The EMBO Journal vol.6 no.13 pp.4185–4192, 1987

Specific pre-cleavage and post-cleavage complexes involved in the formation of SV40 late mRNA 3′ termini *in vitro*

David Zarkower and Marvin Wickens

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

Communicated by W.Keller

Complexes form between processing factors present in a crude nuclear extract from HeLa cells and a simian virus 40 (SV40) late pre-mRNA which spans the polyadenylation [poly(A)] site. A specific 'pre-cleavage complex' forms on the pre-mRNA before cleavage. Formation of this complex requires the highly conserved sequence AAUAAA: it is prevented by mutations in AAUAAA, and by annealing DNA oligonucleotides to that sequence. After cleavage, the 5′ half-molecule is found in a distinct 'post-cleavage complex'. In contrast, the 3′ half-molecule is released. After cleavage and polyadenylation, polyadenylated RNA also is released. *De novo* formation of the post-cleavage complex requires AAUAAA and a nearby 3′ terminus. Competition experiments suggest that a component which recognizes AAUAAA is required for formation of both pre- and post-cleavage complexes.

Key words: complexes/downstream element/polyadenylation/mRNA processing

Introduction

Two sequential reactions are required to form mRNA 3′ termini in animal cells: endonucleolytic cleavage at the polyadenylation [poly(A)] site of the pre-mRNA to form a new 3′ terminus, and the addition of 250 adenosine residues to the newly formed end (reviewed in Birnstiel *et al.*, 1985). The cleavage reaction requires two sequences in the pre-mRNA: the sequence AAUAAA, which lies upstream of the poly(A) site, and sequence element(s) which lie downstream. The AAUAAA sequence is more highly conserved during evolution and more strictly required than the downstream element(s) (Wickens and Stephenson, 1984; Conway and Wickens, 1985). AAUAAA also is required for polyadenylation of a synthetic RNA which ends at the poly(A) site (Manley, 1983; Zarkower *et al.*, 1986).

Prior to cleavage, a complex forms between processing components present in a crude HeLa cell nuclear extract and the mRNA precursor (Zhang and Cole, 1987; Skolnik-David *et al.*, 1987). This pre-cleavage complex migrates more slowly than free RNA in a non-denaturing polyacrylamide gel. Thus, as in splicing (Konarska and Sharp, 1986; Pikielny and Rosbash, 1986), 'gel retardation' of labeled pre-mRNA provides a simple and direct assay for the formation of complexes involved in mRNA processing.

After cleavage, a stable complex between the AAUAAA sequence of cleaved RNA and extract factor(s) can be detected using an oligonucleotide/RNase H protection assay (Zarkower and Wickens, 1987). This post-cleavage complex is observed only on cleaved RNAs bearing a 3′ terminal 3′ deoxyadenosine, and may reflect the stable interaction of the presumptive poly(A) polymerase with the AAUAAA sequence and the 3′ end of cleaved RNA. However, a component required to form this complex is also required for cleavage (Zarkower and Wickens, 1987), suggesting that the complex forms initially with the mRNA precursor and that at least part of it remains associated with the RNA after cleavage.

In this report, we identify and characterize pre-cleavage and post-cleavage complexes using a 'gel retardation' assay. We demonstrate, by several criteria, the involvement of these complexes in mRNA 3′ end formation. Finally, we suggest that cleavage occurs in the pre-cleavage complex and that the post-cleavage complex persists until polyadenylation.

Results

RNA substrates

The RNAs used as substrates were prepared by transcribing suitable plasmids with SP6 RNA polymerase (Melton *et al.*, 1984). Each RNA contains a portion of simian virus 40 (SV40) sequence near the 'late' (virion protein) mRNA poly(A) site. RNAs are named by the bases of SV40 they contain: for example, −141/+55 RNA contains SV40 sequences from 141 bases upstream of the poly(A) site to 55 bases downstream. The structures of RNA substrates are described in Materials and methods, and the sequence from −141 to +70 is presented in the accompanying paper (Conway and Wickens, 1987).

Pre-cleavage and post-cleavage complexes

Electrophoresis in non-denaturing gels (Dahlberg *et al.*, 1969) provides a direct assay for the binding of factors present in a HeLa cell nuclear extract to SV40 late pre-mRNA (Figure 1). Labeled −141/+55 RNA was added to nuclear extract containing 3′dATP to prevent polyadenylation (Moore and Sharp, 1985). At various times afterwards, two aliquots of the reaction were removed. One was used to assay the formation of complexes: it was loaded directly, without deproteinization, onto a polyacrylamide gel containing non-denaturing buffer (Figure 1A). The other was used to determine the extent of cleavage: it was deproteinized with phenol and the RNA analyzed by electrophoresis through a polyacrylamide gel containing 7 M urea (Figure 1B).

Immediately after the RNA is added to extract (Figure 1A, 0 min), a complex forms (designated complex A) which migrates slower than naked RNA (Figure 1A, no ext.). This complex forms on any added RNA substrate, even a prokaryotic RNA (not shown) and so is termed 'non-specific'. One minute later, before cleavage has occurred (Figure 1B), a new complex forms, designated complex B (Figure 1A). The patterns of complexes and RNA do not change between 1 and 2 min. By 5 min, 10% of the RNA has been cleaved, and a very small amount of a new complex appears that migrates slightly faster than B. The amount of this new complex, designated B′, progressively increases at 8, 10, 30 and 60 min, with concomitant decreases in complex B. After 1 h, only complex B′ is detectable. The increase in complex B′ correlates well with the increase in the extent of cleavage, which rises from 10% at 5 min to more than 90% after 1 h.

