The EMBO Journal vol.6 no.13 pp.4177–4184, 1987

Analysis of mRNA 3′ end formation by modification interference: the only modifications which prevent processing lie in AAUAAA and the poly(A) site

Laura Conway and Marvin Wickens

Department of Biochemistry, College of Agriculture and Life Sciences, University of Wisconsin–Madison, Madison, WI 53706, USA

Communicated by W.Keller

A modification interference method is described in which chemically modified transcripts are used to identify bases required for any reaction for which synthetic RNA is a substrate. This technique provides information analogous to that obtained from the analysis of a complete set of point mutants. Using SV40 late pre-mRNAs, we determine that modification of any base in the AAUAAA sequence prevents cleavage, polyadenylation and formation of pre-cleavage complexes *in vitro*. Modification of the A to which poly(A) is added prevents polyadenylation, but does not interfere with formation of the pre-cleavage complex. No single modification downstream of the poly(A) site significantly affects cleavage efficiency. Since the region downstream of the poly(A) site is required for cleavage and complex formation (Conway and Wickens, 1985; Zarkower and Wickens, 1987b), we infer that the critical features of this downstream region are either diffuse or redundant.

Key words: AAUAAA/modification interference/mRNA processing/polyadenylation

Introduction

In eukaryotes, the 3′ termini of mRNAs are generated by RNA processing, rather than by transcription termination (reviewed in Nevins, 1983; Birnstiel *et al.*, 1985). Endonucleolytic cleavage of mRNA precursors is generally followed by the addition of up to 250 adenylate residues [poly(A); Nevins and Darnell, 1978].

Cleavage and polyadenylation require at least two separate sequence elements: the highly conserved sequence AAUAAA, found 6–30 nucleotides upstream of the poly(A) site, and a less well conserved sequence found downstream of the poly(A) site (reviewed in Birnstiel *et al.*, 1985). Mutations in AAUAAA greatly reduce cleavage efficiency both *in vivo* (Proudfoot and Brownlee, 1976; Fitzgerald and Shenk, 1981; Montell *et al.*, 1983; Wickens and Stephenson, 1984; Higgs *et al.*, 1984) and *in vitro* (Zarkower *et al.*, 1986). AAUAAA is also required for polyadenylation *in vitro* (Manley *et al.*, 1985; Zarkower *et al.*, 1986; Skolnik-David *et al.*, 1987).

Mutations in the downstream region reduce cleavage efficiency (reviewed in Birnstiel *et al.*, 1985). For SV40 late pre-mRNA, examined in this report, deletion or substitution of the region between 1 and 27 bases downstream of the poly(A) site reduces cleavage efficiency 5-fold in injected oocytes (Conway and Wickens, 1985) and 20-fold *in vitro* in a crude extract of HeLa cell nuclei (Zarkower and Wickens, 1987b).

Prior to cleavage, factors in the extract bind to the mRNA precursor, probably at AAUAAA (Hashimoto and Steitz, 1986; Zarkower and Wickens, 1987a). Formation of this pre-cleavage complex requires both the AAUAAA sequence and the region downstream of the poly(A) site (Zhang and Cole, 1987; Skolnik-David *et al.*, 1987; D.Zarkower and M.Wickens, submitted; Zarkower and Wickens, 1987b; Humphrey *et al.*, 1987).

To identify any other single nucleotides which are critical for 3′ end formation, we have developed a modification interference method. It provides, in a single experiment, information analogous to that which would be obtained by analyzing a complete set of point mutants. The strategy is similar to that used to examine protein–DNA interactions by methylation interference (Siebenlist and Gilbert, 1980), but is here adapted to analyze RNA. Analogous approaches have been used to analyze ribosome assembly (Herr *et al.*, 1979) and formation of a tRNA-ribosome complex (Peattie and Herr, 1981). In principle, modification interference can be used to identify critical residues for any reaction for which synthetic RNA is a substrate.

In this report, we use modification interference to examine the formation of SV40 late mRNA 3′ termini *in vitro*, in a crude extract of HeLa cell nuclei. We identify those nucleotides in the pre-mRNA which are essential for cleavage, polyadenylation and formation of the pre-cleavage complex. B.Rymond and M. Rosbash, who developed the modification interference method independently, have applied it to yeast pre-mRNA splicing (in preparation).

Results

Modification interference approach

Treatment of RNA with diethylpyrocarbonate (DEPC) carboxyethylates purine bases (Henderson *et al.*, 1973; Vincze *et al.*, 1973); hydrazine treatment removes pyrimidine bases (Brown, 1967; Cashmore and Petersen, 1978; Figure 1A). Although neither treatment breaks the RNA chain, the phosphate backbone is rendered susceptible to cleavage by aniline at modified nucleotides (Turchinskii *et al.*, 1969; Peattie, 1979). This susceptibility can be exploited to identify specific single nucleotides which are critical for mRNA processing, using the strategy diagrammed in Figure 1B. The example illustrated identifies nucleotides required for cleavage and polyadenylation, but is equally applicable to other processing reactions. End-labeled RNA is treated with DEPC or hydrazine under conditions which modify, on average, one nucleotide per RNA molecule. The collection of modified RNAs is incubated in nuclear extract, under conditions suitable for cleavage and polyadenylation. RNAs which successfully undergo processing ('processed RNAs') are purified. The processed RNAs, and RNA which has not been incubated in the extract ('starting RNA'), are treated with aniline, and the products analyzed by gel electrophoresis. Gaps in the sequencing 'ladder' of processed RNA correspond to sites which, when modified, prevent cleavage and polyadenylation.

The strategy hinges on the assumption that processing will be blocked by modification of individual critical nucleotides. Since both modifications drastically alter nucleotide structure, this assumption is reasonable (Figure 1A).

