The EMBO Journal vol.7 no.10 pp.3159–3169, 1988

Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding

Claire L.Moore, Jie Chen and John Whoriskey

Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, MA 02111, USA

Communicated by W.Keller

The major proteins crosslinked by UV light to RNA containing the adenovirus-2 L3 poly(A) site are species of 155, 68 and 38 kd mol. wt (p155, p68 and p38). Mutation of AAUAAA to AAGAAA prevented crosslinking of the two larger proteins and destroyed the ability of the RNA to compete for binding of these proteins. However, association of p155 and p68 with precursor was unaffected by deletion of sequences downstream of the poly(A) site critical for *in vitro* polyadenylation. These two proteins are in the polyadenylation-specific, but not the nonspecific complexes detected by electrophoresis in nondenaturing gels. In addition, p155 and p68 are not found on RNA which has been processed. p155 bound a 15-nt oligomer containing AAUAAA, and thus does not require extended RNA sequence for interaction with RNA. Identified by immunoprecipitation with specific antibody, p38 is the C protein of heterogeneous ribonucleoprotein particles (hnRNPs). While p155 has an Sm epitope, it is not associated with snRNPs containing trimethylated guanosine caps.

Key words: mRNA/polyadenylation

Introduction

The ability to examine the polyadenylation of eukaryotic mRNA precursor *in vitro* has greatly enhanced our understanding of the mechanism of this essential processing event. Formation of the mature mRNA 3′ end requires endonucleolytic cleavage of a larger precursor followed by the addition of ~200 adenylate residues to the 3′ end. Like the *in vivo* reaction, proper polyadenylation *in vitro* requires the hexanucleotide sequence, AAUAAA, as well as downstream elements (Hart *et al.*, 1985b; Manley *et al.*, 1985; Sperry and Berget, 1986; Zarkower, *et al.*, 1986; Humphrey *et al.*, 1987; Ryner and Manley, 1987; Skolnik-David *et al.*, 1987; Zhang and Cole, 1987; Green and Hart, 1988; Zarkower and Wickens, 1988). From these and other studies, it is clear that *trans*-acting factors in HeLa nuclear extract are required for the processing event. These must recognize the signal sequences on precursor RNA as well as provide the enzymatic activities of cleavage and polyadenylation. When multiple polyadenylation sites exist on the same precursor RNA, the relative affinity of each site for such factors may influence which site is used. Several experiments have indicated that extract factors are interacting with the AAUAAA sequence and the downstream elements during the polyadenylation reaction. First, the AAUAAA region is protected from ribonuclease digestion after precursor RNA has been incubated with extract (Hashimoto and Steitz, 1986; Humphrey *et al.*, 1987; Stefano and Adams, 1988). In some cases, a large ribonuclease-resistant fragment extending from the AAUAAA to the downstream element can be demonstrated. Secondly, this hexamer is not accessible for basepairing with a complementary oligonucleotide when the polyadenylation reaction is blocked by 3′dATP (Zarkower and Wickens, 1987a) or when isolated polyadenylation-specific complexes are probed with a similar oligonucleotide (Stefano and Adams, 1988).

The observation of a specific complex which contains precursor RNA, proteins and possibly small nuclear RNAs provided additional evidence that extract factors interact with the RNA (Humphrey *et al.*, 1987; Moore *et al.*, 1988; Skolnik-David *et al.*, 1987; Zarkower and Wickens, 1987b; Zhang and Cole, 1987). This complex assembles onto precursor RNA very rapidly after incubation with HeLa nuclear extract. Because the polyadenylation-specific complex requires both the AAUAAA and a downstream element for efficient formation, it is likely that it contains factors which interact with these sequences. Although the majority of precursor is associated with the complex, processing activities have not yet been identified as components. Partial fractionation of processing activities from crude nuclear extracts has been recently reported (Takagaki *et al.*, 1988). However, the nature of the factors responsible for polyadenylation remains undefined.

One approach which has been helpful in the study of RNA–protein interactions is the use of UV light to covalently crosslink RNA to proteins in close contact with the nucleic acid. This technique has been used to localize the binding sites of *Escherichia coli* aminoacyl tRNA synthetase on tRNA (Schoenmaker and Schimmel, 1974), and of purified ribosomal proteins on *E.coli* rRNA (Ehresmann *et al.*, 1977). Crosslinking with UV irradiation also demonstrated highly specific RNA–protein interactions which occur in complex particles such as ribosomes (Moller and Brimacombe, 1975; Rinke *et al.*, 1976) or retroviruses (Sen and Todaro, 1977). This technique has also provided a rigorous definition of proteins associated with eukaryotic mRNA precursor (Dreyfuss, 1986; Chung and Wooley, 1986), and of proteins which bind to the cap structure of eukaryotic mRNA (Pelletier and Sonenberg, 1985; Rozen and Sonenberg, 1987).

We have used UV-induced RNA–protein crosslinking to reveal extract proteins which specifically bound to polyadenylation precursor RNA. In this report we demonstrate that the major proteins readily crosslinked to substrate RNA with the adenovirus L3 polyadenylation site are species of 155, 68 and 38 kd mol. wt. The 68-kd protein probably corresponds to a similar-sized protein recently described by Wilusz and Shenk (1988). In addition, we have further investigated the role of these proteins in the polyadenylation process, and further characterized them by immunoprecipitation with specific antibodies.