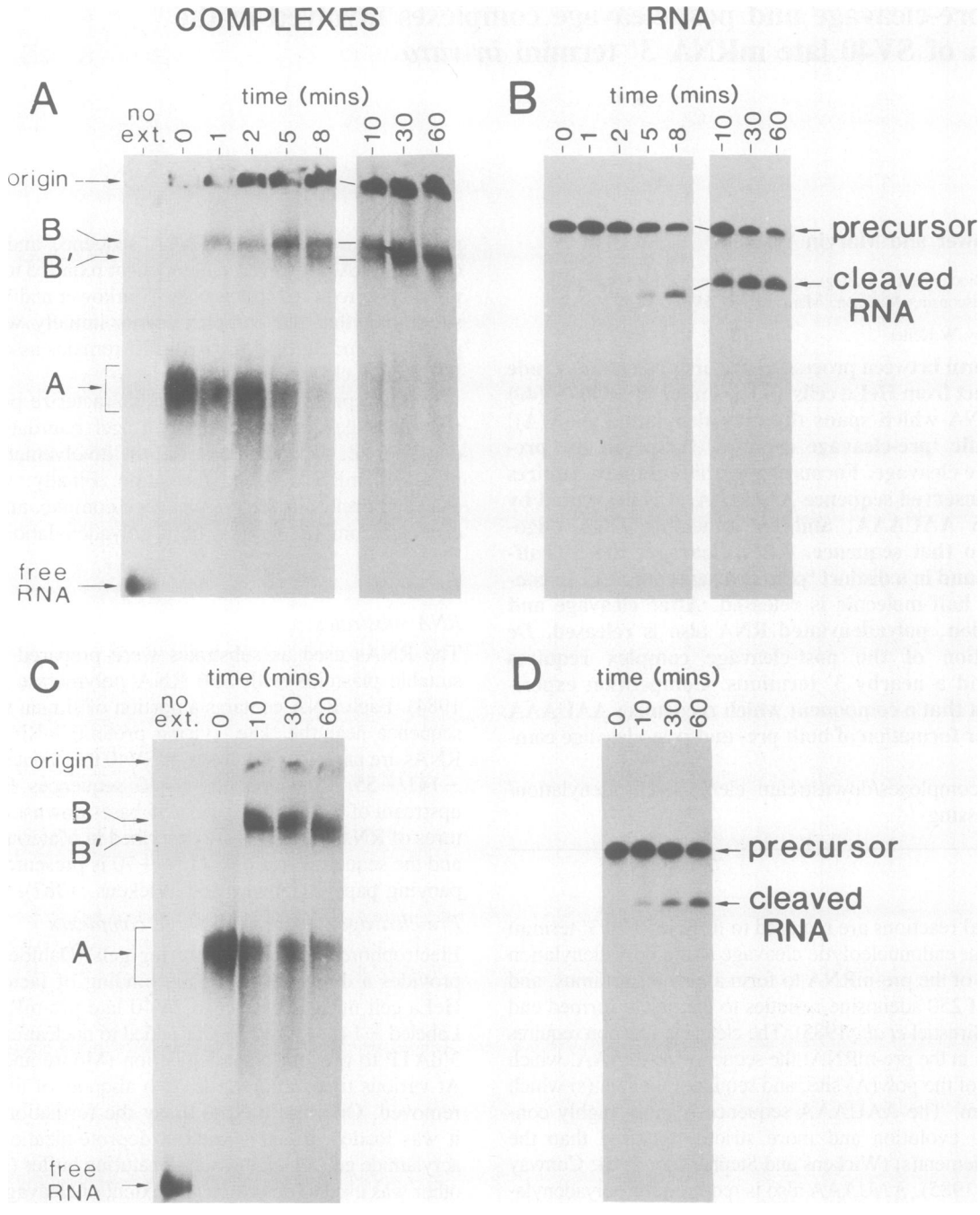

Fig. 1. Pre-cleavage and post-cleavage complexes. **(A)** Time course of complex formation under standard conditions. Each reaction contained uniformly-labeled −141/+55 RNA (2.5 fmol) incubated in nuclear extract under conditions optimal for cleavage (see Materials and methods). Reactions were started at staggered times and stopped simultaneously by addition of heparin. Five microliters of each reaction were loaded directly onto a non-denaturing gel and subjected to electrophoresis. The time of incubation prior to heparin addition is given above each lane. The lane marked 'no ext.' contains naked −141/+55 RNA in the absence of extract. Zero to eight minute and 10−60 min time points are from two separate experiments. **(B)** Time course of cleavage under standard conditions. Five microliters of each reaction in (A) were deproteinized and analyzed by electrophoresis through a denaturing 6% polyacrylamide gel. The positions of −141/+55 RNA ('precursor') and 5′ half-molecules ('cleaved RNA') are indicated. **(C)** Time course of complex formation in the absence of PVA. −141/+55 RNA (2.5 fmol) was incubated in nuclear extract with PVA and with 3′ dATP. Reactions were otherwise performed as in (A). **(D)** Time course of cleavage in the absence of PVA. Five microliters of each reaction in (C) were deproteinized and analyzed by electrophoresis through a denaturing 6% polyacrylamide gel.

Thus, whereas complex B appears to form prior to cleavage, complex B′ is detected only after.

To corroborate these results, we assayed complex formation in reactions from which polyvinyl alcohol (PVA) was omitted (Figure 1C and D). In the absence of PVA, the rate of cleavage is decreased more than 10-fold: even at 30 min, less than 10% of the RNA has been cleaved (Figure 1D). Complex B forms within 10 min (Figure 1C), when cleavage is barely detectable (Figure 1D).

The data in Figure 1 strongly suggest that complex B contains mRNA precursor and that B′ contains the 5′ half-molecule (i.e. 'cleaved RNA'). To test this prediction, two-dimensional gels were used to examine a reaction containing both B and B′ (Figure 2A). In the first dimension, complexes were separated by electro-