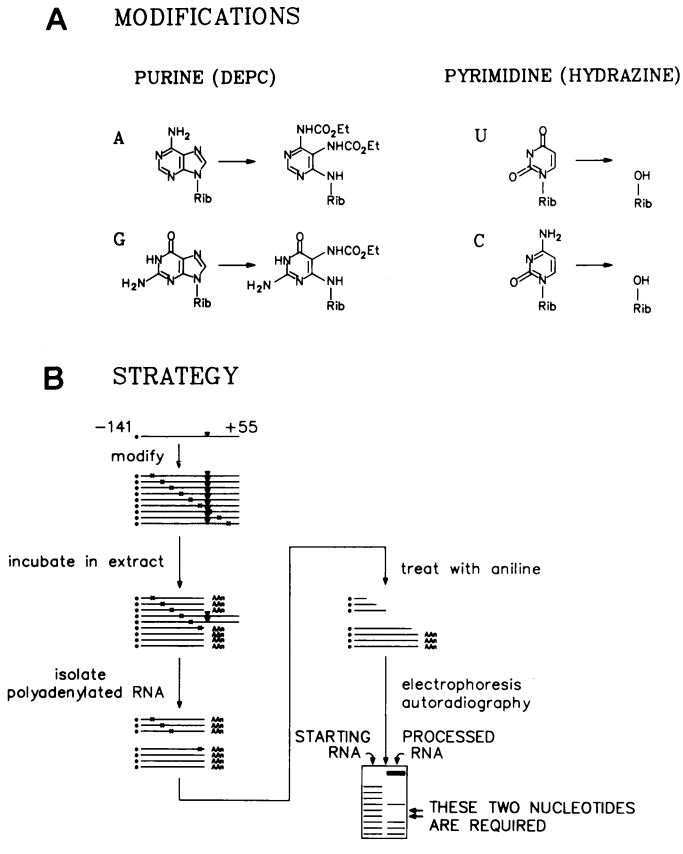

Fig. 1. Chemical modification approach. (**A**) Structure of bases following treatment with either DEPC (purines) or hydrazine (pyrimidines). (**B**) Modification interference strategy. An analysis of cleavage and polyadenylation is depicted, but the same strategy can be applied to any other processing reaction. Modified nucleotides are indicated by ×. The asterisk at the end of the RNA indicates that it is end-labeled. See text for details.

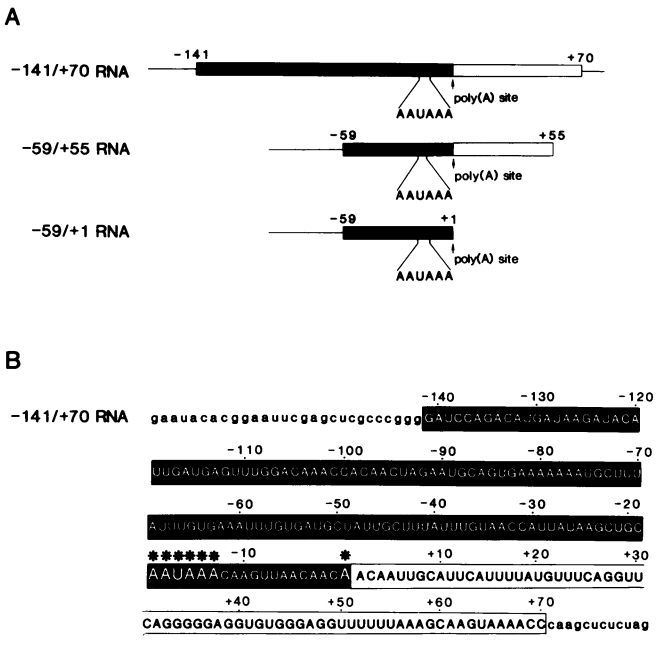

Fig. 2. RNA substrates. (**A**) Structure of transcripts. Thin lines represent the vector sequence. SV40 sequence upstream of the poly(A) site is shown as a black box, SV40 sequence downstream of the poly(A) site as a white box. The positions of AAUAAA and the poly(A) site are marked. (**B**) Sequence of −141/+70 RNA. Vector sequence is in small letters. SV40 sequence is in capital letters. AAUAAA and the A to which poly(A) is added are marked with asterisks. The numbering system used throughout the paper is given above the sequence: the A to which poly(A) is added is designated +1.

Structure of transcripts

In Figure 2A, the structures of the three transcripts used in this report are shown. The sequence of the longest transcript, −141/+70 RNA, is also provided (Figure 2B). Each transcript is named according to the SV40 sequence it contains. For example, −141/+70 RNA extends from 141 bases upstream to 70 bases downstream of the SV40 late poly(A) site. Transcripts were synthesized using RNA polymerase from bacteriophage SP6 (Melton *et al.*, 1984), and possess 5′ triphosphate termini, not caps, unless noted otherwise.

Nucleotides required for cleavage

The region upstream of the poly(A) site. 5′ End-labeled −59/+55 RNA was modified either at purines (with DEPC) or at uridines (with hydrazine). The modified RNA was incubated in nuclear extract containing 0.5 mM EDTA, which permits cleavage but prevents polyadenylation (Moore and Sharp, 1985). After the reaction, 5′ half-molecules and molecules which had not been cleaved were purified by gel electrophoresis. Both RNAs were then treated with aniline and analyzed by electrophoresis. As a control, modified RNA which had not been incubated in the extract was analyzed in parallel.

DEPC-treated RNA that was not processed (Figure 3A, lane 2) yields bands at every purine producing a sequencing 'ladder' almost identical to that obtained with DEPC-treated RNA which was never incubated in extract (lane 1). Similarly, RNA that was modified with hydrazine but not processed yields bands at every uridine (lane 4). In contrast, RNAs that were cleaved, which extend from −59 to +1, yield ladders which lack those bands corresponding to AAUAAA (lanes 3 and 5). −141/+55 RNA generates the same result as −59/+55 RNA (data not shown). Thus, cleavage requires AAUAAA but no other single nucleotides upstream of the poly(A) site.

Several details of the analysis merit discussion. First, we are unable to examine cytidines with 5′ end-labeled RNA (see Materials and methods). Second, although one might expect that bands corresponding to essential residues should be enriched in RNA which is not processed, this is not observed. We do not detect any enrichment because, under the processing conditions used here, 75% of even unmodified RNA is not processed. Third, incubation of carboxyethylated RNA in the extract results in an enhanced sensitivity of G residues to aniline cleavage (compare the G at −51 in lane 1 with lanes 2 and 3, Figure 3A). This enhancement is seen in all RNA fractions under all conditions tested and is unrelated to cleavage and polyadenylation.