©IRL Press Limited, Oxford, England

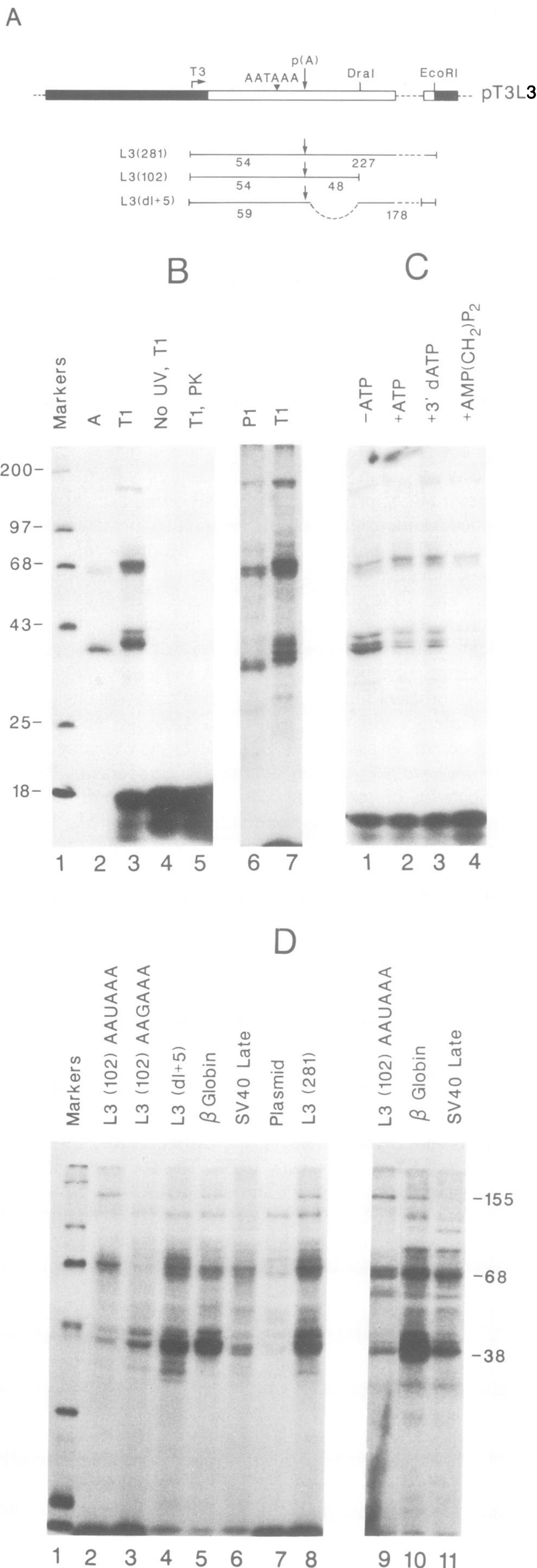


Fig. 1. Proteins crosslinked to polyadenylation precursor by UV light. (**A**) RNA substrates containing the adenovirus-2 L3 polyadenylation site. (**B**) Effects of digestion with different ribonucleases. ^{32}P-labeled L3(102) RNA was incubated with nuclear extract for 5 min at 30°C, exposed to UV irradiation for 10 min at 4°C, and digested for 30 min at 37°C with ribonuclease A (**lane 2**), ribonuclese T1 (**lanes 3** and **7**), or ribonuclease P1 (**lane 6**). **Lane 1** contains ^{14}C-labeled protein markers (the corresponding mol. wts are indicated on the left side). **Lane 4** contains a sample which was digested with ribonuclease T1 without prior UV exposure. **Lane 5** contains sample which was digested with proteinase (PK) following crosslinking and ribonuclease T1 digestion. (**C**) Effects of ATP and ATP analogs on the proteins crosslinked to L3(102) RNA. L3(102) RNA was incubated with nuclear extract in the absence of ATP and creatine phosphate (**lane 1**), or in the presence of 1 mM ATP (**lane 2**), 1 mM 3′dATP (**lane 3**) or 5 mM AMP(CH_2)P_2 (**lane 4**). The reactions represented in **lanes 2–4** contained 20 mM creatine phosphate. After crosslinking, samples were digested with ribonuclease T1. (**D**) Analysis of sequence requirements for binding of polyadenylation-specific proteins. Proteins crosslinked to the indicated RNAs were detected as described for (**B**). Samples in **lanes 2–8** were digested with ribonuclease T1, **lanes 9–11** with ribonuclease P1. **Lane 1**, marker proteins; **lanes 2** and **9**, L3(102) RNA containing the normal AAUAAA sequence; **lane 3**, L3(102) RNA with AAGAAA sequence; **lane 4**, L3(d1+5) RNA; **lanes 5** and **10**, RNA containing the rabbit β-globin poly(A) site; **lanes 6** and **11**, RNA containing the SV40 late poly(A) site; **lane 7**, plasmid RNA containing only prokaryotic sequence; **lane 8**, L3(281) RNA. Samples in (**B–D**) were analyzed by electrophoresis on an SDS-containing 10% polyacrylamide gel.

Results

Radioactive substrate RNA containing the adenovirus-2 L3 polyadenylation site (Figure 1A) was used to detect proteins which can be crosslinked to this RNA by UV light. Substrate RNA was transcribed from DNA template using T3 RNA polymerase and ^{32}P-labeled nucleotide triphosphates. The precursor L3(102) was composed of 102 nucleotides of sequence which flanks the L3 polyadenylation site. This substrate RNA is accurately and efficiently polyadenylated when incubated with HeLa nuclear extract and ATP (Skolnik-David *et al.*, 1987). Two substrate RNAs from plasmids containing mutations were also tested. These were synthesized from DNA templates in which the AATAAA sequence 20 bases upstream of the L3 site had been altered to AAGAAA or in which sequences from +5 to +48 downstream of the L3 site had been deleted. These substrates are almost completely inactive for polyadenylation *in vitro* (Skolnik-David *et al.*, 1987).

Specific proteins crosslinked to precursor RNA by UV light

Radioactive precursor RNA was incubated with HeLa nuclear extract and ATP at 30°C for 5 min. The reaction mixture was then exposed to UV light for 10 min at 4°C. Following this treatment, the RNA was extensively digested with ribonuclease, boiled in SDS and mercaptoethenol, and the sample subjected to electrophoresis in a discontinuous polyacrylamide gel containing SDS. Proteins crosslinked to precursor RNA can be detected by the radioactively labeled nucleotides which remain covalently attached to the proteins after the ribonuclease digestion. Non-crosslinked proteins are removed from the RNA by the denaturing conditions and are not visualized by the autoradiography. Upon treatment with ribonucleases A, T1 or P1, three prominent bands at ~155, 68 and 38 kd mol. wt were apparent (Figure 1B, lanes 2, 3, 6 and 7). Ribonuclease A cleaves naked RNA after cytosine and uridine nucleotides, ribonuclease T1 after guanosine nucleotides and ribonuclease P1 at every

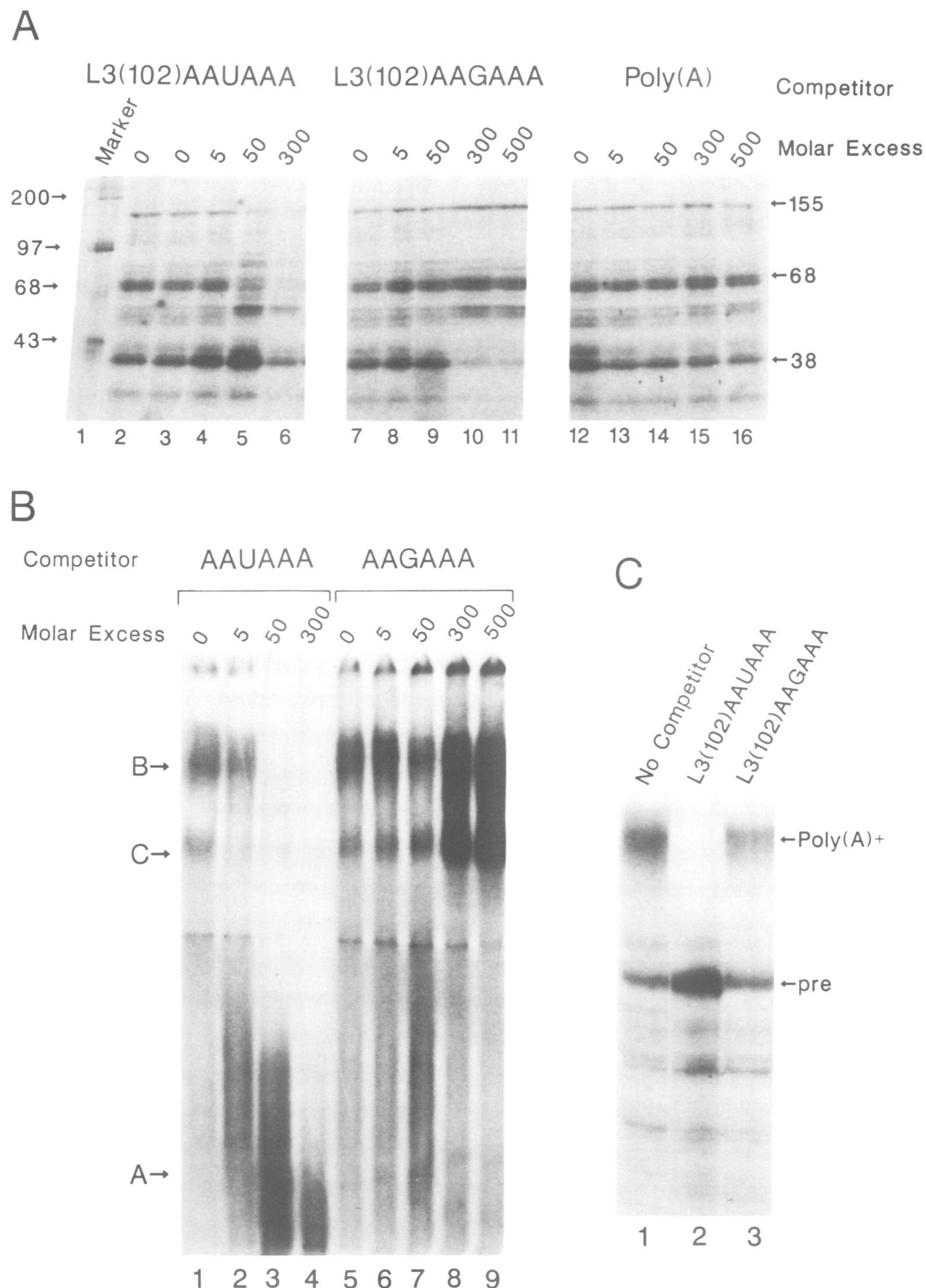


Fig. 2. The effects of excess RNA on protein crosslinking. ^{32}P-labeled L3(102)AAUAAA RNA (0.5 nM) was incubated with nuclear extract and ATP in the presence of the indicated amounts of unlabeled competitor RNA. After 5 min at 30°C, aliquots of each reaction were removed, crosslinked, and digested with both ribonucleases T1 and A, and the proteins separated on a 10% polyacrylamide–SDS gel. The remaining reaction mixture was incubated for an additional 30 min and a second aliquot analyzed for polyadenylation-specific complexes on nondenaturing 4% polyacrylamide gels. The RNA from the remainder of the reaction was purified and electrophoresed on a 15% polyacrylamide/8.3 M urea gel. (**A**) Analysis of crosslinked proteins. **Lane 1**, ^{14}C-labeled marker proteins: **lanes 2–6**, L3(102)AAUAAA as competitor; **lanes 7–11**, L3(102)AAGAAA as competitor; **lanes 12–16**, poly(A) homopolymer (average length 200 bases) as competitor. (**B**) Analysis of complex formation. **Lanes 1–4**, L3(102)AAUAAA as competitor; **lanes 5–9**, L3(102)AAGAAA as competitor. A indicates the position of nonspecific complex; B and C are polyadenylation-specific complexes. (**C**) Analysis of RNA. **Lane 1**, no competitor; **lane 2**, L3(102)AAUAAA RNA as competitor at 300-fold molar excess; **lane 3**, L3(102)AAGAAA as competitor at 300-fold excess. Poly(A)$^+$, polyadenylated RNA; pre, precursor.

nucleotide. By comparing the digestion patterns produced by the different nucleases, it is clear that the size of the covalently attached RNA fragment does have a slight effect on the mobility of a protein. Regardless of the nuclease used for digestion, the 68-kd protein often appeared as a dimer. With T1 digestion, the 38-kd protein was usually visible as a set of two to three bands. The presence of all of these bands was dependent on UV crosslinking, since no label was detected at these positions when the RNA was incubated with extract and digested with ribonuclease T1 without previous

crosslinking (Figure 1B, lane 4). This control also demonstrated that digested, uncrosslinked RNA migrated with or faster than the 18-kd protein marker. The higher bands which appeared after crosslinking represent proteins since they disappeared when crosslinking and T1 digestion is followed by treatment with proteinase (Figure 1B, lane 5).