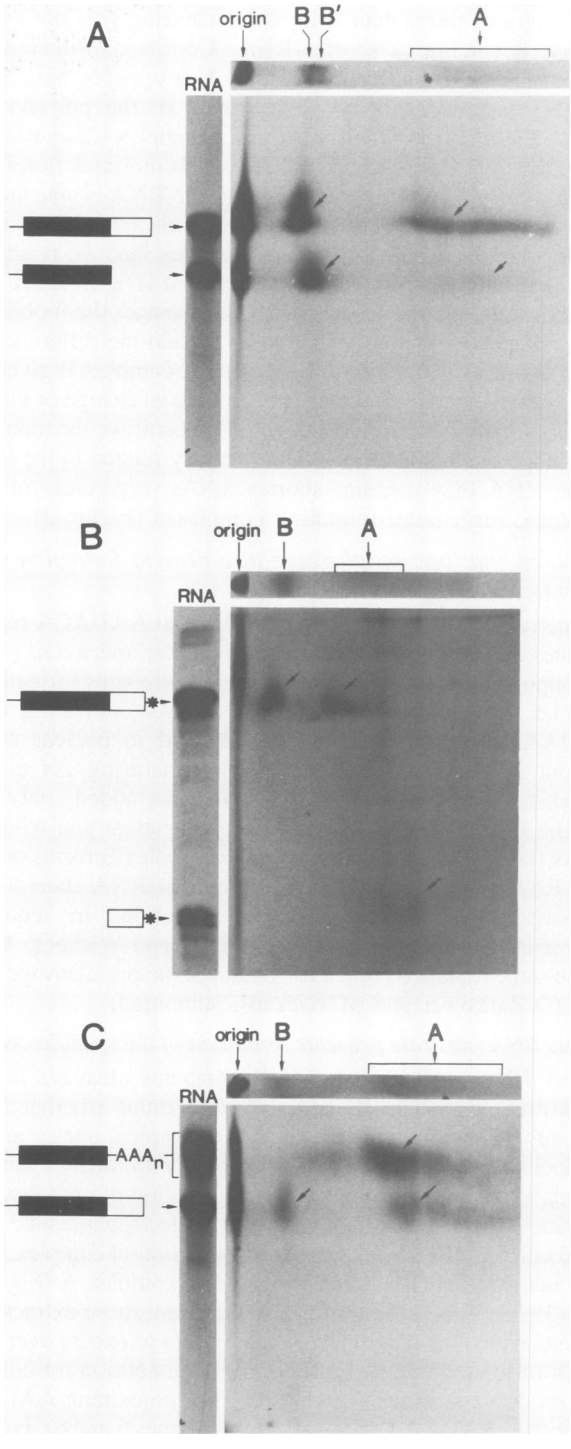

Fig. 2. Analysis of complexes on two-dimensional gels. (**A**) Uniformly-labeled −141/+55 RNA (5 fmol) was incubated in nuclear extract containing 3′ dATP for 15 min under optimal conditions for cleavage (see Materials and methods) in a total volume of 12.5 μl. The reaction was stopped by 10 min incubation at 0°C with heparin and two-dimensional electrophoresis was performed using 5 μl of the reaction. **First dimension** (left to right): gel containing 45 mM Tris−borate, 0.6 mM EDTA, no denaturant; **second dimension** (top to bottom): gel containing 90 mM Tris−borate, 1.2 mM EDTA and 7 M urea. For alignment purposes, a strip of a non-denaturing gel run in parallel is shown above each gel. Similarly, an aliquot of deproteinized RNA was analyzed on the same second dimension gel as the complexes, and is presented to the left of each gel. The structure and position of each RNA is depicted at the left. (**B**) As in (A), except that reactions contained 1 mM EDTA instead of 3′ dATP and the RNA was 3′ end-labeled −141/+70 RNA (10 fmol). (**C**) As in (A), except that 3′ dATP was omitted; instead, only ATP was present.

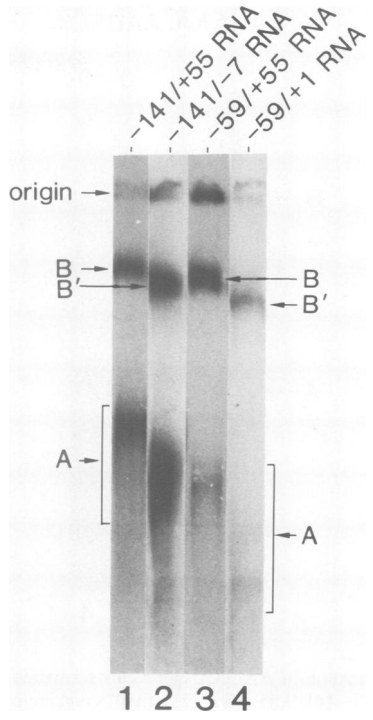

Fig. 3. Formation of post-cleavage complex *de novo*. Each RNA (2.5 fmol) was incubated in nuclear extract without PVA. After 10 min of incubation, reactions were stopped and 5 μl of each reaction analyzed by electrophoresis through a non-denaturing gel. The amount of radioactivity present in each lane varies because 2.5 fmol of each RNA was used and the RNAs vary in length, but are identical in specific activity. The length of each RNA is: −141/+55 RNA, 225 nucleotides; −141/−7 RNA, 164 nucleotides; −59/+55 RNA, 155 nucleotides; −59/+1 RNA, 100 nucleotides.

phoresis through a non-denaturing gel; in the second, the RNA present in each complex was separated by electrophoresis through a gel containing 7 M urea. As predicted, B contains only the precursor and B′ contains only cleaved RNA (Figure 2A). Complex A contains primarily the precursor, and a small amount of cleaved RNA. Material which remains at the origin of the first dimension gel contains a mixture of both RNAs. The 3′ half-molecule generated by cleavage is not detectable since it is degraded rapidly under these reaction conditions.

To determine whether complex B also contains the 3′ half molecule, we incubated 3′ end-labeled −141/+70 RNA in extract containing EDTA rather than 3′ dATP (Figure 2B). The use of 3′ end-labeled RNA simplifies detection of 3′ half-molecules, while EDTA prevents their degradation and inhibits polyadenylation (Moore and Sharp, 1985; Zarkower *et al.*, 1986). The RNA species present in each complex was identified on two-dimensional gels. As expected, precursor is in complex B, and the 5′ half-molecule is not detectable, since it is not labeled. The 3′ half-molecule is found in a complex that migrates slightly faster than complex A, not in complex B or B′ (Figure 2B). The mobility of non-specific complexes is dependent on the length of the RNA they contain—shorter RNAs form faster migrating complexes (Konarska and Sharp, 1986; Figure 3). We therefore propose that the 3′ half-molecule is released after cleavage and then forms a non-specific complex; this complex migrates faster than that formed on the precursor simply because the 3′ half-molecule is shorter.

To determine whether the 5′ half-molecule remains in complex B′ after it has been polyadenylated, −141/+55 RNA was incubated in nuclear extract under conditions which permit

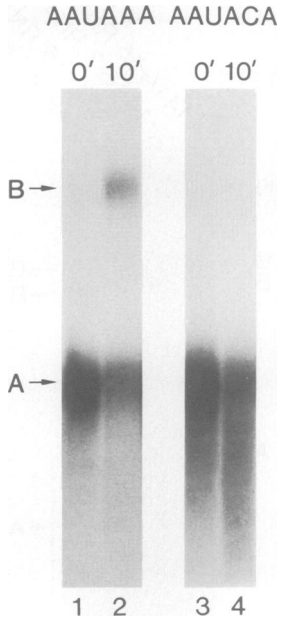

Fig. 4. A point mutation in AAUAAA prevents formation of the pre-cleavage complex. −141/+55 RNA (2.5 fmol) containing either AAUAAA (**lanes 1** and **2**) or AAUACA (**lanes 3** and **4**) was incubated in nuclear extract and analyzed by non-denaturing gel electrophoresis after 0 min or 10 min.

sequential cleavage and polyadenylation, and then was analyzed on two-dimensional gels (Figure 2C). RNA which has been cleaved and polyadenylated is heterogeneous in length due to variation in the length of the poly(A) tails. In the first dimension, the polyadenylated products migrate faster than complex B, and nearly comigrate with complex A formed on the precursor. Either polyadenylated RNAs are released as naked RNA and form non-specific complexes, or they form a new, specific complex with accelerated mobility. In either case, it is clear that neither B nor B′ persist intact after polyadenylation.

