Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal

Carol S. Lutz and James C. Alwine

Department of Microbiology, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6142 USA

An integral component of the splicing machinery, the U1 snRNP, is here implicated in the efficient polyadenylation of SV40 late mRNAs. This occurs as a result of an interaction between U1 snRNP-A protein and the upstream efficiency element (USE) of the polyadenylation signal. UV cross-linking and immunoprecipitation demonstrate that this interaction can occur while U1 snRNP-A protein is simultaneously bound to U1 RNA as part of the snRNP. The dual reactivity of U1 snRNP-A occurs because the protein has two RNA recognition motifs (RRMs). The target RNA of the first RRM (RRM1) has been shown previously to be the second stem-loop of U1 RNA. We have found that a target for the second RRM (RRM2) is within the AUUUGURA motifs of the USE of the SV40 late polyadenylation signal. RNA substrates containing the wild-type USE efficiently bind to U1 snRNP-A protein, whereas substrates fail to bind when motifs of the USE were replaced by linker sequences. The addition of an oligoribonucleotide containing a USE motif to an in vitro polyadenylation reaction inhibits polyadenylation of a substrate representing the SV40 late polyadenylation signal, whereas a mutant oligoribonucleotide, a nonspecific oligoribonucleotide, and an oligoribonucleotide containing the U1 RNA-binding site had much reduced or no inhibitory effects. In addition, antibodies to bacterially produced, purified U1 snRNP-A protein specifically inhibit in vitro polyadenylation of the SV40 late substrate. These data suggest that the U1 snRNP-A protein performs an important role in polyadenylation through interaction with the USE. Because this interaction can occur when U1 snRNP-A protein is part of the U1 snRNP, our data provide evidence to support a link between the processes of splicing and polyadenylation, as suggested by the exon definition model.

[Key Words: U1 snRNP; polyadenylation; cross-linking; RNA recognition motif (RRM); upstream efficiency element (USE); oligoribonucleotides]

Received August 27, 1993; revised version accepted January 14, 1994.

Polyadenylation is an essential processing event for most eukaryotic mRNAs. The signals and components of the polyadenylation reaction have been the subject of extensive study (for review, see Manley 1988; Wickens 1990; Wahle and Keller 1992). Examination of many polyadenylation signals has shown that besides the conserved AAUAAA sequence, there are surrounding RNA sequences required for efficient utilization of the AAUAAA as a polyadenylation signal. Downstream efficiency elements (DSEs) have been identified in many different viral systems (Gil and Proudfoot 1984, 1987; McDevitt et al. 1984, 1986; Sadowsky and Alwine 1984, Cole and Stacy 1985; Conway and Wickens 1985; Sadowsky et al. 1985; Zhang and Cole 1987; Wilusz et al. 1988; Wilusz and Shenk 1990) at distances from 5 to 60 nucleotides downstream of the poly(A) addition site. These elements do not conform to a consensus sequence but are characterized as either U-rich or GU-rich. Upstream efficiency elements (USEs) have also been described in several systems, including the SV40 late RNAs (Carswell and Alwine 1989; Schek et al. 1992), ground squirrel hepatitis virus (Russnak and Ganem 1990; Russnak 1991), cauliflower mosaic virus (Sanfacon et al. 1991), HIV-1 [Brown et al. 1991; DeZazzo et al. 1991; Valsamakis et al. 1991, 1992], and adenovirus type 2 (DeZazzo and Imperiale 1989). The USE of the SV40 late polyadenylation signal has three motifs with the consensus AUUUGURA, located between 15 and 50 nucleotides upstream of the AAUAAA (Fig. 1A,B), which function to positively affect polyadenylation [Schek et al. 1992]. Motifs of similar sequence and function are found in the ground squirrel hepatitis virus and cauliflower mosaic virus (Russnak 1991; Sanfacon et al. 1991). The conservation of these elements between plant and mammalian virus polyadenylation signals suggests a conserved function, possibly

*Corresponding author.
mediated through protein binding. This concept is supported by the observation that the AUUUGURA motif shares homology with two protein binding sites on U1 small nuclear RNA (snRNA), that for the U1 small nuclear ribonucleoprotein (U1 snRNP)-specific A protein and that for the Sm complex of proteins (Fig. 1C).