The region downstream of the poly(A) site. To examine nucleotides downstream of the poly(A) site, 3′ end-labeled −141/+70 RNA was modified with DEPC or hydrazine and incubated in nuclear extract. The most abundant 3′ half-molecule, which extends from +5 to +70 (Sheets *et al.*, 1987), was purified by gel electrophoresis, as was RNA which failed to be processed. After aniline treatment, processed and unprocessed RNAs yield identical ladders, whether modified with DEPC (A + G) or hydrazine (U + C). Each nucleotide in the sequence is represented (Figure 3B). The same result is obtained with the longest 3′ half molecule (+2 to +70; data not shown). We conclude that no single modifications downstream of the poly(A) site prevent cleavage.

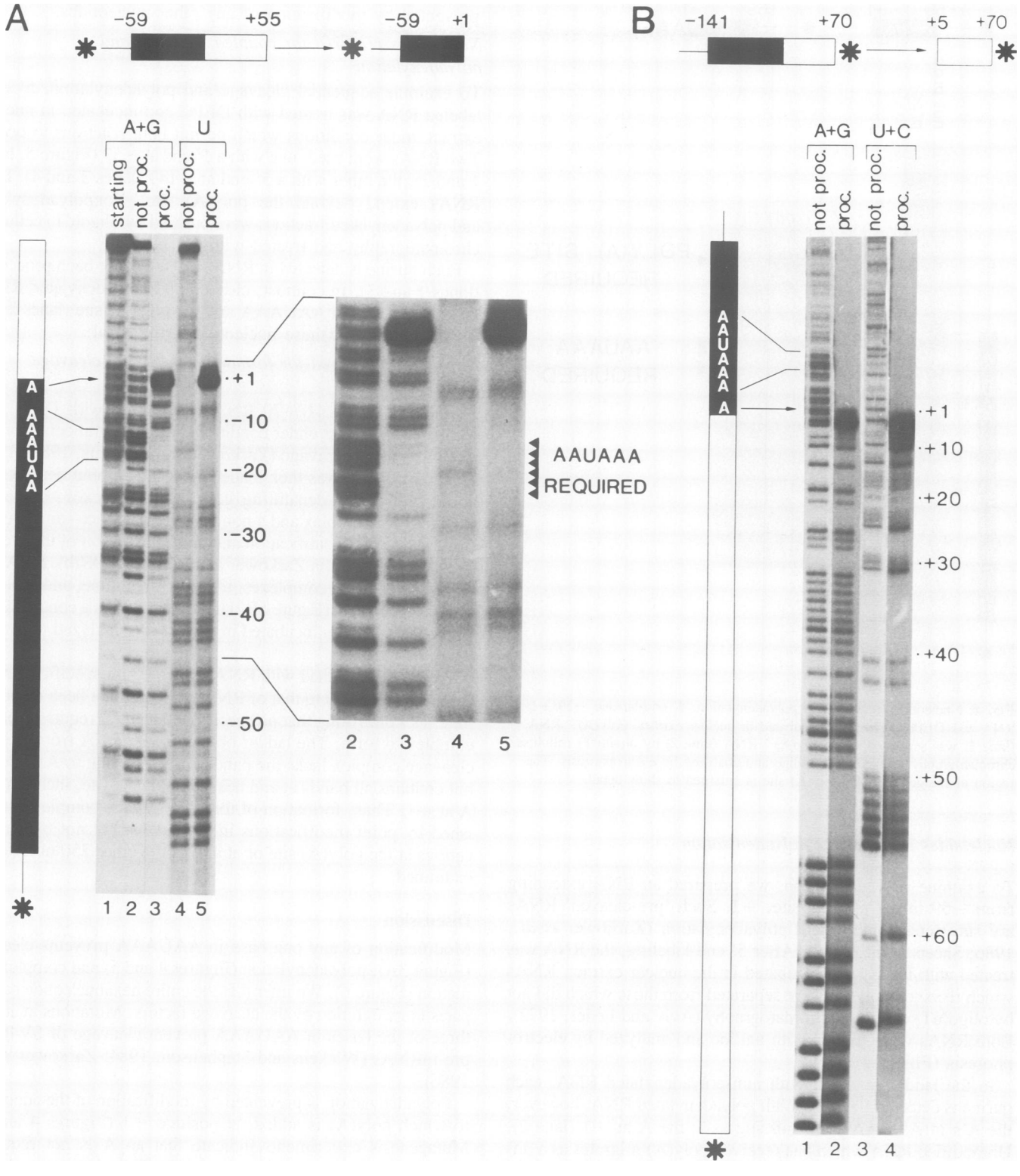

Fig. 3. Nucleotides required for cleavage. **(A)** Nucleotides upstream of the poly(A) site. 5′ end-labeled −59/+55 RNA was treated with DEPC (lanes 1−3) or hydrazine (lanes 4 and 5) and incubated in nuclear extract which contained 0.5 mM EDTA to prevent polyadenylation. 5′ Half-molecules were isolated and treated with aniline. **Lane 1:** DEPC-treated RNA that was not incubated in extract; **lane 2:** DEPC-treated RNA that was not cleaved during incubation in the extract; **lane 3:** DEPC-treated RNA that was cleaved in the extract; **lane 4:** hydrazine-treated RNA that was not cleaved in extract. **lane 5:** hydrazine-treated RNA that was cleaved in extract. The sequence surrounding AAUAAA is enlarged to show detail. AAUAAA is marked with arrowheads. **(B)** Nucleotides downstream of the poly(A) site. 3′ End-labeled −141/+70 RNA was treated with DEPC (lanes 1 and 2) or hydrazine (lanes 3 and 4) and incubated in nuclear extract containing 0.5 mM EDTA. 3′ Half-molecules were isolated and treated with aniline. **Lane 1:** DEPC-treated RNA that was not cleaved in extract; **lane 2:** DEPC-treated RNA that was cleaved in extract; **lane 3:** hydrazine-treated RNA that was not cleaved in extract; **lane 4:** hydrazine-treated RNA that was cleaved in extract.