In most subsequent experiments, complex formation is assayed using reactions which lack PVA, for two reasons. (i) Without PVA, virtually all the labeled RNA enters the non-denaturing gel. In contrast, with PVA present, a variable amount of material fails to enter (Figure 1). Although this trapping at the origin probably is due only to non-specific aggregation (Figure 2), its variability complicates comparisons between samples. (ii) Complex B′ forms much later in reactions without PVA (Figure 1), making it simpler to assay complex B specifically.

To summarize the data thus far, complex B forms prior to cleavage and contains only the precursor. After cleavage, complex B′ is detected. It contains the 5′ half-molecule but not the 3′ half-molecule, which apparently has been released. After polyadenylation, the 5′ half-molecule, now bearing poly(A), leaves this complex. We will refer to complex B as a pre-cleavage complex, and to B′ as a post-cleavage complex.

Formation of post-cleavage complexes de novo

To test whether formation of complex B′ requires cleavage *per se*, we synthesized, by transcription *in vitro*, a 'pre-cleaved' −141/−7 RNA (Figure 3). This transcript closely resembles RNA which has been cleaved, and is a substrate for polyadenylation, but not for cleavage (Zarkower *et al.*, 1986). Upon addition to nuclear extract containing 3′ dATP, it forms complex B′ (lane 2), whereas −141/+55 RNA, analyzed in parallel, forms complex B (lanes 1 and 2). Thus B′ can form *de novo*.

B′ migrates faster than B in non-denaturing gels not merely because it contains a shorter RNA. Complex formation was assayed using RNAs that extend to +1 or to +55, but contain only 59 nucleotides of SV40 upstream of the poly(A) site. −59/+1 RNA (lane 4) forms a specific complex (designated B′, as it forms on a pre-cleaved RNA) which migrates faster than that of −59/+55 RNA (lane 3). −59/+55 RNA is nine nucleotides shorter than −141/−7 RNA, yet forms a specific complex which migrates more slowly (lane 3 versus lane 2), nearly comigrating with the complex formed on −141/+55 RNA (lane 1). Thus although the length of RNA influences the mobility of these complexes, it is insufficient to explain the difference between B and B′. Rather, we propose that complex B gives rise to complex B′ either through a conformational change or through the loss or gain of some component(s). In contrast, the mobilities of non-specific complexes (A) are directly related to the length of the RNA they contain: shorter RNAs, regardless of their sequence, form faster-migrating complexes (Figure 3).

Mutations that prevent cleavage also prevent formation of the pre-cleavage complex

A transversion which changes AAUAAA to AAUACA reduces the rate of cleavage 10- to 20-fold *in vitro* (Zarkower *et al.*, 1986, and unpublished data). This mutation also prevents formation of complex B (Figure 4). −141/+55 RNAs containing either AAUAAA or AAUACA were incubated in nuclear extract without PVA, and analyzed by non-denaturing gel electrophoresis. Immediately after the RNA was added, only non-specific complexes are detected on both mutant and wild type RNAs (lanes 1 and 3). Within 10 min complex B forms on wild type RNA (lane 2), but not on the mutant RNA (lane 4).

Cleavage also is prevented by mutations in sequences downstream of the poly(A) site (Conway and Wickens, 1985). These same mutations block formation of the pre-cleavage complex (D.Zarkower and M.Wickens, submitted).

Excess RNA substrate prevents formation of the specific complex

Excess RNA substrate specifically prevents cleavage *in vitro* (Zarkower and Wickens, 1987). To determine whether it also prevents formation of the pre-cleavage complex, a trace amount of labeled −141/+55 RNA was mixed with various amounts of unlabeled competitor RNA. The formation of complex B on the labeled RNA was then assayed by electrophoresis.

Excess −141/+55 RNA reduces formation of complex B, but does not significantly affect formation of complex A (Figure 5, lanes 1−4). This indicates that, as expected, those extract components which form complex A are present in excess over those which form complex B. Furthermore, competition for complex B is sequence-specific: −141/+55 RNA containing AAUACA rather than AAUAAA does not compete (lanes 5 and 6), nor does a 232 nucleotide RNA containing only prokaryotic sequences (lanes 7 and 8).

RNA in which the SV40 sequence downstream of the poly(A) site has been changed to a 32-nt prokaryotic sequence is a poor competitor (lane 9). This RNA also is cleaved inefficiently and forms little complex (D.Zarkower and M.Wickens, submitted). In contrast, pre-cleaved RNA, which lacks *any* nucleotides downstream of the poly(A) site, competes effectively for complex formation (lanes 10 and 11). This competition also requires AAUAAA: it is abolished by a point mutation in that sequence (lanes 12 and 13).

In summary, the competition data confirm that AAUAAA and the downstream element are required for formation of the pre-cleavage complex, and that formation of the complex is saturable.

