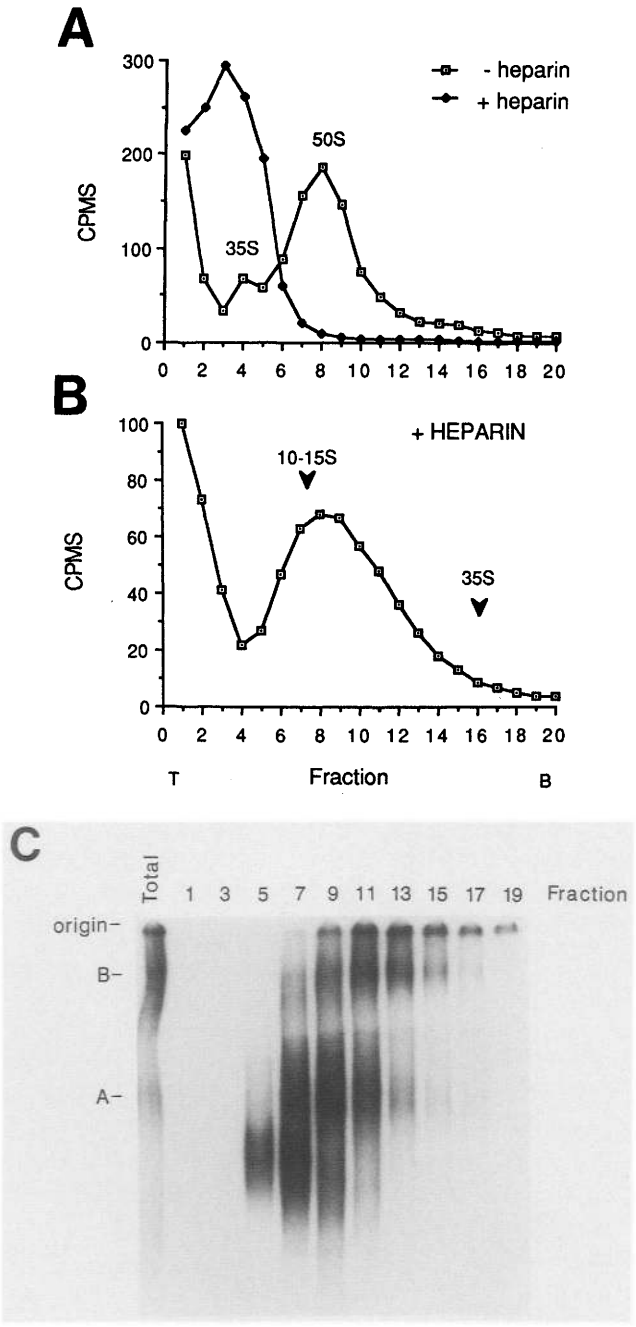


Figure 6. Effect of heparin on the polyadenylation complexes detected by sedimentation analysis. (*A*) cpm profiles of gradients containing complexes treated with or without heparin. ^{32}P-Labeled precursor RNA was incubated with nuclear extract and ATP for 45 min. Half of the reaction mixture was treated with heparin (5 mg/ml) for 10 min at 0°C. The samples were sedimented on parallel 10–30% glycerol gradients. Centrifugation was in an SW50.1 rotor for 2 hr at 49,000 rpm. cpm values on *A* and *B* should be multiplied by 1000. (□) Minus heparin treatment; (■) plus heparin treatment. (*B*) Heparin-treated complexes sedimented at the same conditions as *A*, except that the centrifugation time was twice as long. The 35S complex of a non-heparin-treated sample run on a parallel gradient sediments near the bottom, peaking at fraction 16. snRNP, 10S–15S in size, in the non-heparin-treated gradient peak in fraction 7. (*C*) Autoradiograph of a nondenaturing 4% polyacrylamide gel showing the distribution of ^{32}P-labeled RNA complexes in the gradient represented in *B*. A portion of each odd-numbered gradient fraction was resolved directly without further treatment on a gel in parallel with the total (T) polyadenylation reaction. The lane numbers correspond directly to the fraction numbers.

the RNA is probably limited because long nonprocessed substrate RNAs containing the hexanucleotide element are not efficiently polyadenylated (Manley 1983; Manley et al. 1985). For example, an RNA substrate from the +5 to +48 deletion mutant was neither processed nor extended by poly(A) synthesis, even though it contained an intact AAUAAA sequence. Thus, the 200-nucleotide sequence of RNA between the 3′ terminus and the conserved hexanucleotide motif in this substrate is probably too long for efficient poly(A) synthesis.