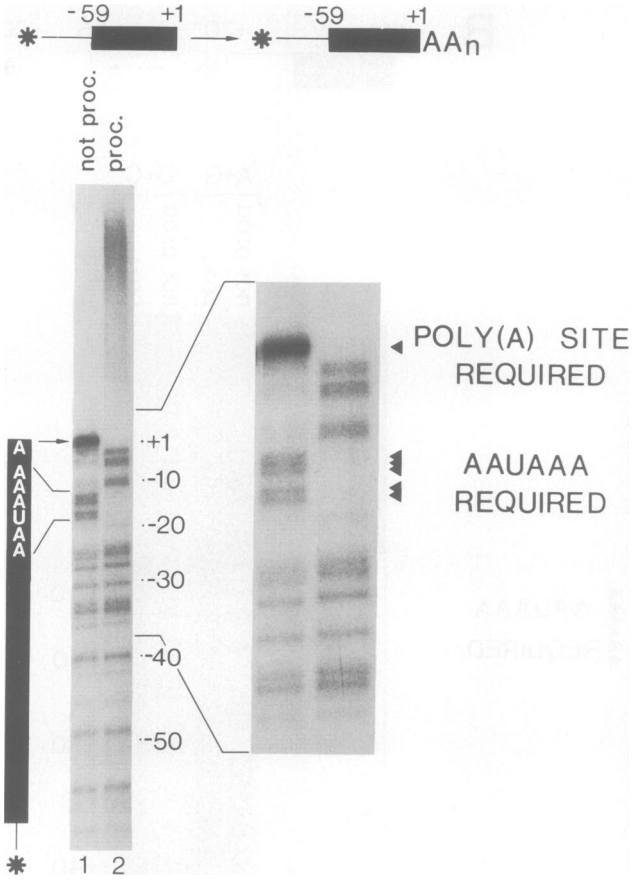

Fig. 4. Nucleotides required for polyadenylation. 5′ end-labeled −59/+1 RNA was DEPC-treated and incubated in nuclear extract. Poly(A)$^-$ RNA (**lane 1**) was separated from poly(A)$^+$ RNA (**lane 2**) by oligo(dT) cellulose chromatography, treated with aniline and analyzed by electrophoresis. The region of the gel near the poly(A) site is enlarged to show detail.

Nucleotides required for polyadenylation

To examine polyadenylation, we prepared an RNA extending from −59 to the poly(A) site, +1. Such 'pre-cleaved' RNAs are substrates for efficient polyadenylation (Zarkower *et al.*, 1986; Sheets *et al.*, 1987). After 5′ end-labeling, the RNA was treated with DEPC and incubated in the nuclear extract. RNAs which received poly(A) were separated from those which did not by oligo(dT) cellulose chromatography (Aviv and Leder, 1972). Both RNAs were treated with aniline and analyzed by electrophoresis (Figure 4).

In the ladder obtained with non-polyadenylated RNA, each purine is detected (lane 1). With polyadenylated RNA, however, the A's in AAUAAA are missing as is the A at +1 (lane 2). Unmodified RNA which has received poly(A) migrates as a diffuse high molecular weight band in lane 2, due to the addition of heterogeneous lengths of poly(A).

Purines between AAUAAA and the poly(A) site are enriched in polyadenylated RNA. We believe that this enrichment is artifactual, both because it is not observed when polyadenylation is coupled to cleavage (Figure 5), and because mutations in this region do not affect polyadenylation significantly (M.Sheets, unpublished data). Our interpretation of the enrichment is presented in the Discussion.

We conclude that polyadenylation is prevented by modification of AAUAAA or of the base to which poly(A) is added. It is not prevented by modification of any other nucleotides.

Nucleotides required for coupled cleavage and polyadenylation

To examine sequential cleavage and polyadenylation, 5′ end-labeled RNA was treated with DEPC and incubated in nuclear extract under conditions which permit both reactions to occur. Two RNAs were used, one with a 5′ end at −59 (Figure 5, lanes 1 and 2), the other with a 5′ end at −141 (lanes 3 and 4). Both RNAs extend beyond the poly(A) site. Nonpolyadenylated and polyadenylated fractions were purified by oligo(dT) cellulose chromatography and treated with aniline.

Each purine is detected in RNA which has not been processed (lanes 1 and 3). In contrast, cleaved and polyadenylated RNAs lack the purines in AAUAAA and the poly(A) site (lanes 2 and 4), indicating that those nucleotides are critical.

Nucleotides required for formation of the pre-cleavage complex

3′ end-labeled −141/+70 RNA was treated with DEPC and incubated in extract under conditions which permit complex formation, but not cleavage (Zarkower and Wickens, 1987b). The reaction mixture was then loaded directly, without deproteinization, onto a non-denaturing gel. Two complexes—a specific pre-cleavage complex and a non-specific complex—are detected (Figure 6A; the identification and characterization of these complexes is presented in Zarkower and Wickens, 1987b). RNA was purified from both complexes, treated with aniline, and analyzed by electrophoresis (Figure 6B, lanes 2 and 3). As a control, RNA which had not been incubated in extract was analyzed in parallel (lane 1).

The ladder obtained with RNA from the non-specific complex (lane 2) is identical to that of RNA which has not been incubated (lane 1), confirming that no specific sequence is required for this complex to form. The ladder obtained with RNA from the pre-cleavage complex lacks the bands corresponding to AAUAAA, but contains all bands at and near the poly(A) site, including the A at +1. Thus, formation of the pre-cleavage complex is abolished by point modifications in AAUAAA but not by modifications at the poly(A) site or at any of the other 250 nucleotides examined.

Discussion

Modification of any one base in AAUAAA prevents cleavage (Figure 3A), polyadenylation (Figures 4 and 5) and complex formation (Figure 6). We infer that the entire hexanucleotide participates in recognition by the cleavage factors. Mutations in at least three of the bases in AAUAAA prevent cleavage of SV40 late pre-mRNAs (Wickens and Stephenson, 1984; Zarkower *et al.*, 1986).

Polyadenylation is prevented by modification of the adenosine to which poly(A) is added, at position +1 (Figures 4 and 5). Mutagenesis experiments indicate that an A is not required, however; U termini are polyadenylated efficiently *in vitro* (Moore *et al.*, 1986; Sheets *et al.*, 1986). Thus, a carboxyethylated adenosine is inert, but a non-adenosine nucleotide is active. Critical features of the terminal nucleotide could be identified by analyzing RNAs which contain nucleotide derivatives at that position.