The sequences required for binding of the different proteins were also determined. Two mutations of the L3 site prevent *in vitro* polyadenylation (Skolnik-David *et al.*, 1987). One is a point mutation in which the AAUAAA upstream of the poly(A) site is changed to AAGAAA. The major proteins crosslinked to this RNA were 38 kd in size (Figure 1D, lane 3). This mutation greatly reduced the amount of 155- and 68-kd proteins detectable with this assay. Traces of label at both positions are present upon a longer exposure of the autoradiograph (data not shown). In the second mutation, sequences from 5 to 48 nt beyond the poly(A) site were deleted. This 237-nt-long substrate, L3(dl +5), produced a labeled protein pattern similar to that of the short precursor L3(102)AAUAAA (Figure 1D, lane 4). Additional bands seen with L3(dl +5) are likely to represent nonspecific binding since they are also detected on a longer normal precursor, L3(281) (Figure 1D, lane 8) and on plasmid RNA of 230 nucleotides, which contained no eukaryotic sequence (Figure 1D, lane 7). The plasmid RNA did not crosslink to proteins of 155 and 68 kd, though a band slightly below the 68-kd protein was visible. These results show that that the 155-, 68- and 38-kd proteins are the major proteins of nuclear extract crosslinked to the L3 polyadenylation precursor. These three proteins have been designated p155, p68 and p38 respectively.

RNAs containing the rabbit β-globin poly(A) site (C.Moore and A.Gill, unpublished results) or the Simian virus 40 (SV40) late poly(A) site (Zarkower *et al.*, 1986) are efficiently cleaved and polyadenylated in HeLa nuclear extracts. The proteins which crosslinked to these precursor RNAs were also investigated. The β-globin RNA yielded a pattern very similar to that of the normal L3(281) RNA (Figure 1D, lanes 5 and 8). Most importantly, like the L3 RNAs, this RNA crosslinked to a 155-kd protein and a 68-kd protein, and showed strong binding of a protein of 38 kd. In contrast, the SV40 late RNA crosslinked to a 68-kd protein and a 38-kd protein, but not a 155-kd protein (Figure 1D, lanes 6 and 11). The same proteins were detected when P1 ribonuclease was used after crosslinking (Figure 1D, lanes 9–11).

Effects of excess precursor

The results of the previous experiment suggested that binding of p155 and p68 required AAUAAA. However, it is possible that these proteins interact with the AAGAAA substrate but cannot be crosslinked. Therefore, we tested the ability of excess RNA substrates of different types to compete for binding of these two proteins.

^{32}P-labeled precursor RNA, L3(102)AAUAAA, was incubated with nuclear extract and ATP in the presence of varying amounts of unlabeled substrates. When L3(102)AAUAAA was used as a competitor, 50-fold excess of this RNA decreased crosslinking of both p68 and p155 to the test RNA. This RNA in 300-fold excess completely blocked binding of p155 and p68 and reduced binding of p38 (Figure 2A, lanes 3–6). When L3(102)AAGAAA was used as a competitor, 500 times molar excess had no effect on the binding of p68 and p155, but did reduce the amount of p38 crosslinked to the labeled RNA (Figure 2A, lane 11). These results indicate that the mutated RNA, L3(102)-AAGAAA had little affinity for p68 and p155.

Eukaryotic cells have two proteins which bind poly(A) tracts: a 60-kd protein located in the nucleus (Setyono and Greenberg, 1981) and a 72-kd protein confined to the cytoplasm (Adam *et al.*, 1986; Sachs *et al.*, 1986; and references therein). Since the poly(A) tail formed *in vitro* was not radioactively labeled, these proteins would probably not be detected in the experiment as designed. To test whether these or other proteins associated with poly(A) somehow contributed to the configuration of proteins crosslinked to L3 precursor, poly(A) was used as a competitor. Five hundred times molar excess of this polynucleotide had no effect on the proteins crosslinked to L3(102)AAUAAA RNA (Figure 2A, lanes 12–16).

Complexes which assemble onto L3(102) precursor RNA during a polyadenylation reaction can be detected by electrophoresis of reaction mixtures through a nondenaturing 4% polyacrylamide gel (Skolnik-David *et al.*, 1987). These complexes include a nonspecific complex A, which also forms on RNA lacking polyadenylation signals, and a specific complex B, which requires intact upstream and downstream polyadenylation signals for efficient assembly. For example, mutation of the AAUAAA element to AAGAAA prevents complex B formation (Figure 5A, lanes 1 and 2). Complex B contains both unrreacted precursor RNA and polyadenylated RNA and is thought to be a processing intermediate. Processed RNA is also found in complex C, which migrates between A and B (Figure 2B, lane 1). Including excess quantities of the unlabeled wild-type precursor in the reaction prevented the formation of the specific complexes B and C and also reduced the mobility of complex A (Figure 2B, lanes 1–4). As would be expected from these results, 300-fold excess of this competitor also prevented processing of the test RNA (Figure 2C, lane 2). There was no detrimental effect on complex formation (Figure 2B, lane 8) or processing (Figure 2C, lane 3) with 300-fold molar excess of L3(102)AAGAAA as competitor.

Effects of ATP and kinetics of binding

Cleavage and poly(A) addition *in vitro* requires ATP in the reaction mixture (Moore and Sharp, 1985). However, cleavage alone occurs if analogs are substituted for ATP. These analogs include α,β-methylene adenosine 5′ triphosphate [AMP(CH_2)P_2], a nonhydrolyzable analog and cordycepin triphosphate (3′dATP), which aborts polyadenylation by the incorporation of an adenosine residue lacking a 3′ hydroxyl group. p155, p68 and p38 were crosslinked to L3(102) precursor RNA equally well in the presence or absence of ATP or its analogs (Figure 1C). The concentration of endogenous ATP in the extract is $< 1\ \mu$M, and it is unlikely that ATP is required for binding of these proteins.

In the previous experiments, precursor and extract were incubated at 30°C for only 5 min. It was of interest to know how the association of these proteins with precursor and product RNA changed during the course of the reaction. Reaction mixtures were incubated for different intervals at 30°C. The reactions were started at staggered times so that all reactions were complete at the same time. At the end of the incubation, half of each sample was analyzed for crosslinked proteins, and the other half for RNA. A full

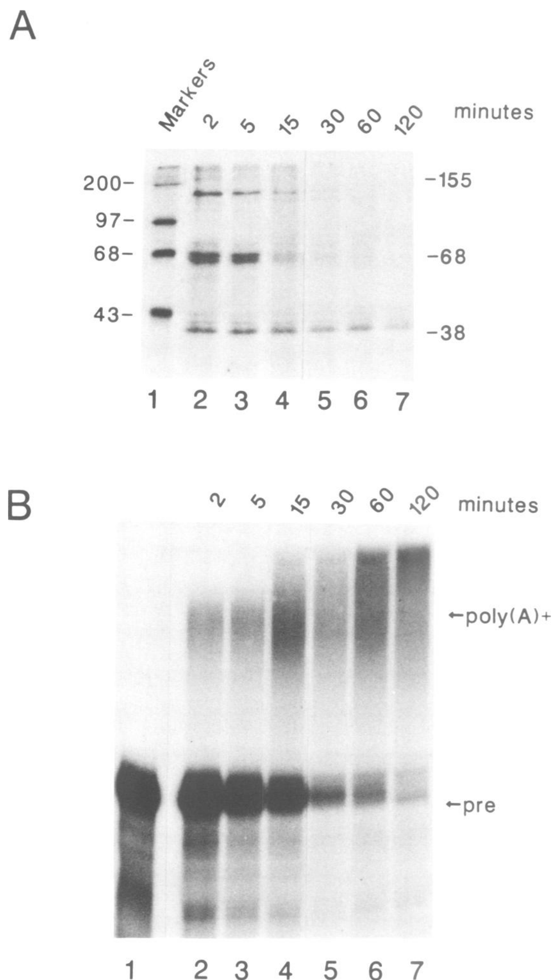


Fig. 3. Kinetics of binding of crosslinked proteins to L3 precursor RNA. L3(102) RNA was incubated with nuclear extract under polyadenylation conditions for the indicated times. Half of each reaction mixture was exposed to UV light, digested with ribonuclease T1, and the proteins separated on a 10% polyacrylamide–SDS gel (**A**). **Lane 1** of (**A**) contains ^{14}C-labeled marker proteins. The RNA in the remaining samples was purified and analyzed by electrophoresis on a 15% polyacrylamide/8.3 M urea gel (**B**). **Lane 1** of (**B**) is unreacted precursor RNA.