The similarity with binding sites for protein components of snRNP particles brings to light a long standing proposal that polyadenylation and splicing may be linked [Moore and Sharp 1984; Hashimoto and Steitz 1986; Raju and Jacob 1988; Pandey et al. 1990; Wasserman and Steitz 1993]. This possibility has been best presented in the model of exon definition [Robberson et al. 1990]. In this model, exons are defined by factors binding at the 3′ splice site of an upstream intron that communicate across the exon with factors at the 5′ splice site of the downstream intron resulting in definition of the exon. It has been proposed that the U1 snRNP, known to interact via base-pairing at the 5′ splice site, may also interact with factors at the 3′ splice site, including the U2 snRNP, thus poising the U1 snRNP to search across the exon for the 5′ splice site. The exon definition model is challenged when the definition of the final exon and removal of the final intron is considered. Although the final exon has the required 3′ splice site, there is no corresponding 5′ splice site on the other end of the exon. This has led to the suggestion that the polyadenylation signal may play the role of the 5′ splice site in this instance to allow appropriate removal of the final intron. Data from coupled in vitro splicing and polyadenylation reactions showing that the presence of a 3′ splice site can stimulate polyadenylation [Niwa et al. 1990] and that mutations in the AUUAAA of the polyadenylation signal can depress splicing of a final intron [Niwa and Berget 1991] support this notion. Suggestions that polyadenylation and splicing may be coregulated are also supported by in vivo experiments [Chiou et al. 1991; Nesic et al. 1993].

The U1 snRNP-A protein has two RNA recognition motifs (RRMs; Query et al. 1989; Kenan et al. 1991), also known as RNA-binding domains (RBDS; Bandziulis et al. 1989). The amino-terminal RRM provides the tether to U1 RNA through specific binding to the second stem-loop of U1 RNA [Lutz-Freyermuth and Keene 1989; Scherly et al. 1989; Lutz-Freyermuth et al. 1990; Tsai et al. 1991]. The target RNA for the carboxy-terminal RRM had not been identified previously. We show that one target RNA for the carboxy-terminal RRM of U1 snRNP-A protein is within the motifs of the USE of the SV40 late polyadenylation signal. In addition, we show that this interaction with the USE can occur while U1 snRNP-A protein is bound to U1 RNA. That the interaction of A protein with the USE is functionally significant for polyadenylation is indicated by the finding that in vitro polyadenylation of an SV40 late polyadenylation substrate is specifically inhibited by [1] an oligoribonucleotide representing the USE sequence, and [2] antibodies to bacterially produced, purified U1 snRNP-A protein. Overall, our data suggest a means by which components of the splicing apparatus can directly affect the polyadenylation reaction.

Results

UV cross-linking of HeLa proteins to the SV40 USE

The SV40 late polyadenylation signal is shown in Figure 1A. Two downstream efficiency elements, DSE19/40 (Conway and Wickens 1985) and DSE59/67 (Sadofsky et al. 1985), have been identified by deletion mutagenesis and analysis in different systems. The redundant DSEs may assure efficient polyadenylation under a variety of cellular conditions. However, the elements are GU- and U-rich, respectively; hence they are similar to other known DSEs.

The upstream efficiency elements of the SV40 late polyadenylation element were defined by linker substitution analysis by Schek et al. (1992). Figure 1B shows the sequence of the upstream region of the SV40 late polyadenylation signal with the three USE motifs that are essentially identical (consensus AUUUGURA). Mutation of all three motifs reduces polyadenylation by >70%. The motifs were also shown to function in a distance-related manner; for example, when a single motif is present, it has a greater positive effect on polyadenylation the closer it is to the AAUAAA. Multiple motifs function additively with respect to their distance from the AAUAAA [Schek et al. 1992]. Figure 1C shows the sequence of USE motif-3 in comparison with the binding sites on U1 snRNA for the U1 snRNP-specific A protein and the Sm complex of proteins. The similarity of these sequences led us to ask whether any U1 snRNP protein(s) could bind the USE motifs.