Further analysis of the in vitro polyadenylation reaction requires a rapid and high-resolution method for resolving complexes involved in this process. Electrophoresis of complexes in native gels has been important in the analysis of the splicing process (Konarska and Sharp 1986; Pikielny and Rosbash 1986). Resolution of polyadenylation reactions by electrophoresis, under conditions involving treatment with the polyanion heparin, has revealed the presence of three complexes. Immediately upon addition to a reaction containing nuclear extract, substrate RNA is organized into a heterogeneous complex, A. This complex forms in the absence of ATP and on substrate RNA lacking polyadenylation-specific sequences. A similar complex has been described during analysis of splicing reactions, and it is probably due to sequence-nonspecific interactions of proteins with precursor RNA. A slower migrating complex, B, assembles when precursor is incubated under conditions favorable for polyadenylation, that is, in the presence of ATP and at 30°C. Complex B forms within 5 min of incubation and, by 15–30 min, before significant polyadenylation has occurred, most of the RNA migrates in the position of this complex. As is the case for the cleavage/poly(A) synthesis reaction, formation of complex B requires both the conserved AAUAAA sequence and a downstream sequence element. Complex B contains both unreacted precursor and polyadenylated RNA. Thus, the kinetics of formation, sequence specificity, and RNA composition suggest that complex B is intimately related to polyadenylation. As the polyadenylation reaction progresses, a new complex, C, appears and contains primarily polyadenylated RNA. The faster migration of complex C compared to complex B suggests that this complex has probably lost components present in complex B.

Complexes formed with substrates active in polyadenylation have also been detected by immunoprecipitation with antisera specific for snRNP (Hashimoto and Steitz 1986), by sedimentation in glycerol gradients (C. Moore, H. Skolnik-David, and P. Sharp, unpubl.), and by protection of substrate RNA sequences against degradation by a complementary oligodeoxynucleotide and RNase H (Zarkower and Wickens 1987). After 10 min of incubation at 30°C, Hashimoto and Steitz (1986) found a fraction of precursor RNA sequences encompassing the AAUAAA element complexed with a component immunoprecipitable with either a Sm-specific serum or a serum specific for trimethylated cap of snRNAs. The identity of the snRNP particle responsible for this binding was not determined, though it is probably neither U1 nor U2 snRNP. Furthermore, Berget and Robberson (1986) have shown that specific degradation of either U1, U2, U4, or U5 snRNAs inactivates splicing but not polyadenylation. Some have questioned whether any snRNPs are involved in polyadenylation, given the persistence of this activity after digestion of extracts by micrococcal nuclease (Ryner and Manley 1987).

A complex that sediments at 50S in glycerol gradients has been resolved from a polyadenylation reaction containing the L3 substrate RNA (C. Moore, H. Skolnik-David, and P. Sharp, unpubl.). Formation of this complex requires both the conserved hexanucleotide sequence and downstream sequences. Treatment of gradient fractions from the 50S region with heparin yields a complex that comigrates during electrophoresis with complex B. Both composition and conformation probably differ between the 50S gradient complex and complex B. One indication of this is that treatment of polyadenylation reactions with heparin before gradient analysis shifts the fraction that generates complex B to the 22S region. Heparin-treated splicing complexes containing one or more snRNPs have been analyzed previously by both gel electrophoresis and sedimentation (Konarska and Sharp 1986, 1987). Comparison of the mobilities of these splicing complexes with complex B material is consistent with the possibility that complex B could contain one or more snRNP.

The accessibility of precursor RNA sequences to cleavage by digestion with RNase H and an oligodeoxynucleotide complementary to sequences containing AAUAAA have been used recently as an assay for polyadenylation complexes (Zarkower and Wickens 1987). Surprisingly, complexes that protected the AAUAAA region were only observed when the poly(A) synthesis reaction was blocked by addition of 3′ dATP. The protected RNA was accurately processed, suggesting that a complex only remained associated with the RNA when poly(A) polymerization was terminated prematurely by incorporation of the 3′ dAMP. Protection was not observed when substrate RNA was challenged at early periods of incubation, conditions where a majority of the L3 substrate would be assembled into complex B. Whether this suggests that the AAUAAA region is accessible in complex B-type structures or that the reaction conditions and substrate [poly(A) site of SV40 late mRNA] used in these previous studies are significantly different from the ones used here remains to be determined. It is also possible that multiple types of complexes form on the substrate RNA during polyadenylation and that complex B represents only one of these structures.