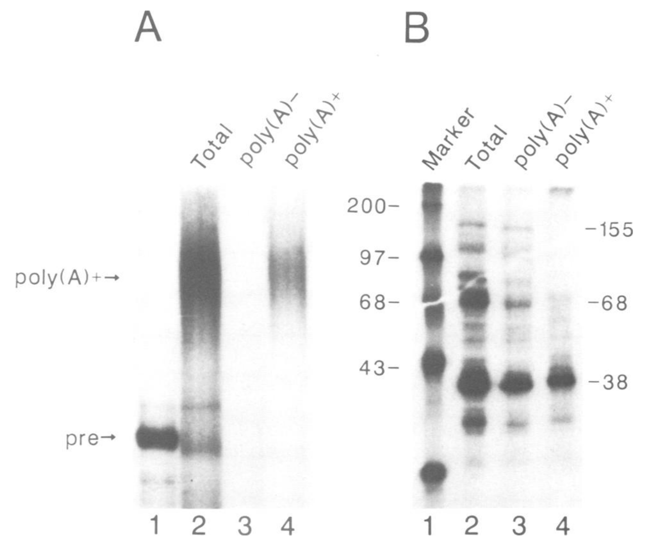


Fig. 4. Proteins crosslinked to polyadenylated RNA. L3(102) RNA was incubated with nuclear extract and ATP for 60 min at 30°C. The sample was exposed to UV light and then fractionated on poly(U) Sephadex. Aliquots of poly(A)$^-$ and poly(A)$^+$ RNAs, as well as an aliquot of the original unfractionated sample (Total), were analyzed by electrophoresis on a 15% polyacrylamide/8.3 M urea gel (**A**). The remaining aliquots of total, poly(A)$^-$, and poly(A)$^+$ RNAs were digested with ribonucleases T1 and A and the crosslinked proteins examined by electrophoresis on a 10% polyacrylamide–SDS gel (**B**). Equal amounts of radioactivity were loaded onto the protein gel.

complement of proteins was crosslinked to precursor RNA within 2 min of incubation (Figure 3A, lane 2). By 15 min, the levels of p68 and p155 had decreased, and were greatly reduced at the later time points (Figure 3A, lanes 4–7). The level of p38 was hardly affected.

At the earliest times, a small amount of RNA is polyadenylated (Figure 3B, lanes 2 and 3) and migrates more slowly than precursor. After 15 min the levels of processed RNA steadily increased (Figure 3B, lanes 4–7). Our previous studies have shown by hydridization/ribonuclease protection analysis, that this processed RNA is correctly cleaved at the internal polyadenylation site (Moore *et al.*, 1986; Skolnik-David *et al.*, 1987). In these studies, retention of the processed RNA on poly(U) Sephadex also showed that the additional length was due to formation of a poly(A) tail several hundred nucleotides long.

Proteins associated with poly(A)$^+$ RNA and with polyadenylation-specific complexes

The inverse relationship between binding of the two specific proteins, p68 and p155, and the appearance of polyadenylated product suggested that these proteins might dissociate from RNA once processing occurs. To explore this possibility further, RNA was incubated with extract for 60 min, followed by UV crosslinking. The RNA was then fractionated into poly(A)$^+$ and poly(A)$^-$ species by chromatography on poly(U) Sephadex. Aliquots examined by electrophoresis on a 15% polyacrylamide/8.3 M urea gel showed that the precursor RNA had been efficiently polyadenylated and that the fractionation was successful (Figure 4A, lanes 2–4). The remaining samples were digested with ribonucleases T1 and A, and the composition of crosslinked proteins determined by SDS–PAGE. p68 and p155 were found only on the poly(A)$^-$ RNA. p38 was found in both the poly(A)$^+$ and poly(A)$^-$ populations (Figure 4B, lanes 3 and 4). Thus, either the two specific proteins dissociate from the RNA product, or are present in a way which cannot be crosslinked.

As discussed previously, early in the *in vitro* polyadenylation reaction, two complexes form on L3(102) precursor RNA. These are a nonspecific complex, A, and a polyadenylation-specific complex, B (Figure 5A, lane 2). The proteins crosslinked to RNA in the two complexes were determined. After incubation of precursor RNA and extract for 5 min at 30°C, the mixture was crosslinked and the complexes separated on a nondenaturing 4% polyacrylamide gel. The UV light treatment did not affect the proportion of RNA in complexes A and B or the mobility of these complexes (Figure 5A, lanes 2 and 3). RNA was electroeluted from the regions of the gel containing A and B, ethanol precipitated, digested with ribonucleases T1 and A, and the proteins analyzed by gel electrophoresis. The A complex contained primarily p38 and a set of proteins between 70 and 97 kd (Figure 5B, lane 3). The polyadenylation-specific complex B contained p38 and more importantly, p68 and p155 (Figure 5B, lane 4).

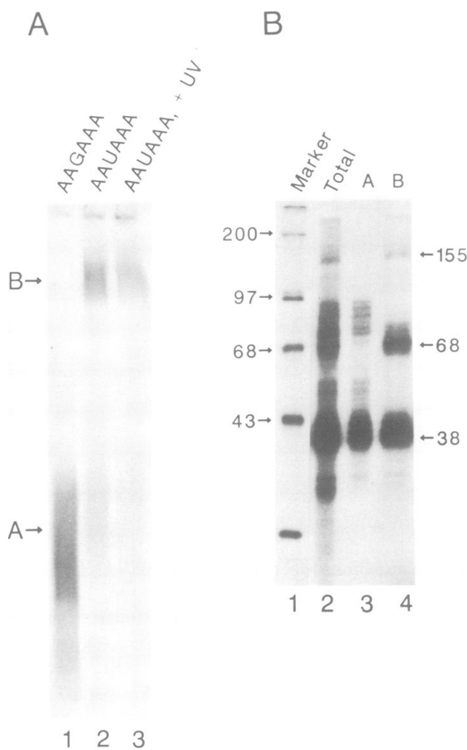
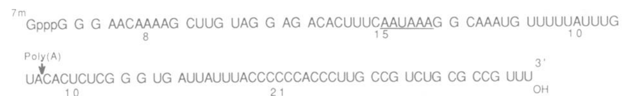


Fig. 5. Crosslinked proteins found in polyadenylation-specific complex. (A) Analysis of complexes. RNAs were incubated with nuclear extract and ATP for 15 min at 30°C and complexes in the reaction mixtures separated on a nondenaturing 4% polyacrylamide gel. **Lane 1**, complexes formed on L3(102)AAGAAA RNA; **lane 2**, complexes formed on L3(102)AAUAAA RNA; **lane 3**, complexes from a reaction containing L3(102)AAUAAA RNA which was exposed to UV light before gel electrophoresis. Complex A is non-specific complex. Complex B is the polyadenylation-specific complex. (B) Analysis of proteins in complexes A and B. Complexes were exposed to UV light and separated by gel electrophoresis. RNA was electroeluted from the regions of the gel corresponding to complexes A and B. These RNAs were digested with ribonucleases T1 and A, and the radioactively labeled proteins examined by electrophoresis on a 10% polyacryalmide−SDS gel. **Lane 1**, ^{14}C-labeled marker proteins; **lane 2**, crosslinked proteins from the reaction mixture before electrophoretic separation of complexes A and B; **lane 3**, proteins crosslinked to RNA from complex A; **lane 4**, proteins crosslinked to RNA from complex B.



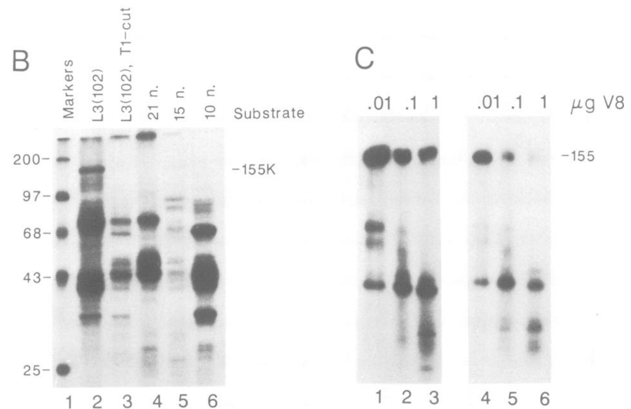


Fig. 6. Proteins crosslinked to specific RNA fragments. (A) Sequence of the L3(102) precursor RNA. Spaces indicate cleavage sites of ribonuclease T1. (B) Analysis of proteins crosslinked to oligomers derived from L3(102) RNA. Purified oligomers were incubated with nuclear extract, exposed to UV light and the crosslinked proteins detected by electrophoresis on a 10% polyacrylamide SDS-containing gel. **Lane 1**, ^{14}C-labeled marker proteins; **lane 2**, proteins crosslinked to intact precursor; **lane 3**, proteins crosslinked to precursor which had been digested with ribonuclease T1 prior to incubation with extract; **lanes 4−6**, proteins crosslinked to the indicated oligomers. (C) Comparison of the partial V8 protease digestion patterns of the 155-kd proteins. The 155-kd proteins derived from intact L3(102) precursor (**lanes 1−3**) or from the 15-nt oliigomer containing AAUAAA (**lanes 4−6**) were partially digested with the indicated amounts of V8 protease and electrophoresed on a 15% polyacrylamide−SDS gel.