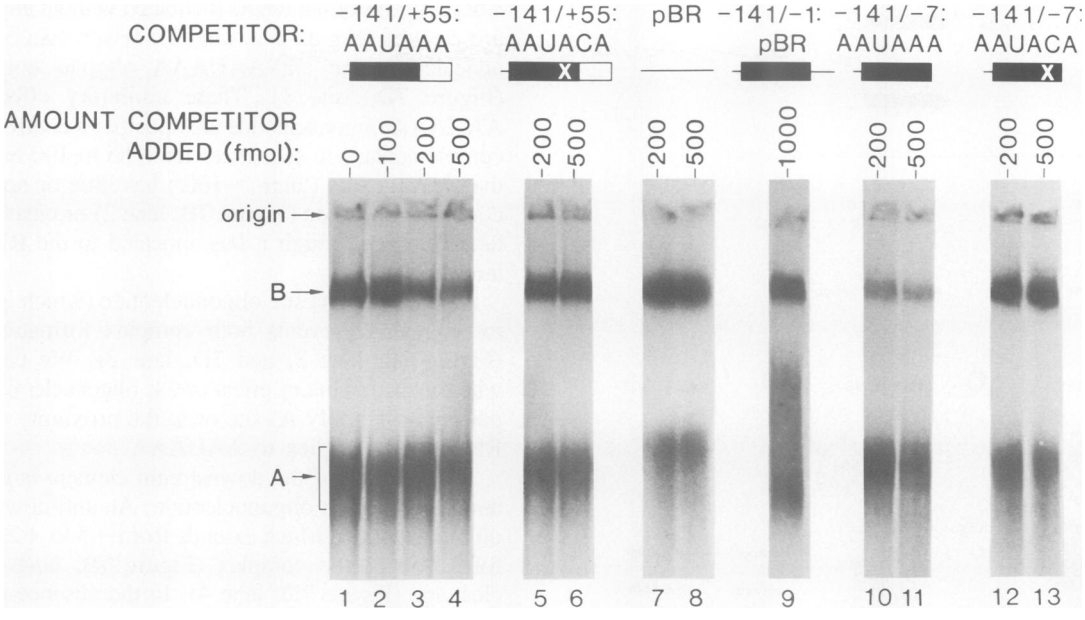

Fig. 5. Excess RNA substrate prevents formation of specific complexes. Labeled −141/+55 RNA (2.5 fmol) was mixed with unlabeled competitor RNAs. The amount (fmol) of competitor added is indicated above each lane. After 15 min of incubation, complex formation was assayed. Competitors are −141/+55 RNA (**lanes 1–4**); −141/+55:AAUACA RNA (**lanes 5** and **6**); a 232-nt-long RNA containing only vector-derived sequences ('pBR'; **lanes 7** and **8**); a −141/−1 RNA with 32 bases of prokaryotic sequence at the 3′ terminus ('−141/−1:pBR'; **lane 9**); −141/−7 RNA (**lanes 10** and **11**); and −141/−7:AAUACA RNA (**lanes 12** and **13**). The structure of each RNA is illustrated; a white X indicates the mutation changing AAUAAA to AAUACA. The structure and preparation of these RNAs is described in Materials and methods.

Excess cleaved RNA, which forms only the post-cleavage complex, prevents formation of the pre-cleavage complex. From this we infer that formation of both complexes requires a common component.

Dissociation of the pre-cleavage complex

To examine whether the pre-cleavage complex is stable, we determined whether it can exchange between pre-mRNA molecules (Figure 6). Complexes were allowed to form on labeled −141/+55 RNA for 15 min. Excess unlabeled competitor RNA, containing either AAUAAA or AAUACA, was then added and the incubation continued for 15 min. For comparison, both competitors were also added simultaneously with the labeled RNAs, at the beginning of the incubation. The amount of complex present on the labeled RNA after 30 min was determined by electrophoresis.

In the absence of any competitor, the expected pattern of complexes A and B is observed (lane 1). Similarly, as expected, excess unlabeled −141/+55 RNA added simultaneously with the labeled RNA prevents formation of complex B on the labeled RNA (lane 2), but only if the unlabeled RNA contains AAUAAA (lane 3). Strikingly, the same results are obtained even if the unlabeled RNA is added 15 min after the labeled RNA (lanes 4 and 5): complexes already formed on the labeled RNA dissociate and are sequestered on the unlabeled RNA as long as it contains AAUAAA. This result demonstrates that the pre-cleavage complex readily dissociates from one RNA and forms a complex with another. We suggest that, under the conditions used here, the pre-cleavage complex may not be stable in solution.

DNA oligonucleotides base-paired to AAUAAA, but not to the downstream element, prevent formation of specific complex

In the following experiments, we examine whether hybridization of a DNA oligonucleotide to various regions of the pre-mRNA interferes with complex formation: the general protocol is outlined in Figure 7A. A DNA oligonucleotide is hybridized to labeled −141/+55 RNA and the heteroduplex added to the extract. EDTA is included to inhibit RNase H (Zarkower and

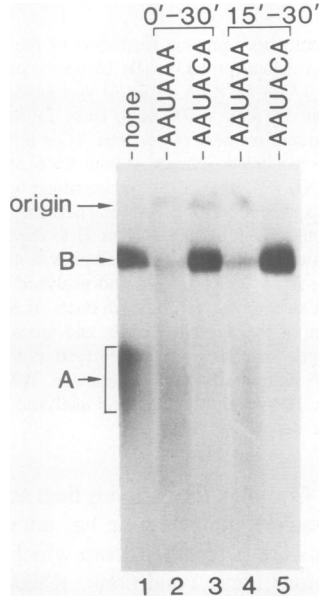

Fig. 6. Dissociation of pre-cleavage complexes. Uniformly labeled −141/+55 RNA (5 fmol) was incubated in nuclear extract without PVA. In **lane 1**, no competitor was added and the incubation was carried out for 30 min. To other reactions, 1000 fmol unlabeled −141/+55 RNA, containing either AAUAAA or AAUACA, was added. This labeled RNA was added either simultaneously with the labeled RNA, or 15 min later. After a total incubation of 30 min, all reactions were stopped by adding heparin and analyzed for complex formation. **Lane 2**: AAUAAA-containing 141/+55 RNA competitor, added at 0 min; **lane 3**: AAUACA-containing −141/+55 RNa competitor, added at 0 min; **lane 4**: AAUAAA-containing −141/+55 RNA competitor, added after 15 min; **lane 5**: AAUACA-containing −141/+55 RNA competitor, added after 15 min.