To determine specific protein binding to the USE motifs, we performed UV cross-linking experiments with HeLa cell nuclear extracts and 32P-labeled RNA substrates. UV light generates highly reactive species of RNA that will react with proteins in direct contact with the RNA. If the RNA is uniformly labeled, for example, with [32P]UTP, the protein(s) could bind the USE motifs.

Cross-linked complexes were immunoprecipitated by use of the monoclonal antibody Y12 [Lerner et al. 1981], which is specific for Sm proteins. This approach would precipitate both Sm proteins bound to the RNA, as well as any other proteins that may be bound to both the substrate RNA and to the snRNP. Following UV cross-linking and prior to immunoprecipitation, the cross-linked complexes were RNase treated to remove the bulk of the substrate RNA, but not enough to disrupt the integrity of the snRNPs. The immunoprecipitated proteins were visualized by SDS-PAGE and autoradiography (Fig. 2). The wild-type SV40 RNA substrate used was an in vitro produced transcript representing the SV40 sequences between the the BamHI and DraI sites (Fig. 1A; SV40 nucleotides 2533–2729). This transcript contains the three USE motifs as well as all other known polyadenylation signal elements. In addition, a control transcript (triple mutant) was used that was identical in sequence except the three USE motifs were mutated by linker substitution as indicated in Figure 1B.

By use of the wild-type substrate, a number of proteins

GENES & DEVELOPMENT 577
Figure 1. (A) Diagram of the SV40 sequences between the BamHI site (SV40 nucleotide 2533) and the DraI site (SV40 nucleotide 2729) containing all of the known elements of the late polyadenylation signal. These include the consensus AAUAAA, the cleavage and polyadenylation site (A^), the downstream efficiency elements, DSE19/40 and DSE59/67 (see text), and upstream efficiency elements (USEs; see text). This fragment has been inserted in front of the T7 promoter in pGEM for in vitro transcription of polyadenylation substrate RNA (template plasmids were linearized at the DraI site, SV40 nucleotide 2729). Flanking the 3'- and 5'-side of the AAUAAA the naturally occurring Alu and HpaI restriction sites have been replaced in our constructions with XbaI and XhoI linkers. These have been used for ease of replacement of USE and DSE sequences with mutant sequences (Schek et al. 1992). (B) The sequence of the upstream region with the three upstream efficiency element motifs denoted in boxes. Below each motif is indicated the specific linker substitution that mutates the three motifs in the triple mutant substrate used in these experiments. (C) The sequence of the USE motif 3 is compared to binding sites on U1 RNA for the U1 snRNP-A protein and for the Sm complex of proteins.

were labeled (lane 5). When compared with snRNP marker proteins [immunoprecipitated by anti-Sm antibody from ^35S-labeled HeLa cell extract (lane 7)] the most notable protein labeled by cross-linking was one of ~32-kD (lane 5), which comigrated with the U1 snRNP-A marker (lane 7). The 32-kD protein was not visualized when the triple mutant substrate RNA was used (lane 6), nor when a nonspecific RNA generated from pGEM was used as the substrate (lane 4). Other proteins are clearly visible in lane 5, some of these are present, though less intense, in lane 2 where normal human serum was used instead of anti-Sm. Under these conditions the band migrating with the 32-kD protein is not present, suggesting that at least some of the other bands may be nonspecific. However, because the wild-type substrate contains the entire polyadenylation element (Fig. 1A) it is possible that some of the other proteins are significant for the polyadenylation process and are in a complex that can be precipitated (see Discussion). Overall, the band migrating with U1 snRNP-A protein was the only readily identifiable snRNP protein band. We have not seen specifically labeled proteins that correspond to known components of the Sm complex of proteins by examination of cross-linked products on gels that fractionate low molecular weight proteins (data not shown). Because Sm proteins and the U1 snRNP-A protein are not antigenically cross-reactive, we concluded that the putative U1 snRNP-A protein was immunoprecipitated by the anti-body because it was simultaneously bound to U1 RNA in the snRNP and to the substrate RNA containing the USE motifs.