Proteins crosslinked to fragments of precursor RNA

The two polyadenylation-specific proteins may not require the entire precursor for binding. Ribonuclease T1 digestion of L3(102)AAUAAA produces several large oligomers (Figure 6A). One of these, a 15-nt fragment, contains the AAUAAA sequence. Two others, a 10-nt oligomer just upstream of the cleavage site and a 21-nt oligomer downstream of this point contain sequences (UUAUU) complementary to the hexanucleotide. A second decamer spans the cleavage site. These oligonucleotides were purified by gel electrophoresis. The two decamers could not be separated by this technique. Each oligomer was then incubated with nuclear extract, the mixture was exposed to UV light, and the proteins resolved by gel electrophoresis. Each oligomer bound a characteristic set of proteins (Figure 6b) which are a subset of the proteins observed when a total T1 digest of precursor RNA is used as substrate (Figure 6B, lane 3). All of the oligomers bound proteins which migrate near 68 kd. Our preliminary experiments using partial V8 protease digestion indicate that these proteins are not the same as p68 crosslinked to intact precursor (data not shown).

The 15-nt oligomer bound a more diverse set of proteins than either the 21- or 10-nt fragments. In addition, the spectrum of proteins varied with the batch of extract (compare Figure 6B, lane 5 and Figure 7B, lane 1). However, all the extracts tested had good polyadenylation activity and yielded a crosslinked protein of 155 kd. To determine if this protein is identical to the one crosslinked to intact precursor, the bands corresponding to these two proteins were cut out of the gel and the partial V8 protease patterns compared using the method of Cleveland *et al.* (1977). Both proteins show the same array of polypeptide products after digestion (Figure 6C).

Immunoprecipitation of crosslinked proteins by specific antisera

To further characterize the proteins crosslinked to precursor containing the L3 poly(A) site, specific antisera were tested for their ability to immunoprecipitate these proteins. RNA was incubated with nuclear extract, exposed to UV light, and digested with ribonuclease T1. Antiserum was added to this mixture and fixed *Staphylococcus* A bacteria used to precipitate the antibody−antigen complexes.

We tested monoclonal antibodies known to recognize two proteins in the heterogeneous ribonucleoprotein particles (hnRNPs) found on nuclear precursor RNA (Dreyfuss *et al.*, 1984). These are the C protein and a protein of 120 kd. The anti-120-kd antibody precipitated a 120-kd protein which

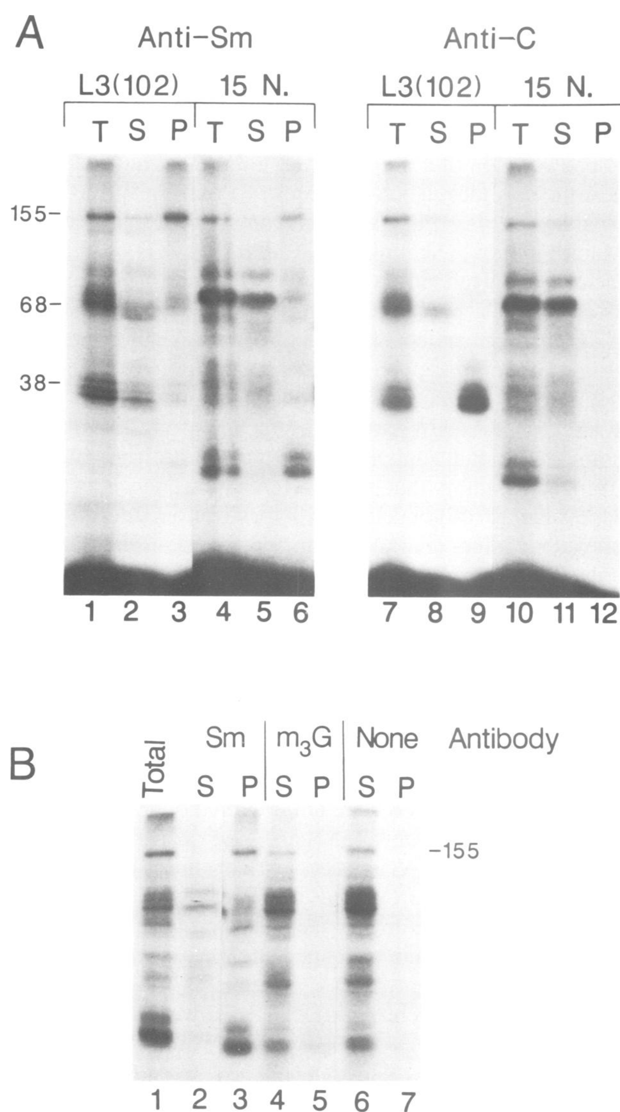


Fig. 7. Immunoprecipitation of crosslinked proteins. (**A**) Immunoprecipitation with anti-Sm and anti-C antibodies. Full length L3(102) RNA or the 15-nt fragment derived from this RNA was incubated with nuclear extract and exposed to UV light. The sample containing L3(102) RNA was then digested with ribonuclease T1. Proteins crosslinked to these substrates were immunoprecipitated with either anti-Sm (**lanes 1–6**) or anti-C (**lanes 7–12**) antibodies. Aliquots of the preparation before immunoprecipitation (T), of the precipitate (P) and of the supernatant (S) were resolved on a 10% polyacrylamide–SDS gel. (**B**) Immunoprecipitation with anti-trimethyl cap antiserum. The 15-nt oligomer was incubated with nuclear extract and exposed to UV light. Crosslinked proteins were immunoprecipitated with the indicated antibodies. **Lanes 2** and **3**, anti-Sm; **lanes 4** and **5**, anti-m_3G; **lanes 6** and **7**, without antibody. Total, proteins before immunoprecipitation; S, supernatant; P, precipitate.

corresponded to a minor band in the set of proteins crosslinked to L3 precursor (data not shown). When the anti-C antibody was used, almost all of p38 was immunoprecipitated (Figure 7A, lane 9). When this antibody was incubated with proteins crosslinked to the 15-nt oligomer, little protein was immunoprecipitated (Figure 7, lane 12). Of the purified oligomers which were tested, the only one which bound any C protein was the 10-nt fragment UUUUUAUUUG (data not shown).

A monoclonal antibody against the Sm proteins of the U class of snRNPs (Lerner *et al.*, 1981) was also tested. This antibody reacted strongly with p155 which had been crosslinked either to full-length precursor or the 15-nt oligomer (Figure 7A, lanes 3 and 6). When the 15-nt oligomer was used as substrate, a doublet at 32 kd was also efficiently immunoprecipitated with the Sm antibody (Figure 7A, lane 6). The nature of this protein is not known. p68 and p155 were not immunoprecipitated by autoimmune antisera (Padgett *et al.*, 1983) specific for the U1 or U2 snRNPs or the La nuclear antigen (data not shown).

To determine if p155 was a constituent of a U snRNP, we attempted to immunoprecipitate this protein with antibody against the trimethyl cap structure found on the U snRNAs (Luhrmann *et al.*, 1982). The stringent ribonuclease digestion used in the analysis of proteins crosslinked to full-length precursor would probably destroy these small RNAs. By using the 15-nt oligomer, it was possible to avoid this step. Also, the immunoprecipitation used in the previous experiments involved washing the pellet in a solution containing 0.5 M NaCl and 0.1% SDS (Goding, 1978) —conditions which could destabilize snRNP–protein interactions. Instead, a milder protocol was used which had been designed for immunoprecipitation of intact snRNPs from nuclear extracts (Padgett *et al.*, 1983). Under these conditions, p155 was again efficiently immunoprecipitated with the Sm antibody (Figure 7B, lane 3). However, none of this polypeptide was immunoprecipitated with the anti-trimethyl cap antiserum (Figure 7B, lane 5). This suggests that even though p155 has an Sm epitope, it is not strongly associated with a U snRNP.

Discussion

Components in nuclear extract from HeLa cells accurately and efficiently polyadenylate precursor RNAs containing mammalian polyadenylation sites (Moore and Sharp, 1985; Zarkower *et al.*, 1986). When precursor RNA containing the adenovirus L3 poly(A) site is incubated with this extract, exposed to UV light, and digested with ribonuclease, radioactive label from the RNA is transferred predominantly to three proteins of ~155, 68 and 38 kd mol. wt, designated p155, p68 and p38 respectively. Several lines of evidence from this study suggest that the two larger proteins may have a direct role in the processing reaction. They both require an intact AAUAAA sequence for binding, quickly associate with precursor upon incubation with extract, are found in polyadenylation-specific complexes thought to be processing intermediates, and are lost from the substrate RNA once it has been cleaved and polyadenylated. Wilusz and Shenk (1988) have recently reported a 64-kd protein which is crosslinked to precursor containing the SV40 early, the SV40 late or the adenovirus L3 site. This protein requires AAUAAA for binding and is probably analagous to the 68-kd protein described here.