Modification of the A at +1 does not detectably inhibit formation of a pre-cleavage complex. Complex formation is reduced 2- to 3-fold by substitution of the 9 bases surrounding the poly(A) site, and by deletion of the natural poly(A) site (Conway and

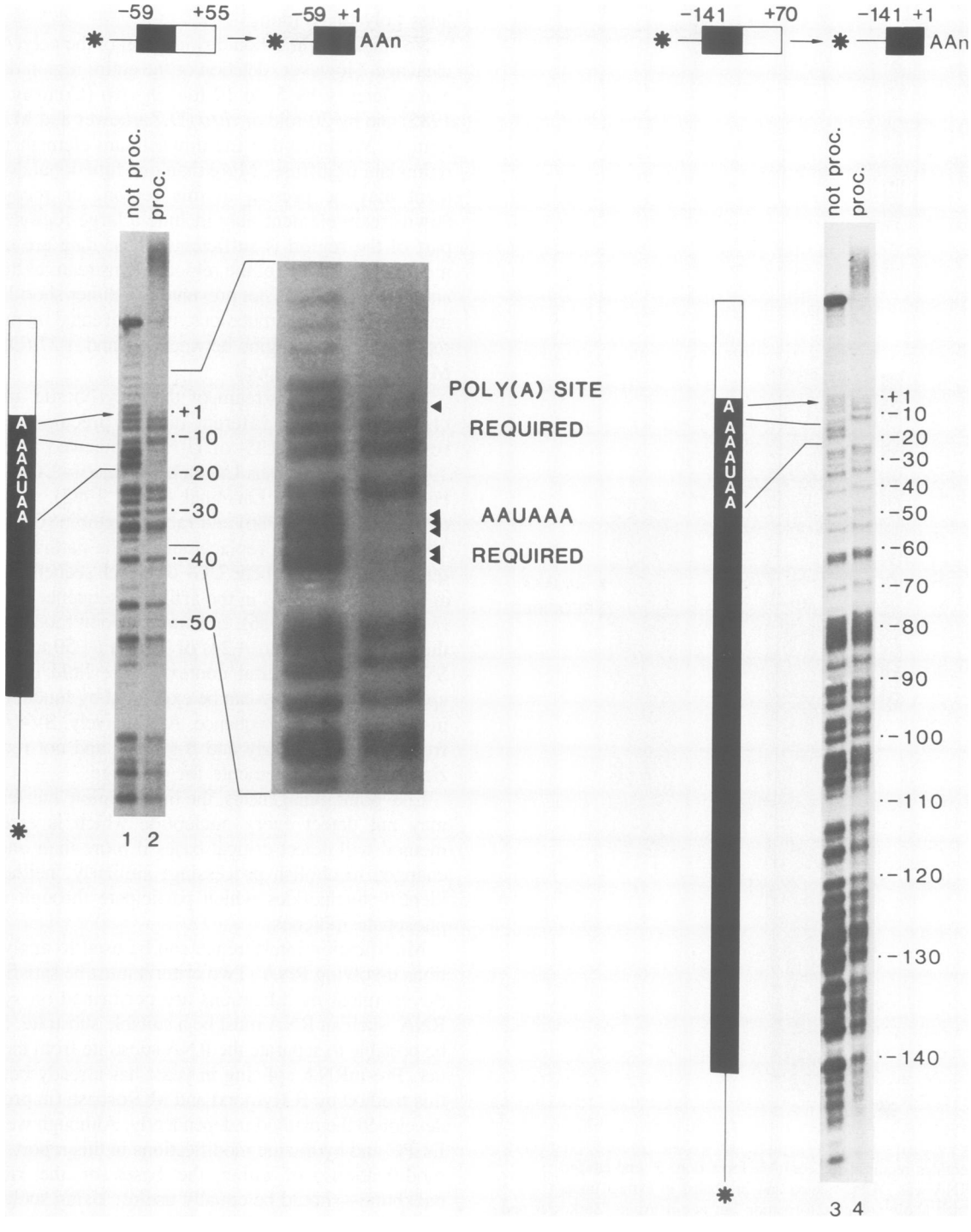

Fig. 5. Nucleotides required for coupled cleavage and polyadenylation. 5′ end-labeled RNA was DEPC-treated and incubated in nuclear extract under conditions which permit sequential cleavage and polyadenylation. Poly(A)$^+$ and poly(A)$^-$ RNAs were separated by oligo(dT) cellulose chromatography. −59/+55 RNA (lanes 1 and 2) or −141/+55 RNA (lanes 3 and 4) were used. **Lanes 1** and **3:** RNA that was not processed in extract [poly(A)$^-$]; **lanes 2** and **4:** RNA that was processed in extract [poly(A)$^+$].

Wickens, 1985; D.Zarkower and M.Wickens, submitted). The modification interference method cannot reliably detect reductions of less than 3-fold. Thus, modification of the +1 A might have less than a 3-fold effect on complex formation; regardless, that position clearly is less significant than AAUAAA in forming the pre-cleavage complex.

We propose that factors in the extract first recognize AAUAAA, then select a nearby NA at which to cleave. This hypothesis is consistent with mutational data; when the natural poly(A) site, a CA dinucleotide, is deleted or replaced with a foreign sequence, cleavage and polyadenylation now occur at a new, 'cryptic' NA dinucleotide (e.g. Fitzgerald and Shenk, 1981; Conway and Wickens, 1985). We cannot assess by modification interference whether the +1 A is essential for cleavage because many RNAs remain unmodified after DEPC or hydrazine treatment, and so generate unmodified, cleaved RNA in the extract. These RNAs are unaffected by aniline treatment and, because they are abundant, obscure bands near +1.