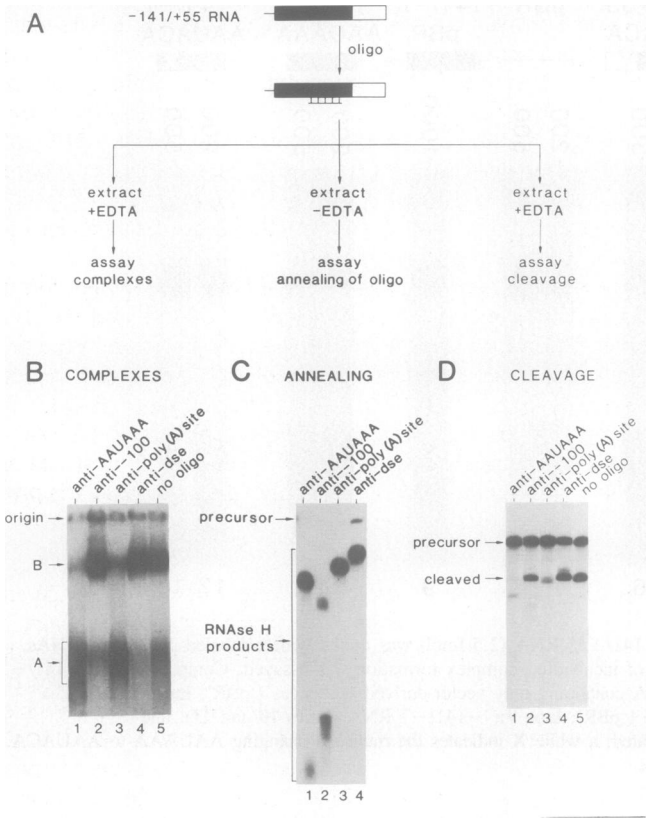

Fig. 7. DNA oligonucleotides prevent formation of pre-cleavage complex. **(A)** Diagram of experimental protocol. **(B)** Effect on complex formation. Uniformly labeled −141/+55 RNA (5 fmol) was annealed to anti-AAUAAA oligo (**lane 1**); anti- −100 oligo (**lane 2**); anti-poly(A) site oligo (**lane 3**); anti-downstream element oligo (**lane 4**) or no oligo (**lane 5**), incubated for 15 min in nuclear extract without PVA and analyzed for complex formation. **(C)** RNase H assay to determine whether oligonucleotides had annealed. Reactions were performed as in (B), except that EDTA was omitted. This permits RNase H to digest the RNA at regions which are base-paired to the DNA oligonucleotide. After 15 min incubation in extract, RNA was prepared and analyzed by electrophoresis on a 6% polyacrylamide denaturing gel. The products of RNase H digestion vary with the position of the oligonucleotide, and are indicated by brackets. **(D)** Effect on cleavage. Reactions were performed as in (A) except that PVA was included to increase the rate of cleavage. After 30 min of incubation in extract, RNA was prepared and analyzed by electrophoresis on a 6% polyacrylamide denaturing gel.

Wickens, 1987). Complex formation is then assayed (Figure 7B). To ensure that each oligonucleotide has annealed to the RNA, identical reactions are performed from which EDTA is omitted (Figure 7C). Under these conditions, RNase H is active and degrades the RNA complementary to any annealed oligonucleotide. Thus, by gel electrophoresis of the deproteinized RNA, one can assess whether the oligonucleotide has annealed. Finally, to determine whether the annealed oligonucleotide affects 3′ end cleavage, each RNA/oligonucleotide heteroduplex is incubated in nuclear extract containing EDTA, and the extent of 3′ end cleavage is measured by gel electrophoresis (Figure 7D). To increase the rate of cleavage, PVA is included in these reactions.

In the absence of any oligonucleotide, complex B forms as expected (Figure 7B, lane 5). In the presence of an oligonucleotide complementary to AAUAAA ('anti-AAUAAA'), it does not form efficiently; instead the RNA is found primarily in complex A (lane 1). This oligonucleotide also blocks cleavage; whereas 50% of the RNAs incubated without any oligonucleotide are cleaved (Figure 7D, lane 5), fewer than 5% of the RNAs annealed to the anti-AAUAAA oligonucleotide are cleaved (Figure 7D, lane 1). These inhibitory effects of the anti-AAUAAA oligonucleotide are specific: a control oligonucleotide complementary to sequences from 85 to 100 bases upstream of the cleavage site ('anti- −100') has little or no effect on either complex formation (Figure 7B, lane 2) or cleavage (Figure 7D, lane 2), even though it has annealed to the RNA (Figure 7C, lane 2).

An anti-poly(A) site oligonucleotide (which extends from −9 to +7) also prevents both complex formation and cleavage (Figure 7B, lane 3, and 7D, lane 3). We cannot distinguish whether the inhibitory effect of this oligonucleotide is due to base-pairing to the poly(A) site or to the proximity of the end of the RNA−DNA duplex to AAUAAA.

The function of the downstream element is not prevented by annealing it to an oligonucleotide. An anti-downstream element oligonucleotide (which extends from +5 to +23) neither blocks formation of the complex (Figure 7B, lane 4) nor prevents cleavage (Figure 7D, lane 4). In the absence of EDTA, more than 90% of the RNA is truncated by RNase H (Figure 7C, lane 4), suggesting that, at least at the start of the incubation, the DNA oligonucleotide is base-paired to the RNA.

The sequence between +27 and +55 exhibits some downstream element function when sequences between +1 and +27 are missing (D.Zarkower and M.Wickens, submitted). Thus, in principle, +27 to +55 sequences might direct formation of the pre-cleavage complex on RNAs base-paired to an oligonucleotide from +5 to +23. This is not the case, however. In oligo-interference experiments such as those in Figure 7, −141/+23 RNA annealed to the anti-downstream element oligonucleotide (which extends from +5 to +23) forms the pre-cleavage complex and is cleaved efficiently (not shown).

In summary, DNA oligonucleotides pre-annealed to AAUAAA or to the poly(A) site prevent cleavage and formation of the pre-cleaved complex, confirming that the AAUAAA sequence is required for complex formation. In contrast, an oligonucleotide annealed to the downstream element(s) has little or no effect. Either the downstream region can function even when base-paired to DNA, or the oligonucleotide annealed downstream is actively removed from the RNA during formation of the pre-cleavage complex.

Discussion

Four lines of evidence support the conclusion that complex B is a specific pre-cleavage complex formed between the pre-mRNA and factors required for cleavage. (i) It is detected prior to cleavage (Figure 1) and contains only the precursor, not the cleaved product (Figure 2). (ii) Mutant RNAs which are not cleaved fail to produce this complex (Figure 4 and D.Zarkower and M.Wickens, submitted). In contrast, formation of complex A is sequence-independent. (iii) Formation of complex B is saturable by substrate, so long as the substrate contains AAUAAA and a functional downstream element (Figure 5). (iv) DNA oligonucleotides which block cleavage when annealed to the pre-mRNA also prevent formation of complex B (Figure 7).

Formation of pre-cleavage complexes on adenovirus L3 (Skolnik-David *et al.*, 1987) and on herpes virus TK (Zhang and Cole, 1987) pre-mRNAs probably requires some of the same components as on SV40 late pre-mRNA. In particular, since formation of each complex requires AAUAAA, it is likely that at least the component which discriminates this sequence is com-

mon. However, since the downstream elements differ (D.Zarkower and M.Wickens, submitted), the complexes may not be identical.

Previously we identified a complex which contains cleaved RNA and which is stable in solution (Zarkower and Wickens, 1987). It was detected using an oligonucleotide/RNase H protection assay; the AAUAAA sequence in cleaved RNA, but not precursor, is inaccessible to a complementary oligonucleotide and RNase H. The complex was detected only when the cleaved RNA contained a 3′ terminal deoxyadenosine; RNA with a 3′ hydroxyl terminus formed no detectable complex.