The AAUAAA hexanucleotide is indispensable for formation of polyadenylated mRNA 3′ ends *in vivo* (for review, see Birnstiel *et al.*, 1985). The three steps of the *in vitro* polyadenylation reaction [the assembly of the polyadenylation processing complex, cleavage of the precursor, and the subsequent polymerization of a poly(A) tract] all depend on AAUAAA. Deletions and mutations (Manley *et al.*, 1985; Zarkower *et al.*, 1986; Humphrey

et al., 1987; Ryner and Manley, 1987; Skolnik-David *et al.*, 1987; Zarkower and Wickens, 1987a,b; Zhang and Cole, 1987) as well as chemical modifications (Conway and Wickens, 1987) of AAUAAA inhibit these functions. This signal sequence is also necessary for proper termination of RNA polymerase II transcription downstream of the polyadenylation site (Logan *et al.*, 1987). Because of their interaction with AAUAAA (discussed below), p155 and p68 could have a role in any of these events.

Mutation of AAUAAA to AAGAAA prevented binding of p155 and p68 to precursor RNA. It is unlikely that these proteins interact with the mutated RNA since excess amounts of the AAGAAA substrate cannot compete for binding of p155 and p68 to normal precursor. These two proteins may actually contact the RNA at the AAUAAA sequence. However, it is also possible that other factors which bind to AAUAAA, but which cannot be detected by crosslinking, direct binding of p155 and p68 to other parts of the precursor. Given this consideration, the AAUAAA sequence has not been confirmed as the binding site of p68. p155, on the other hand, binds to a 15-nt oligomer which contains the AAUAAA element. Thus, if p155 is involved in recognition of this signal sequence, it does not require extensive RNA structure for this interaction. While p155 did not bind to other fragments of the L3 precursor, it may contact other regions when they are part of an intact substrate RNA.

In addition to AAUAAA, sequences located downstream of this hexamer are important for the cleavage of polyadenylation precursor. These sequences are less conserved than the AAUAA, vary in sequence and type, and are often redundant. Common downstream regions include GU or U-rich elements (McLauchlan *et al.*, 1985; Hart *et al.*, 1985a; Proudfoot, 1987), a CAYUG pentanucleotide (Berget, 1984) and sequences that may potentially base pair with the AAUAAA (LeMoullec *et al.*, 1983). Deletion of such a downstream element from the L3 poly(A) site prevents cleavage and polyadenylation of precursor and reduces the efficiency of specific complex formation (Skolnik-David *et al.*, 1987; Moore *et al.*, 1988). Extract components may interact with downstream sequences since the downstream regions of SV40 late (Hasimoto and Steitz, 1986) and adenovirus L3 pre-mRNA (Humphrey *et al.*, 1987; Stefano and Adams, 1988) are protected from mild ribonuclease digestion. However, deletion of downstream sequences from L3 substrate did not affect the binding of p155 or p68. In addition, the AAGAAA substrate, which contained complete downstream sequences, could not compete for binding of these proteins. These findings support the conclusion that p155 and p68 interact primarily with sequences 5′ of the downstream element.

Also detected on precursor containing the rabbit β-globin and SV40 late poly(A) sites were 68-kd proteins. p155 was not detected on the latter RNA. SV40 late pre-mRNAs are efficiently processed *in vitro* (Zarkower *et al.*, 1986) by a mechanism similar to that used for L3 precursor (Moore *et al.*, 1986), and form complexes (Zarkower and Wickens, 1987b) similar to those assembled on L3 precursor (Skolnik-David *et al.*, 1987). However, a difference in the behavior of the L3 and SV40 late sites has been demonstrated after fractionation of crude extracts (Takagaki *et al.*, 1988). Cleavage at the SV40 late site required only one fraction. Cleavage at the L3, SV40 early, and adenovirus E2A sites required this fraction plus a fraction, which suprisingly, also contained nonspecific poly(A) polymerase. It will be interesting to see if the affinity of p155 for certain poly(A) sites correlates with differences in how these precursors are processed.

The addition of a poly(A) tract to RNAs which end at or near the polyadenylation site requires a functional AAUAAA (Manley *et al.*, 1985; Zarkower *et al.*, 1986; Ryner and Manley, 1987; Skolnik-David *et al.*, 1987). The same factors involved in AAUAAA recognition during the assembly of a processing complex may also mediate this AAUAAA-dependent polyadenylation event. The final elongation of the poly(A) tail appears not to depend on AAUAAA (M.Wickens, personal communication), and it is likely that AAUAAA-binding factors are released at this point. Additional support for this type of mechanism comes from the finding that the AAUAAA sequence becomes accessible to oligonucleotide hybridization after *in vitro* polyadenylation (Zarkower and Wickens, 1987a). Also, polyadenylated RNAs appear in a complex which migrates more quickly than the original processing complex on a nondenaturing polyacrylamide gel (Humphrey *et al.*, 1987; Skolnik-David *et al.*, 1987; Zhang and Cole, 1987). As proposed by Zarkower and Wickens (1987a), this dissociation would reasonably allow the processing activities to catalyze subsequent rounds of cleavage and polyadenylation. In light of this possibility, it is interesting to note that p155 and p68 were not observed on polyadenylated product. Given that p155 can bind a short fragment containing AAUAAA, it is not clear how this protein or other factors would be prevented from reassociating with the AAUAAA sequence on polyadenylated product. The poly(A) tail may have a role in this, since the AAUAAA of exogenously added polyadenylated RNA remains unprotected after incubation with nuclear extract (Zarkower and Wickens, 1987a).

In the formation of histone mRNA 3′ ends (Schaufele *et al.*, 1986) and in the removal of introns via splicing (Zhuang and Weiner, 1986; Parker *et al.*, 1987), small ribonucleoprotein particles (snRNPs) interact with critical processing signals via base pairing of sequence in the snRNA with that of precursor. However, recognition of specific RNA sequence in eukaryotes can also be mediated by proteins. A non-snRNP protein interacts with the conserved hairpin loop structure upstream of the human H3 histone cleavage site (Mowry and Steitz, 1987). Proteins of 150 kd (Tazi *et al.*, 1986) and 70 kd (Gerke and Steitz, 1986) bind to the polypyrimidine stretch upstream of splice acceptor sites. U snRNPs contain a common set of proteins which recognize and bind the sequence $AU_{3-6}G$ (Hamm *et al.*, 1987). Interestingly, like p155, the 70-kd splicing protein, the histone hairpin loop binding protein, and the shared U-snRNP proteins are all immunoprecipitated by a monoclonal anti-Sm antibody described by Lerner *et al.* (1981). This Sm epitope may be a common feature for a family of proteins which recognize specific but different RNA sequences.

Under certain conditions, the 70- and 150-kd splicing proteins are associated with snRNPs, but their ability to bind to intron sequence does not depend on this association. A similar situation may apply to factors which interact with polyadenylation sites. After short incubation of polyadenylation precursor with nuclear extract, ribonuclease-protected fragments containing the AAUAAA hexamer can

be immunoprecipitated with Sm antibody and with antibodies (anti-m_3G) against the trimethyl cap structure common to the U snRNAs (Hashimoto and Steitz, 1986; Stefano and Adams, 1988). It is important to note that the latter antibody recognizes the RNA component and not a protein constituent of snRNPs. When considered with our results, this finding implies that minimally a snRNP, p155 and possibly p68 are in close proximity to the AAUAAA sequence of intact precursor in the processing complex. When p155 was crosslinked to the 15-nt oligomer containing AAUAAA, it was not immunoprecipitated with anti-m_3G antibodies. This suggests that the snRNP implicated above is not complexed with p155 and that both of these entities cannot interact stably and simultaneously with such a short sequence. A factor with properties similar to p155 has also been reported by Stefano and Adams (1988). This factor remains bound to AAUAAA of L3 precursor after ribonuclease T1 digestion, is precipitated by the anti-Sm monoclonal, and is found in 25S polyadenylation-specific complex purified on sucrose gradients. Its association with AAUAAA does not require ATP or the downstream signal element and is resistant to micrococcal nuclease digestion. The size of this factor was not determined.