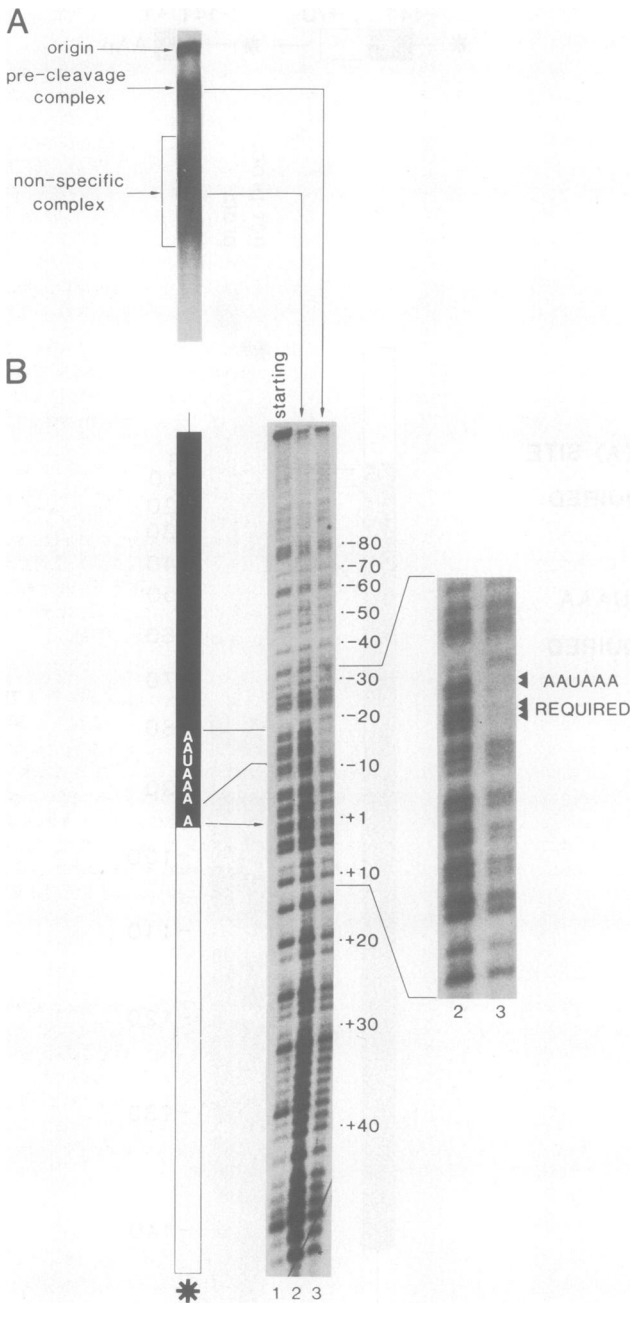

Fig. 6. Nucleotides required for complex formation. 3′ end-labeled −141/+70 RNA was DEPC-treated and incubated in extract under conditions which allow complex formation but not cleavage (Zarkower and Wickens, 1987b). (**A**) To separate specific, pre-cleavage complexes and non-specific complexes, extract reactions (without deproteinization) were subjected to electrophoresis through a non-denaturing polyacrylamide gel. (**B**) RNA was purified from both complexes, treated with aniline and analyzed by electrophoresis. **Lane 1:** starting RNA; **lane 2:** RNA from the non-specific complex; **lane 3:** RNA from the pre-cleavage complex.

Although mutations between AAUAAA and the poly(A) site have little or no effect on polyadenylation of 'pre-cleaved' RNAs (M.Sheets, unpublished data), modification of any of the purines in this region enhances polyadenylation (Figure 4). This enhancement is not seen in coupled cleavage/polyadenylation reactions (Figure 5). We suggest that an extract factor binds to RNA 3′ ends non-specifically; carboxyethylation of adenosines within 10 nucleotides of the 3′ end interferes with this binding and so stimulates polyadenylation.

No single modification downstream of the poly(A) site prevents cleavage. However, deletion of the entire region decreases cleavage efficiency by 5- to 10-fold *in vivo* (Conway and Wickens, 1985) and by 20-fold *in vitro* (D.Zarkower and M.Wickens, submitted). We infer that the downstream element must be either redundant or diffuse. More than one functional downstream element may be present in this region, or recognition of the downstream element may include a large region, such that any part of the region is sufficient. These data are consistent with mutational analyses of the region downstream of the poly(A) site: although cleavage is not prevented by either short linker-scanning mutants or by point mutations, it is prevented by mutations which replace the entire region between +1 and +27 (D.Zarkower and M.Wickens, submitted).

Sequences downstream of the poly(A) site are required for cleavage of many different mRNA precursors. The required regions tend to be UG- or U-rich (reviewed in Birnstiel *et al.*, 1985), and a consensus UG-rich sequence —YGUGUUYY— has been proposed (McLauchlan *et al.*, 1985). With SV40 early (McDevitt *et al.*, 1986) and rabbit β-globin (Gil and Proudfoot, 1987) pre-mRNAs, replacement of the natural downstream sequence by short synthetic UG- or U-rich sequences restores cleavage. Point mutations in the artificial sequences reduce cleavage (McDevitt *et al.*, 1986). In contrast, point modifications in either the UG-rich (+15 to +24) or U-rich (+50 to +55) regions of SV40, in their natural context, have little or no effect. The apparent discrepancy can be explained by functional redundancy in the natural SV40 sequence. Alternatively, SV40 late may differ from both SV40 early and β-globin, and not require either U-rich or UG-rich elements for cleavage.

Like point mutagenesis, the modification interference approach may not detect every nucleotide which is essential. Neither method will detect critical bases if more than one base must be changed to abolish processing; similarly, neither method will identify nucleotides which participate through their ribose or phosphate moieties.

Modification interference can be used to analyze many reactions involving RNA. Two criteria must be satisfied. First, since the chemical modifications are performed on synthetic, naked RNA, such an RNA must be a suitable substrate. Second, it must be possible to separate the RNA substrate from the reaction product. Pre-mRNA splicing in yeast has already been analyzed by this method by B.Rymond and M.Rosbash (in preparation), who developed the method independently. Although we have used only DEPC and hydrazine modifications in this report, other chemical modifications—of either the bases or the ribose-phosphate backbone—should be equally useful. Using such modifications, it should be possible to identify critical features of required bases more precisely, and to identify phosphates which form essential ionic contacts.

Materials and methods

Materials

T4 RNA ligase was purchased from U.S. Biochemical Corp. Calf intestinal alkaline phosphatase, bovine serum albumin, glycogen and proteinase K were purchased from Boehringer Mannheim. T4 polynucleotide kinase, SP6 polymerase and all restriction enzymes were from New England Biolabs. Dimethyl sulfoxide (DMSO), diethylpyrocarbonate (DEPC), oligo(dT) cellulose and yeast RNA were from Sigma. Hydrazine (Baker brand) and aniline (Baker brand) were purchased from VWR. [5′^{32}P]pCp and [γ-^{32}P]ATP were obtained from Amersham.

Aniline was redistilled once under nitrogen and stored at −20°C in the dark. Yeast RNA was purified by proteinase K digestion and repeated phenol extraction. A new bottle of DEPC was opened each week and stored at 4°C. Hydrazine

was stored at room temperature in the hood. To keep it anhydrous, the bottle was stored inside a jar containing Drierite. All manipulations were carried out in silanized microfuge tubes. Capillary pipettes were silanized before use. Solutions were made using DEPC-treated water.