The gel retardation data presented here and the previous oligonucleotide/RNase H protection results differ in two respects. The protection assay detects no complexes on the precursor; the retardation assay does. Similarly, the protection assay fails to detect complexes on cleaved RNAs which lack a 3′ deoxyadenosine; these RNAs do form complexes in the retardation assay (Figure 3). These two differences probably reflect the different sensitivity of each assay to complex instability. The pre-cleavage complex may be unstable in solution since it can exchange between precursors (Figure 6). Transient dissociation would result in a lack of protection at AAUAAA, but would not prevent detection in the gel retardation assay. Similarly, although the complex formed on cleaved RNA protects AAUAAA only in RNAs which contain a 3′ deoxyadenosine, it presumably forms transiently on RNAs with a 3′ hydroxyl, since such RNAs are polyadenylated in an AAUAAA-dependent reaction. These transient complexes are detected in the retardation assay.

The exceptional stability of the post-cleavage complex formed on 3′ deoxyadenosine-terminated RNAs is not due simply to the inability of the RNA to receive poly(A); RNAs with non-extendable 3′ termini, but without a 3′ deoxyadenosine, form only unstable complexes (V.Bardwell and M.Wickens, unpublished data). Thus we infer that the stability which results in protection of AAUAAA may be due to a close and specific interaction between the presumed poly(A) polymerase and the end of a substrate containing 3′ deoxyadenosine. Binding to AAUAAA may persist during that interaction, or be required only to establish it.

Formation of pre- and post-cleavage complexes requires the AAUAAA sequence (Zarkower and Wickens, 1987 and Figure 5). Cleaved RNA, which is a substrate for polyadenylation but not for cleavage, nevertheless competes both for formation of the pre-cleavage complex and for cleavage. We infer that a common component is required both for cleavage and for polyadenylation *in vitro*. Since both reactions are AAUAAA-dependent, as is competition by cleaved RNA, this component may discriminate AAUAAA (Manley, 1983; Manley *et al.*, 1985; Zarkower *et al.*, 1986). In a coupled cleavage/polyadenylation reaction, the factor which recognizes AAUAAA may dissociate either before or after cleavage has occurred.

After cleavage, the 3′ half-molecule apparently is released, while the 5′ half-molecule remains in a specific, post-cleavage complex (Figure 2). 'Release' is inferred from the accelerated mobility of the 3′ half-molecule in the gel retardation assay: it comigrates with 'non-specific' complexes formed on RNAs of the same length. By the same criterion, the 5′ half-molecule is released after it receives poly(A) (Figure 2 and Skolnik-David *et al.*, 1987). Similarly, its AAUAAA sequence becomes accessible to oligonucleotides and RNase H (Zarkower and Wickens, 1987).

RNAs which end at the poly(A) site (pre-cleaved RNAs) behave very differently than RNAs with non-SV40 sequences immediately downstream of that site. Whereas both precursor RNA and cleaved RNA form specific complexes, RNAs which carry a prokaryotic sequence of 32 nucleotides downstream do not (D.Zarkower and M.Wickens, submitted). Similarly, precursor RNA and cleaved RNAs undergo processing in the extract (cleavage or polyadenylation, respectively), while RNA with irrelevant downstream sequences does not. Thus, as inferred previously (Zarkower and Wickens, 1987), formation of specific complexes requires AAUAAA and either a downstream element, or a 3′ terminus near AAUAAA.

The downstream region is essential both for complex formation and for cleavage. Humphrey *et al.* (1987) recently have shown that it is protected from RNase T_1 digestion in the pre-cleavage complex. However, the isolated downstream region, freed from AAUAAA, is released after cleavage (Figure 2); similarly, it does not form a specific complex after addition to extract (not shown). Together these results suggest that binding to the downstream region occurs only when AAUAAA is present on the same molecule.

Working model

Our data lead to the following working model of cleavage and polyadenylation. The factor that recognizes AAUAAA is physically associated with a component required for polyadenylation (Zarkower and Wickens, 1987). This factor binds to the AAUAAA sequence, forming a 'pre-cleavage complex' in which cleavage then occurs. Assembly of this complex requires the region downstream of the poly(A) site. After cleavage, a 'post-cleavage complex' persists in which the 5′ half-molecule is retained while the 3′ half-molecule is released. This complex may contain the components which recognize AAUAAA and are required for polyadenylation, though this is unproven. The post-cleavage complex dissociates during polyadenylation. Dissociation liberates the processing activities to cleave and polyadenylate a second mRNA precursor (Zarkower and Wickens, 1987).

Materials and methods

Plasmid construction and synthesis of RNA substrates

−141/+55 RNA and −141/+55:AAUACA RNA were prepared by transcription of *Dra*I-cleaved pSPSV −141/+70 and pSPSV −141/+70:AAUACA (Zarkower and Wickens, 1987) by SP6 RNA polymerase (Melton *et al.*, 1984; Konarska *et al.*, 1984). These RNAs contain 29 bases of vector-derived sequence followed by 196 bases of SV40 sequence.

−141/−7 RNA and −141/−7:AAUACA RNA were synthesized by transcription of the same plasmids after cleavage at −7 with *Hpa*I. −141/−1 RNA contains SV40 sequences from −141 to −1, followed by a 3′-terminal 32 nt of vector-derived sequence, from the *Xba*I site to the *Eco*RI site of pSP64 (Melton *et al.*, 1984).

−59/+55 and −59/+1 RNAs were prepared as follows: to generate pSPSV −59/+70, a *Sfa*NI/*Hin*dIII fragment of pSPSV −141/+70, containing SV40 sequences from −59 to +70, was inserted into *Sma*I/*Hin*dIII pSP65 (Melton *et al.*, 1984). This plasmid was cleaved with *Dra*I and transcribed, yielding −59/+55 RNA. This RNA contains nucleotides −59 to +55 of SV40 and 39 vector-derived nucleotides at the 5′ terminus. pSPSV −59/+1 contains SV40 sequences from −59 to the poly(A) site, fused to a *Hae*II linker. Transcription of *Hae*II-cleaved pSPSV −59/+1 yields −59/+1 RNA, which is identical to −59/+55 RNA but extends only to +1.

RNAs used in competition experiments (Figure 5) were prepared as follows: the competitor RNA containing only prokaryotic sequences ('pBR' in Figure 5) was transcribed from *Pvu*II-cleaved pSP64. Other competitor RNAs were transcribed from the templates described above.