A specific set of proteins, called hnRNP proteins, is associated with mRNA precursor in the nucleus. The major constituents of hnRNP complexes are proteins ranging in mol. wt from 34 to 43 kd and have been classified into three distinct groups called A, B and C (Dreyfuss, 1986; Chung and Wooley, 1986). UV irradiation of hnRNP in intact cells (Dreyfuss *et al.*, 1984; Chung and Wooley, 1986; Dreyfuss, 1986) or of hnRNP assembled *in vitro* in HeLa extract (Economidis and Pederson, 1983) indicated that these core proteins are in close contact with RNA polymerase II transcripts. The C protein has the greatest affinity for RNA (Beyer *et al.*, 1977).

p38, one of the proteins efficiently crosslinked to L3 polyadenylation precursor, did not require AAUAAA for binding. This protein was immunoprecipitated by monoclonal antibody against the C hnRNP protein (Choi and Dreyfuss, 1984). Depletion of C protein from extract inhibits the *in vitro* splicing reaction (Choi *et al.*, 1986). It has been proposed that the C protein may have an important structural role in splicing. The binding of the C protein to L3 polyadenylation precursor can be greatly reduced by including excess AAGAAA substrate in the reaction. This does not slow down processing of the precursor, and suggests the C protein may not be as essential for polyadenylation as for splicing.

In summary, the major proteins crosslinked to polyadenylation precursor containing the L3 site are the C hnRNP protein and p68 and p155. The chemistry of UV crosslinking is only partly understood. Pyrimidine bases are more reactive than purines (Patrick and Rahn, 1976) and only some amino acids are involved (Smith, 1976). Thus, UV crosslinking may favor a particular type of RNA–protein interaction. For example, in the case of the cap-binding proteins of eukaryotic mRNA, a comparison of chemical and UV crosslinking techniques revealed different, but overlapping sets of proteins (Pelletier and Sonenberg, 1985). In a similar way, other polyadenylation-specific proteins may be demonstrated by alternative strategies to detect RNA–protein interactions.

p155 and p68 depend on AAUAAA for binding to polyadenylation precursor, are found in specific processing complexes and dissociate from the RNA once it has been processed. From the results described above, it is likely that these two proteins have a critical role in the polyadenylation process, such as recognizing the AAUAAA sequence and thus initiating formation of the processing complex. The role of these proteins should become clearer as crude extracts are fractionated into active components.

Materials and methods

Preparation of substrate RNA, polyadenylation reactions and analysis of RNA products

For synthesis of RNA substrate, plasmids pT3L3 and pT3L3G (Skolnik-David *et al.*, 1987) were digested with either *Dra*I or *Eco*RI to give substrate RNA 102 or 281 nt in length respectively. pSPL3(dl+5) was digested with *Eco*RI. Plasmids pβSP65, containing the rabbit β-globin poly(A) site, and pSPSVL, containing the SV40 late poly(A) site, were generous gifts from A.Gil and S.Berget respectively. Run-off 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% (v/v) HeLa nuclear extract as described previously (Moore and Sharp, 1985). RNAs from reaction mixtures were purified by proteinase K digestion, phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation, and separated on 10 or 15% polyacrylamide/8.3 M urea gels. The concentration of ATP in the extract was determined using firefly lantern extract from Sigma (Kimmich *et al.*, 1975).

UV crosslinking of proteins to RNA

Polyadenylation reaction mixtures containing substrate RNA (typically 100 000 c.p.m. equivalent to ~0.5 ng RNA) were incubated at 30°C for various times. tRNA was added at a final concentration of 200 μg/ml to dissociate weakly bound proteins. Proteins were crosslinked to RNA by exposure to UV light at 254 nm for 10 min at 4°C. The UV light source (Ultraviolet Products Model UVG-54) was supported 4.5 cm above the sample. The intensity at this distance was 1.8 mW/m^2. Samples were digested with either ribonuclease T1 (2.5 U/μl), ribonuclease A (1 μg/μl), ribonuclease P1 (0.4 U/μl) or a combination of ribonucleases T1 and A for 30 min at 37°C. Protein gel loading buffer was mixed into the reaction and the samples were heated at 100°C for 3 min, centrifuged for 1 min to remove precipitates, and electrophoresed on an SDS-containing discontinuous PAGE system (Laemmli, 1970). The separating gel had a final acrylamide concentration of 10%. The tracking dye was typically run 8–9 cm. ^{14}C-labeled proteins (BRL) were used as markers. Gels were fixed, enhanced with Autofluor (National Diagnostics) when ^{14}C-labeled markers were used, dried and the proteins visualized by autoradiography.

Analysis of proteins in complexes

HeLa nuclear extract (100 μl) was preincubated at 30°C for 30 min to remove endogenous ATP. Precursor RNA L3(102)AAUAAA, ATP and creatine phosphate were added to the extract and complexes were allowed to form for 20 min at 30°C. After this period, heparin to 2 mg/ml and tRNA to 200 μg/ml were added and the mixture further incubated on ice for 10 min. Samples were crosslinked with UV light and loaded onto a 1.5-mm-thick nondenaturing 4% acrylamide gel containing 45 mM Tris-borate, pH 8.3, and 1.2 mM EDTA (Skolnik-David *et al.*, 1987). The gel was pre-electrophoresed for 30 min at 18 V/cm. Electrophoresis was carried out at 4°C at the same voltage gradient for 3–4 h until the xylene cyanol dye reached 12 cm. The gel was autoradiographed for 30–60 min at room temperature to determine the position of the complex. The specific complex (complex B) and the nonspecific complex (complex A) were cut out and the RNA electroeluted at 4°C in 45 mM Tris-borate and 0.5% SDS for 15 h using 10 mA of current on an Isco electroelutor. The electroeluted sample was ethanol precipitated, resuspended in 25 μl of a solution containing 0.5% *N*-lauryl sarcosine (Sigma), 0.5 mM DTT, 10 mM Tris (pH 7.5), 1 mM EDTA and boiled for 3 min. Insoluble material was removed by centrifugation and the sample was digested with ribonuclease T1 (2.5 U/μl) for 60 min at 37°C. Samples were analyzed by electrophoresis on 10% polyacrylamide–SDS gel.

Poly(U) Sephadex chromatography

A 100-μl polyadenylation reaction was carried out for 60 min. After crosslinking with UV light, 5 μl was removed for RNA analysis and 20 μl was digested with ribonucleases T1 and A as described previously. To the remaining sample, 1 ml of binding buffer mix (200 mM NaCl, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.2% SDS, 1 mM DTT), and 100 μl of poly(U) Sephadex beads (BRL) were added, and the tubes rotated 10 min at room temperature. The beads were spun down and the supernatant containing the poly(A)$^-$ RNA was ethanol precipitated. After rinsing the beads once with 1 ml binding buffer and twice with 1 ml each of a wash solution (10% formamide, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.2% SDS), poly(A)$^+$ RNA was released in 200 μl of a solution containing 90% formamide, 10 mM Tris, pH 7.5, 1 mM EDTA, 0.2% SDS, and then ethanol precipitated. Both poly(A)$^+$ and poly(A)$^-$ samples were resuspended in 0.5% *N*-lauryl-sarcosine, 10 mM Tris, pH 7.5, and 1 mM EDTA. Aliquots were removed for RNA analysis. The remaining samples were boiled, centrifuged to pellet insoluble material, and digested with ribonucleases T1 (2.5 U/μl) and A (1 μg/μl) for 30 min at 37°C. Protein gel loading buffer was added and the samples were boiled, spun, and electrophoresed on a discontinuous SDS–polyacrylamide gel. For RNA analysis, samples from the total reaction mixture as well as the poly(A)$^+$ and poly(A)$^-$ fractions were digested with proteinase K at 125 μg/ml for 10 min at 30°C, extracted once with phenol:chloroform:isoamyl alcohol (50:48:2), ethanol precipitated and electrophoresed on a 15% polyacryalmide/8.3 M urea gel.

Partial peptide analysis

RNA was crosslinked to protein with UV light and the proteins resolved by gel electrophoresis. The appropriate bands were cut out of the gel and partially digested with V8 protease as described previously (Cleveland *et al.*, 1977).

Immunoprecipitations

Immunoprecipitations were performed with the monoclonal anti-Sm antibody Y12 (Lerner *et al.*, 1981), human autoimmune antisera against the U1 snRNP (AG), the U2 snRNP (MG), or the La nuclear antigen (ON), the monoclonal anti-C antibody 4F4 (Choi and Dreyfuss, 1984), the monoclonal anti-120-kd antibody 3G3 (Dreyfuss *et al.*, 1984), and the polyclonal anti-m_3G antibody (Lührman *et al.*, 1982). These antibodies were generous gifts of Drs J.Steitz, G.Dreyfuss and R.Lührmann respectively. The activity of antibodies against snRNAs were verified by visualizing the immunoprecipitated RNAs on ethidium-bromide-stained polyacrylamide gels. After crosslinking proteins to RNA and if necessary, ribonuclease digestion, reactions were diluted 1:3 in 2% Triton, 50 mM Tris, pH 8, 150 mM NaCl, 1 mM EDTA, 20 μg/ml PMSF and 0.5 μg/ml leupeptin. Antibody was added and the mixture incubated for 5 h on ice. Seventy microliters of a solution containing 10% fixed *Staphylococcus* A bacteria (Igsorb, The Enzyme Center, Malden, MA) was added and the incubation continued for 30 min. After centrifugation, the pellet was washed twice with 10 mM Tris, pH 7.5, 0.5 M NaCl, 1% Triton X-100, 0.1% SDS, and 1 mM EDTA, and once with 10 mM Tris, pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS and 1 mM EDTA. The pellet was then resuspended in 50 μl of protein gel loading buffer for electrophoresis.