Construction of templates

To construct pSPSV −141/+70, a portion of SV40 containing 141 bases upstream and 70 bases downstream of the late poly(A) site (from the natural *Bam*HI site to a synthetic *Hin*dIII site) was cloned into pSP65. A synthetic *Xba*I linker was inserted at the *Hin*dIII site. pSPSV −59/+70 contains SV40 sequence cloned into the polylinker of pSP64. The polylinker is present from the SP6 promoter to the *Sma*I site. The SV40 sequence begins there and extends from the SV40 *Sfa*NI site at −59 to the synthetic *Hin*dIII sequence at +70. *Dra*I cuts this plasmid at +55 within the SV40 sequence. pSPSV −59/+1 is identical to pSPSV −59/+70 upstream of the poly(A) site. It contains a *Hae*II linker at the poly(A) site. After digestion with *Hae*II, transcription yields an RNA that ends at the poly(A) site. pIBI76 sequence lies downstream of the poly(A) site beginning at the *Sma*I site and extending to the *Hin*dIII site.

Preparation of RNA

Transcription. DNA templates were linearized with *Xba*I (pSPSV −141/+70), *Dra*I (pSPSV −59/+70) or *Hae*II (pSPSV −59/+1). Transcription was performed as described (Melton *et al.*, 1984) except no RNasin was used. Diguanosine triphosphate was included in the transcription reaction when −141/+70 RNA was prepared for use in assaying complex formation. All other precursors possess polyphosphate termini, not caps. Transcripts were purified by polyacrylamide gel electrophoresis before end-labeling.

5′ end-labeling. To 5′ end-label the RNA, 5−10 pmol purified RNA were incubated in 5 mM Tris−HCl, pH 9 with 20 units calf intestinal alkaline phosphatase for 1 h at room temperature. The RNA was extracted with phenol/chloroform (50:50, v/v), ethanol precipitated and redissolved in 7 μl water. 2 μl 10 × T4 polynucleotide kinase buffer (0.5 M Tris−HCl, pH 7.6, 0.1 M $MgCl_2$, 50 mM dithiothreitol, 1 mM spermidine, 1 mM EDTA; Maniatis *et al.*, 1982), 1 μl (10 units) T4 polynucleotide kinase and 10 μl (100 μCi) [γ-^{32}P]ATP were added to the RNA. The reaction mix was incubated at 37°C for 1 h. Full length end-labeled RNA was purified by polyacrylamide gel electrophoresis.

3′ End-labeling (Romaniuk and Uhlenbeck, 1983). 5—10 pmol RNA were ethanol precipitated and redissolved in: 2 μl of 250 μM ATP, 2 μl 0.1 mg/ml bovine serum albumin, 2 μl 10X T4 RNA ligase buffer (0.5 M Tris, pH 7.6, 0.15 M $MgCl_2$, 33 mM dithiothreitol), 2 μl DMSO, 2 μl (8−20 units) T4 RNA ligase and 10 μl (100 μCi) [5′^{32}P]pCp. The reaction mix was incubated at 4°C for at least 10 h. End-labeled RNA was purified by polyacrylamide gel electrophoresis.

Eluting RNA from sequencing gels (Peattie, 1979). RNA was eluted from sequencing gels by soaking the gel slice in 0.5 M ammonium acetate, pH 6.5, 0.1% SDS, 0.1 mM EDTA for at least 2 h at 37°C in the presence of either glycogen (for unlabeled RNA) or yeast RNA (for end-labeled RNA) as carrier.

Chemical modification and aniline cleavage

Modification and aniline cleavage were performed essentially as described by Peattie (1979), but are here described to provide relevant changes and technical details. All centrifugations were performed in a Brinkmann microcentrifuge (13 000 *g*).

Purine modifications. 1 × 10^6 c.p.m. end-labeled RNA was ethanol precipitated with 12.5 μg carrier RNA. The RNA was redissolved in 200 μl of 50 mM sodium acetate, pH 4.5, 1 mM EDTA (pH adjusted with acetic acid). Two microliters fresh DEPC were added. The sample was vortexed and incubated at 90°C for 2.5 min. To stop the reaction, 75 μl 1 M sodium acetate, pH 4.5 and 750 μl 100% ethanol were added. The sample was centrifuged for 20 min at 0°C to pellet RNA. The pellet was redissolved in 200 μl 0.3 M sodium acetate, pH 3.8. 600 μl ethanol were added and the sample was centrifuged again. The pelleted RNA was redissolved in 27 μl water.

Pyrimidine modifications—U only. End-labeled RNA, 1 × 10^6 c.p.m., was ethanol precipitated in the presence of 12.5 μg carrier RNA. The pellet was rinsed twice with 75% ethanol and resuspended in 10 μl water. 10 μl anhydrous hydrazine were added. The RNA was incubated on ice for 10 min. To stop the reaction, 200 μl 0.3 M sodium acetate, pH 3.8 and 750 μl ethanol were added. The RNA was pelleted and redissolved in 200 μl 0.3 M sodium acetate. 600 μl ethanol were added and the RNA was pelleted again. The RNA was redissolved in 27 μl water.

Pyrimidine modifications—U + C. This reaction is only useful with 3′ end-labeled RNA. It consistently yields unreadable 'ladders' when 5′ end-labeled RNA is used. 3′ End-labeled RNA, 1 × 10^6 c.p.m., was ethanol precipitated with 12.5 μg carrier. The RNA was redissolved in 20 μl anhydrous hydrazine/0.5 M NaCl. The reaction mix was incubated on ice for 30 min. To stop the reaction, 200 μl 0.3 M NaOAc, pH 3.8 and 750 μl ethanol were added. The sample was spun for 20 min at 0°C to pellet the RNA. The RNA was redissolved in 200 μl 0.3 M NaOAc, pH 3.8, 600 μl ethanol were added and the RNA was pelleted again. The RNA was redissolved in 27 μl water.

Cleavage of modified RNA with aniline. Purified RNA was ethanol precipitated and redissolved in 20 μl 1 M aniline that had been diluted in 0.3 M sodium acetate, pH 3.8. The sample was incubated at 60°C for 20 min in the dark. To stop the reaction, 1.4 ml of *n*-butanol was added. The tube was vortexed and centrifuged for 10 min to pellet the RNA. The pellet was redissolved in 150 μl 1% SDS. Butanol, 1.3 ml, was added. The tube was vortexed and centrifuged for 10 min. The pellet was rinsed with 1 ml 100% ethanol, dried and resuspended in 5 μl loading buffer containing 8 M urea, 20 mM Tris, pH 7.9, 1 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue (Peattie, 1979). The sample was boiled for 90 s, put on ice immediately and loaded onto a sequencing gel.