Materials and methods

Plasmid construction

Plasmid pSPL3S was described previously (Moore et al. 1986). The adenovirus sequence from this recombinant was subcloned into a Bluescript M13(−) plasmid (Stratagene). In this construct, pT3L3, the L3 poly(A) site was positioned downstream of the T3 bacteriophage RNA polymerase promoter. Site-directed mutagenesis was performed on single-stranded DNA, derived from pT3L3 according to the gapped duplex method (Fritz

1985) to give pT3L3G and pL3UUGUU. The synthetic oligonucleotide primers were a gift from Dr. K.L. Ramachandran (Biogen). pL3(dl + 5) was generated by cleavage of pSPL3S with *Dra*I and *Ava*I, and the single-stranded tails produced by *Ava*I digestion were degraded by S1 nuclease treatment. The large fragment was isolated by agarose gel electrophoresis and treated with ligase to give a plasmid with a deletion between +5 and +48 downstream of the L3 site. All of the mutations were confirmed by sequencing.

Polyadenylation reactions and analysis of products

For synthesis of RNA substrate, pT3L3, pT3L3G, and pL3UUGUU were digested with either *Rsa*I, *Dra*I, or *Eco*RI to give substrate RNA 53, 102, or 281 nucleotides in length, respectively. pSPL3(dl + 5) was digested with *Eco*RI. Runoff RNAs were generated in vitro with SP6 or T3 RNA polymerase in the presence of the dinucleotide primer G(5′)ppp(5′)G (Konarska et al. 1984) and labeled by incorporation of [α-^{32}P]-UTP. In vitro polyadenylation reactions were performed in a total volume of 25 μl and contained substrate RNA, 1 mM ATP, 20 mM creatine phosphate, 44 mM KCl, 0.7 mM $MgCl_2$, 8.8% glycerol, 8.8 mM HEPES (pH 7.6), 0.1 mM EDTA, 0.2 mM dithiothreitol (DTT), and 32% (vol/vol) HeLa nuclear extract, as described previously (Moore and Sharp 1985). RNAs from reaction mixtures were purified and separated on 10% or 15% polyacrylamide/8.3 M urea gels. Analysis by hybridization to the M13 cDNA clone was performed as described previously (Moore and Sharp 1985).

Analysis of complex formation

To consume endogenous ATP, nuclear extracts were preincubated at 30°C for 30 min. Precursor RNA was then added to the extract in the presence of 1 mM ATP and 20 mM creatine phosphate at 30°C. For time courses, reactions were started at staggered times so that all of the incubations were complete at the same time. At the end of the incubation time, the samples were treated with heparin (5 mg/ml) for 10 min on ice and immediately loaded on 1.3-mm thick, 4% polyacrylamide gels (acrylamide/bisacrylamide weight ratio of 80 : 1), containing 45 mM Tris–borate and 1.2 mM EDTA (pH 8.3). The gel was preelectrophoresed for 30 min at 18 V/cm. Electrophoresis was carried out at the same voltage gradient for 4–5 hr (until xylene cyanol dye had migrated 15–16 cm) at 4°C. The gel was transferred to Whatman 3MM paper, dried, and autoradiographed at −70°C with a screen. To examine the RNA composition of different regions of the gel, RNAs were transferred to DEAE-membrane by electroblotting and eluted, as described by Pikielny and Rosbash (1986).

For sedimentation analysis, a 125-μl polyadenylation reaction was loaded onto a 5-ml gradient containing 10–30% glycerol in 20 mM HEPES (pH 7.5), 1 mM $MgCl_2$, 25 mM KCl, and 0.1 mM EDTA (Grabowski et al. 1985). Gradients were centrifuged in a SW50.1 rotor at 49,000 rpm, 5°C, for 2–3 hr. Twenty 250-μl fractions were collected and counted to determine the ^{32}P cpm Cerenkov profile of the gradient.