To immunoprecipitate snRNPs, samples were incubated with antibody for 60 min on ice, followed by the addition of 200 μl of Igsorb solution, and a further incubation of 30 min (Padgett *et al.*, 1983). The immunoprecipitate was washed four times with 1 ml of 150 mM NaCl, 50 mM Tris, pH 7.5, 0.05% NP40 and 0.5 mM EDTA. RNAs in the pellet were purified by phenol:chloroform:isoamyl alcohol extraction, ethanol precipitated, separated on a 10% polyacrylamide/8.3 M urea gel, and visualized by ethidium bromide staining. To examine proteins, the pellet was resuspended in 50 μl protein gel loading buffer.

Acknowledgements

We are grateful to G.Dreyfuss, R.Lurhmann and J.Steitz for gifts of antibodies and S.Berget and A.Gil for sharing DNA clones. We also wish to thank L.Hyman, M.Kessler, A.Gil, M.Garcia and P.Sharp for critical review of the manuscript and many helpful discussions. This research was supported by N.S.F. Grant No. DMB-86067757.

References

Adam,S.A., Nakagawa,T., Swanson,M.S., Woodruff,T.K. and Dreyfuss,G. (1986) *Mol. Cell. Biol.*, **6**, 2932–2943.

Berget,S.M. (1984) *Nature*, **309**, 179–182.

Beyer,A.L., Christensen,M.E.,, Walker,B.W. and LeStourgeon,W.M. (1977) *Cell*, **11**, 127–138.

Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349–359.

Choi,Y.D. and Dreyfuss,G. (1984) *J. Cell Biol.*, **99**, 1977–2004.

Choi,Y.D., Grabowski,P.J., Sharp,P.A. and Dreyfuss,G. (1986) *Science*, **231**, 1534–1539.

Chung,S.Y. and Wooley,J. (1986) *Proteins*, **1**, 195–210.

Cleveland,D.W., Fischer,S.G., Kirschner,M.W. and Laemmli,U.K. (1977) *J. Biol. Chem.*, **252**, 1102–1106.

Conway,L. and Wickens,M. (1987) *EMBO J.*, **6**, 4177–4184.

Dreyfuss,G. (1986) *Annu. Rev. Cell Biol.*, **2**, 459–498.

Dreyfuss,G., Choi,Y.D. and Adam,S.A. (1984) *Moll. Cell. Biol.*, **4**, 1104–1114.

Economidis,I.V. and Pederson,T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1599–1602.

Ehresmann,B., Backendorf,C., Ehresmann,C. and Ebel,J.P. (1977) *FEBS Lett.*, **78**, 261–266.

Gerke,V. and Steitz,J.A. (1986) *Cell*, **47**, 973–984.

Gil,A. and Proudfoot,N.J. (1987) *Cell*, **49**, 399–406.

Goding,J.W. (1978) *J. Immunol. Methods*, **20**, 241–253.

Green,T.L. and Hart,R.P. (1988) *Mol. Cell. Biol*, **8**, 1839–1841.

Hamm,J., Kazmaier,M. and Mattaj,I.W. (1987) *EMBO J.*, **6**, 3479–3485.

Hart,R.P., McDevitt,M.A., Ali,H. and Nevins,J.R. (1985a) *Mol. Cell. Biol.*, **5**, 2975–2983.

Hart,R.P., McDevitt,M.A. and Nevins,J.R. (1985b) *Cell*, **43**, 677–683.

Hashimoto,C . and Steitz,J.A. (1986) *Cell*, **45**, 581–591.

Humphrey,T., Christofori,G., Lucijanic,V. and Keller,W. (1987) *EMBO J.*, **6**, 4159–4168.

Kimmich,G.A., Randles,J. and Brand,J.S. (1975) *Anal. Biochem.*, **69**, 187–206.

Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1984) *Cell*, **38**, 731–736.

Laemmli,U.K. (1970) *Nature*, **277**, 680–685.

LeMoullec,J.M., Akusjarvi,G., Stalhandske,R., Pettersson,U., Chambraud,B., Gilardi,P., Nasri,M. and Perricaudet,M. (1983) *J. Virol.*, **48**, 127–134.

Lerner,E.A., Lerner,M.R., Janeway,C.A. and Steitz,J.A. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2737–2741.

Logan,J., Ralph-Pedersen,E., Darnell,J.E. and Shenk,T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8306–8310.

Lührmann,R., Appel,B., Bringmann,P., Rinke,J., Reuter,R., Tothe,S. and Bald,R. (1982) *Nucleic Acids Res.* **10**, 7103–7113.

Manley,J.L., Yu,H. and Ryner,L.C. (1985) *Mol. Cell. Biol.*, **5**, 373–379.

Mclauchlan,J., Gaffney,D., Whiton,J.L. and Clements,J.B. (1985) *Nucleic Acids Res.*, **13**, 1347–1368.

Moller,K. and Brimacombe,P. (1975) *Mol. Gen. Genet.*, **141**, 343–335.

Moore,C.L. and Sharp,P.A. (1985) *Cell*, **41**, 845–855.

Moore,C.L., Skolnik-David,H. and Sharp,P.A. (1986) *EMBO J.*, **5**, 1929–1938.

Moore,C.L., Skolnik-David,H. and Sharp,P.A. (1988) *Mol. Cell. Biol.*, **8**, 226–233.

Mowry,K.L. and Steiz,J.A. (1987) *Science*, **238**, 1682–1687.

Padgett,R.A., Mount,S.M., Steitz,J.A. and Sharp,P.A. (1983) Cell, **35**, 101–107.

Parker,R., Siliciano,P.G. and Guthrie,C. (1987) *Cell*, **49**, 229–239.

Patrick,M.H. and Rahn,R.O. (1976) In Wang,S. (ed.), *Photochemistry and Photobiology of Nucleic Acids*. Academic Press, New York, Vol. 2, pp. 35–95.

Pelletier,J. and Sonenberg,N. (1985) *Mol. Cell. Biol.*, **5**, 3222–3230.

Rinke,J., Yuki,A. and Brimacombe,R. (1976) *Eur. J. Biochem.*, **64**, 77–89.

Rozen,R. and Sonenberg,N. (1987) *Nucleic Acids Res.*, **15**, 6489–6500.

Ryner,L.C. and Manley,J.L. (1987) *Mol. Cell. Biol.*, **7**, 495–503.

Sachs,A.B., Bond,M.W. and Kornberg,R.D. (1986) *Cell*, **45**, 827–835.

Schaufele,F., Gilmartin,G.M., Bannwarth,W. and Birnstiel,M.L. (1986) *Nature*, **323**, 777–781.

Schoenmaker,H.J. and Schimmel,P.R. (1974) *J. Mol. Biol.*, **84**, 503–513.

Sen,A. and Todaro,G. (1977) *Cell*, **10**, 91–99.

Setyono,B. and Greenberg,J.R. (1981) *Cell*, **24**, 775–783.

Skolnik-David,H., Moore,C.L. and Sharp,P.A. (1987) *Genes Dev.*, **1**, 672–682.

Smith,K.C. (1976) In Wang,S. (ed.), *Photochemistry and Phtobiology of Nucleic Acids*. Academic Press, New York, Vol. 2, pp. 187–218.

Sperry,A.O. and Berget,S.M. (1986) *Mol Cell. Biol.*, **6**, 4734–4741.

Stefano,J. and Adams,E.E. (1988) *Mol. Cell Biol.*, **8**, 2052–2062.

Takagaki,Y., Ryner,L.C. and Manley,J.L. (1988) *Cell*, **52**, 731–742.

Tazi,J., Alibert,C., Temsamani,J., Reveillaud,I., Cathala,G., Brunel,C. and Jeanteur,P. (1986) *Cell*, **47**, 755–766.

Wilusz,J. and Shenk,T. (1988) *Cell,* **52**, 221–228.
Zarkower,D. and Wickens,M. (1987a) *EMBO J.*, **6**, 177–186.
Zarkower,D. and Wickens,M. (1987b) *EMBO J.*, **6**, 4185–4192.
Zarkower,D., and Wickens,M. (1988) *J. Biol. Chem.*, **263**, 5780–5788.
Zarkower,D., Stephenson,P., Sheets,M. and Wickens,M. (1986) *Mol. Cell. Biol.*, **6**, 2317–2323.
Zhang,F. and Cole,C.N. (1987) *Mol. Cell. Biol.*, **7**, 3277–3286.
Zhuang,Y. and Weiner,A.M. (1986) *Cell,* **46**, 827–835.

Received on May 17, 1988; revised on July 1, 1988