Preparation and use of nuclear extract

HeLa cell nuclear extract was prepared as described (Dignam *et al.*, 1983) except that $MgCl_2$ was omitted from buffer A, phenylmethylsulfonyl fluoride (PMSF) was omitted from all buffers, 0.125 mM EDTA was used in buffer D and the extract was dialyzed against buffer D for 3.5 h. The extract was then frozen without centrifugation.

Cleavage and polyadenylation reactions were performed as described by Moore and Sharp (1985). 25 μl modified RNA were incubated in a 125 μl reaction containing 55 μl nuclear extract, 40 mM KCl, 20 mM phosphocreatine, 0.25 mM ATP and 2.8% polyvinyl alcohol. 0.5 mM EDTA was included when necessary to prevent polyadenylation and degradation of the 3′ half molecules. RNA was isolated from the reaction as described by Krainer *et al.* (1984). Processed and unprocessed RNAs were separated either by polyacrylamide gel electrophoresis or by oligo(dT) cellulose chromatography (Aviv and Leder, 1972).

Analysis of pre-cleavage complexes

300 fmol of capped, 3′ end-labeled −141/+70 RNA were incubated in a 75 μl reaction containing 33 μl nuclear extract and 40 mM KCl. After 10 min at 30°C, heparin was added to a final concentration of 5 mg/ml and the sample incubated for 10 min at 0°C. The sample was loaded onto a 1.5 mm thick 4% acrylamide:0.05% bis-acrylamide non-denaturing gel and electrophoresed at 18 V/cm for 3 h. Pre-cleavage and non-specific complexes were detected by autoradiography. The gel slice containing each complex was excised, inserted into a single horizontal lane of a 6% sequencing gel and electrophoresed at 40 mA for 2 h. The region of the gel containing the RNA was excised and the RNA eluted as described (Peattie, 1979).

Acknowledgements

We thank Brian Rymond and Michael Rosbash for their advice and communication of results prior to publication, and Jim Dahlberg for reading the manuscript. We thank the members of the Wickens laboratory for helpful discussions, and Steve Ogg and Ray Smith for technical assistance. L.C. is supported by a NSF graduate fellowship and a Peterson Fellowship from the University of Wisconsin. This work is supported by the NIH (GM31892), a Basil O'Conner Starter Research Grant from the March of Dimes and a Steenbock Career Development Award to M.W.

References

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA.*, **69**, 1408−1412.
Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349−359.
Brown,D.M. (1967) *Methods Enzymol.*, **12**, 31−34.
Cashmore,A.R. and Peterson,G.B. (1978) *Nucleic Acids Res.*, **5**, 2485−2491.
Conway,L. and Wickens,M. (1985) *Proc. Natl. Acad. Sci. USA.*, **82**, 3949−3953.
Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475−1488.
Fitzgerald,M. and Shenk,T. (1981) *Cell*, **24**, 251−260.
Gil,A. and Proudfoot,N.J. (1987) *Cell*, **49**, 399−406.
Hashimoto,C. and Steitz,J.A. (1986) *Cell*, **45**, 581−591.
Henderson,R.E.L., Kirkegaard,L.H. and Leonard,N.J. (1973) *Biochim. Biophys. Acta*, **294**, 356−364.
Herr,W., Chapman,N.M. and Noller,H.F. (1979) *J. Mol. Biol.*, **130**, 433−449.
Higgs,D.R., Goodbourn,S.E.Y., Lamb,J., Clegg,J.B., Weatherall,D.J. and Proudfoot,N.J. (1984) *Nature*, **306**, 398−400.
Humphrey,T., Christofori,G., Lucijanic,V. and Keller,W. (1987) *EMBO J.*, **6**, 4159−4168.
Krainer,A.R., Maniatis,T., Ruskin,B. and Green,M.R. (1984) *Cell*, **36**, 993−1005.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
Manley,J.L., Yu,H. and Ryner,L. (1985) *Mol. Cell. Biol.*, **5**, 373−379.

McDevitt,M.A., Hart,R.P., Wong,W.W. and Nevins,J.R. (1986) *EMBO J.*, **5**, 2907–2913.
McLauchlan,J., Gaffney,D., Whitten,J.L. and Clements,J.B. (1985) *Nucleic Acids Res.*, **13**, 1347–1369.
Melton,D., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Montell,C., Fisher,E.F., Caruthers,M.H. and Berk,A.J. (1983) *Nature*, **305**, 600–605.
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.
Nevins,J.R. and Darnell,J.E. (1978) *Cell*, **15**, 1477–1493.
Nevins,J.R. (1983) *Annu. Rev. Biochem.*, **52**, 441–466.
Peattie,D.A. (1979) *Proc. Natl. Acad. Sci. USA.*, **76**, 1760–1764.
Peattie,D.A. and Herr,W. (1981) *Proc. Natl. Acad. Sci. USA.*, **78**, 2273–2277.
Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211–214.
Romaniuk,P.J. and Uhlenbeck,O.C. (1983) *Methods Enzymol.*, **100**, 52–59.
Sheets,M.D., Stephenson,P. and Wickens,M. (1987) *Mol. Cell. Biol.*, **7**, 1518–1529.
Siebenlist,U. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 122–126.
Skolnik-David,H., Moore,C.L. and Sharp,P.A. (1987) *Genes and Development*, **1**, 672–682.
Turchinskii,M.F., Gus'kova,L.I., Khazai,I., Budovskii,E.I. and Kochetov,N.K. (1969) *Mol. Biol. (USSR)*, **4**, 343–348.
Vincze,A., Henderson,R.E.L., McDonald,J.J. and Leonard,N. (1973) *J. Am. Chem. Soc.*, **95**, 2677–2682.
Wickens,M.P. and Stephenson,P. (1984) *Science*, **226**, 1045–1056.
Zarkower,D. and Wickens,M. (1987a) *EMBO J.*, **6**, 177–186.
Zarkower,D. and Wickens,M. (1987b) *EMBO J.*, **6**, 4185–4192.
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.

Received on October 6, 1987