Acknowledgments

H.S.-D. acknowledges postdoctoral fellowship support from EMBO, and C.L.M., from the Life Sciences Research Foundation. This work was supported by grant DMB-8606757 from the National Science Foundation (to C.L.M.) and by Public Health Service grants GM32467 and PO1-CA42063 from the National Institutes of Health, grant DCB-8502718 from the National Science Foundation, and, partially, a National Cancer Institute Cancer Center core grant P30-CA14051 (to P.A.S.). In addition, we are grateful to M. Siafaca for secretarial help, to A. Gil, P. Grabowski, M. Konarska, S. Berget, and A. Virtanen for much helpful discussion and advice, and to John Whoriskey for technical assistance.

References

Berget, S.M. and B.L. Robberson. 1986. U1, U2, and U4/6 small nuclear ribonucleoproteins are required for splicing but not polyadenylation. *Cell* **46:** 691–696.

Birnstiel, M.L., M. Busslinger, and K. Strub. 1985. Transcription termination and 3′ processing: The end is in site! *Cell* **41:** 349–359.

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.

Fritz, H.J. 1985. The oligonucleotide-directed construction of mutations in recombinant filamentous phage. In DNA cloning (ed. D.M. Glover), pp 151–171. IRL Press, Oxford.

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.

Grabowski, P., S.R. Seiler, and P.A. Sharp. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell* **42:** 345–353.

Hart, R.P., M.A. McDevitt, H. Ali, and J.R. Nevins. 1985a. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and the early simian virus 40 polyadenylation sites. *Mol. Cell. Biol.* **5:** 2975–2983.

Hart, R.P., M.A. McDevitt, and J.R. Nevins. 1985b. Poly(A) site cleavage in a HeLa nuclear extract is dependent on downstream sequences. *Cell* **43:** 677–683.

Hashimoto, C. and J.A. Steitz. 1986. A small nuclear ribonucleoprotein associates with the AAUAAA polyadenylation signal *in vitro*. *Cell* **45:** 581–591.

Konarska, M.M. and P.A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46:** 845–855.

———.1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* **49:** 763–774.

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.

LeMoullec, S.M., G. Akusjarvi, P. Stalhardske, U. Pettersson, B. Chambraud, P. Gilardi, M. Nasri, and M. Perricaudet. 1983. Polyadenylic acid addition sites in the adenovirus type 2 major late transcriptional unit. *J. Virol.* **48:** 122–134.

Manley, J.L. 1983. Accurate and specific polyadenylation of mRNA precursors in a soluble whole-cell lysate. *Cell* **33:** 595–605.

Manley, J.L., H. Yu, and L. Ryner. 1985. RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation *in vitro*. *Mol. Cell. Biol.* **5:** 373–379.

McLauchlan, J., D. Gaffney, J.C. Whitton, and J.B. Clements. 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. *Nucleic Acids Res.* **13:** 1347–1368.

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* **305:** 600–605.

Moore, C.L. and P.A. Sharp. 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. *Cell* **41:** 845–855.

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.

Pikielny, C.W. and M. Rosbash. 1986. Specific nuclear RNAs are associated with yeast spliceosomes. *Cell* **45:** 869–877.

Ryner, L.C. and J.L. Manley. 1987. Requirements 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.

Sperry, A.O. and S.M. Berget. 1986. *In vitro* cleavage of the simian virus 40 early polyadenylation site adjacent to a required TG sequence. *Mol. Cell. Biol.* **6:** 4734–4741.

Wickens, M. and P. Stephenson. 1984. Role of the conserved AAUAAA sequence: Four AAUAAA point mutations prevent messenger RNA 3′ end formation. *Science* **226:** 1045–1051.

Zarkower, D. and M. Wickens. 1987. Formation of mRNA 3′ termini: Stability and dissociation of a complex involving the AAUAAA sequence. *EMBO J.* **6:** 177–186.

Zarkower, D., P. Stephenson, M. Sheets, and M. Wickens. 1986. The AAUAAA sequence is required both for cleavage and for polyadenylation of SV40 pre-mRNA in vitro. *Mol. Cell. Biol.* **6:** 2317–2323.

Electrophoretic separation of polyadenylation-specific complexes.

H Skolnik-David, C L Moore and P A Sharp

Genes Dev. 1987, **1:**
Access the most recent version at doi:10.1101/gad.1.7.672