Transcription by SP6 RNA polymerase was carried out in the presence of [^{32}P]UTP and GpppG (Melton *et al.*, 1984; Konarska *et al.*, 1984). RNA was purified by elution from a polyacrylamide gel slice in 0.5 M ammonium acetate, 0.1% sodium dodecyl sulfate, and 0.5 mM EDTA at 25°C for 2−12 h. The eluted RNA was precipitated with ethanol.

Preparation of nuclear extract and in vitro processing

Nuclear extract was prepared from HeLa cells by the method of Dignam *et al.*

(1983) except that $MgCl_2$ was omitted from buffer A, 0.125 mM EDTA was included in buffer D, extract was dialyzed 3.5 h, and phenylmethylsulfonyl fluoride (PMSF) was omitted from all buffers. Reactions contained 5.5 μl nuclear extract, 40 mM KCl, 1 mM 3′ dATP, and 1–10 fmol labeled pre-mRNA in a total volume of 12.5 μl. Certain reactions (noted in the text) also included 2.8% polyvinyl alcohol (PVA), 20 mM phosphocreatine (Sigma), and 0.1 mM ATP. These conditions are optimal for cleavage and polyadenylation. RNA was prepared after incubation as described (Zarkower and Wickens, 1987).

Assay for complex formation by gel electrophoresis

Nuclear extract reactions were stopped by the addition of heparin to a final concentration of 5 mg/ml. After 10 min on ice, 5 μl of each reaction was analyzed by electrophoresis through a non-denaturing polyacrylamide gel (4% acrylamide, 0.05% *N,N′* methylenebisacrylamide) at 15 V/cm for 2–6 h (Konarska *et al.*, 1986; Zhang and Cole, 1987; Skolnik-David *et al.*, 1987). Gels were 1.5 mm thick and 12–16 cm long. Both the gel and the buffer contained 45 mM Tris, 45 mM boric acid, and 0.6 mM EDTA, pH 8.3. Gels were dried and exposed to X-ray film.

Two-dimensional analysis of complexes

In the first dimension, complexes were separated by electrophoresis through a non-denaturing gel, as described above. After autoradiography to detect complexes, strips of this gel were excised and laid onto a 1.5 mm thick 6% polyacrylamide gel containing 7 M urea, and electrophoresed in the second dimension at 45 mA for 2–4 h. The gels were dried and exposed to X-ray film.

Competition experiments

RNA competitors were transcribed in the presence of [^{3}H]UTP and eluted from a polyacrylamide gel after UV shadowing. The mass of competitor RNA was determined as described (Zarkower and Wickens, 1987). Except as noted in the text, labeled RNA and unlabeled competitors were mixed and then added together to nuclear extract.

Oligonucleotide interference experiments

DNA oligonucleotides were annealed by 10 min incubation at 30°C of reactions to which all components except nuclear extract had been added. Nuclear extract then was added, and the reactions incubated an additional 15–30 min. To assay complex formation, reactions contained 5.5 μl nuclear extract, 40 mM KCl, 1 mM EDTA, 35 ng oligonucleotide and 5 fmol RNA, in a total volume of 12.5 μl. After annealing and 15 min incubation, 5 μl of each reaction was electrophoresed on a non-denaturing gel. To assay annealing, reactions were performed as above except that EDTA was omitted and 3′ dATP (1 mM) and $MgCl_2$ were added. Incubation was for 15 min. RNA then was prepared and analyzed by electrophoresis on a denaturing 6% polyacrylamide gel. To assay the effects of oligonucleotides on cleavage, reactions were performed under conditions optimal for cleavage (see above) except that EDTA (1 mM) replaced 3′ dATP. After 30 min incubation, RNA was analyzed as before. The sequences of the oligonucleotides are: anti-AAUAAA, 3′GACGTTATTTGTTCAA5′; anti- −100, 3′TTGATCTTACGTCACT5′; anti-poly(A) site, 3′CAATTGTTGTTGTTAA5′; and anti-downstream element, 3′TAACGTAAGTAAAAGACAAAGTC5′.

Acknowledgements

We are grateful to Walter Keller, Gerhard Christofori, Tim Humphrey, Vlatka Lucijanic, Phillip Sharp, Claire Moore, Hagit Skolnik-David, Susan Berget, Charles Cole and James Stephano for communication of results prior to publication. We thank the members of the Wickens laboratory for helpful discussion and Ray Smith for technical help. DNA oligonucleotide synthesis was performed by the Protein Sequence–DNA Synthesis Facility, supported by the PHS, NIH (S10-RR01684 and CA-07175), a GRS Grant to the University of Wisconsin Medical School, and by the University of Wisconsin Graduate School. D.Z. is supported by a predoctoral fellowship from the NSF and a predoctoral training grant from the NIH. This work was supported by the NIH (GM31892) and a Steenbock Career Development Award to M.W.

References

Birnstiel,M.L., Busslinger,M. and Strub,K. (1985) *Cell*, **41**, 349–359.
Conway,L. and Wickens,M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3949–3953.
Conway,L. and Wickens,M. (1987) *EMBO J.*, **6**, 4177–4184.
Dahlberg,A.E., Dingman,C.W. and Peacock,A.C. (1969) *J. Mol. Biol.*, **41**, 139–147.
Dingnam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1488.
Humphrey,T., Christofori,G., Lucijanic,V. and Keller,W. (1987) *EMBO J.*, **6**,
Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1984) *Cell*, **38**, 731–736.
Konarska,M.M. and Sharp,P.A. (1986) *Cell*, **46**, 845–855.
Manley,J.L. (1983) *Cell*, **33**, 595–605.
Manley,J.L., Yu,H. and Ryner,L. (1985) *Mol. Cell. Biol.*, **5**, 373–379.
Melton,D., Krieg,P.A., Rebagliati,M.A., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Moore,C. and Sharp,P.A. (1985) *Cell*, **41**, 845–855.
Pikielny,C.W. and Rosbash,M. (1986) *Cell*, **45**, 869–877.
Skolnik-David,H., Moore,C. and Sharp,P.A. (1987) *Genes and Development*, **1**, 672–682.
Wickens,M. and Stephenson,P. (1983) *Science*, **226**, 1045–1056.
Zarkower,D. and Wickens,M. (1987) *EMBO J.*, **6**, 177–186.
Zarkower,D., Stephenson,P., Sheets,M. and Wickens,M. (1986) *Mol. Cell. Biol.*, **6**, 2317–2323.
Zhang,F. and Cole,C.C. (1987) *Mol. Cell. Biol.*, **7**, 3277–3286.

Received on October 6, 1987