In vitro-transcribed/translated U1 snRNP-A protein specifically binds to USE-containing RNAs

The structure of the U1 snRNP-A protein is shown diagrammatically in Figure 3. The protein has two RRMs, which could suggest a mechanism whereby the U1 snRNP-A protein binds simultaneously to U1 RNA and to SV40 substrates. The amino-terminal RRM, RRM1, has been shown previously to bind U1 RNA (Scherly et al. 1989; Lutz-Freyermuth et al. 1990), hence providing the tether to U1 RNA. The target RNA for the carboxy-terminal RRM, RRM2, had not been identified previously. It was therefore important to (1) confirm that U1 snRNP-A protein itself bound to the USE motifs; and (2) to determine whether the carboxy-terminal RRM was capable of direct binding to USE motifs.

To perform this experiment, we prepared the intact U1 snRNP-A protein or the individual RRMs by use of a reticulocyte lysate coupled transcription/translation system. The templates for these proteins were engineered such that the g10 epitope tag would be placed on the amino terminus of each protein (Fig. 3; Lutz-Freyermuth et al. 1990). The resulting protein or protein fragments were tested with various ^32P-labeled RNA sub-
Direct binding of U1 snRNP-A to SV40 USE

Figure 2. UV cross-linking of proteins to RNA followed by immunoprecipitation. HeLa nuclear extracts were incubated with substrate RNAs (as indicated above), reacted with UV light, treated with RNase, and immunoprecipitated as described with antibodies as indicated. (Lanes 1–3, NEG) Negative sera; (lanes 4–7, Sm) Y12. (Lane 7) An Sm immunoprecipitation of an 35S-labeled HeLa cell extract.

Figure 3. Schematic representation of the U1 snRNP-A protein. Each RRM is indicated. The highly conserved RNP-2 and -1 elements within each RRM (Bandziulis et al. 1989) are indicated by vertical bars. The full-length A protein [top] and the two separate RRMs [bottom] were produced by in vitro transcription and translation with the g10 epitope tag [g10] on the amino-terminal end of each protein.

Inhibition of polyadenylation by specific oligonucleotides

If the binding of U1 snRNP-A protein to the USE is a significant event in polyadenylation, then the polyadenylation reaction should be specifically inhibited by oligoribonucleotides representing the USE motifs. In vitro polyadenylation reactions were performed using HeLa
Figure 4. Binding and coprecipitation of the U1 snRNP-A protein to RNAs. [Input] Input samples of the four substrate RNAs used in the binding experiments. In this and all other panels, lane 1 represents the wild-type SV40 substrate RNA containing the USEs; lane 2 represents the triple mutant; lane 3 represents full-length U1 RNA; and lane 4 represents a 98-nucleotide control transcript from the pGEM plasmid (see text). Each of these RNAs was tested for binding with in vitro-synthesized full-length U1 snRNP-A (lanes 1–4, Full-length A), RRM1 of U1 snRNP-A (lanes 1–4, RRM1), and RRM2 of U1 snRNP-A (lanes 1–4, RRM2). Each protein carried the g10 epitope tag as indicated in Fig. 3. An unprogrammed translation reaction was used as a negative control in the experiments depicted in the unprogrammed panel. The protein and RNAs were mixed and incubated at 30°C for 15 min. Following the binding reaction, epitope tag antibody was added and incubated for 15 min. Pansorbin (Calbiochem) was then added to precipitate antibody-bound complexes. Coprecipitated RNAs were then phenol extracted, ethanol precipitated, and subjected to electrophoresis on 10% polyacrylamide/8 M urea gels. [Unbound] The RNA remaining in the supernatant after binding to full-length U1 snRNP-A protein and immunoprecipitation. These data show that the input RNA was not substantially degraded during the procedure.

cell nuclear extracts and wild-type SV40 RNA substrates in the presence of oligoribonucleotides representing the USE motif or nonspecific sequences (described in Materials and methods). These oligoribonucleotides were tested for their ability to bind to U1 snRNP-A protein and to the separated RRMs. The USE oligoribonucleotide bound with the same specificity as the full-length wild-type substrate, whereas the nonspecific oligoribonucleotide failed to bind (data not shown). Figure 5A (lane 1) shows a polyadenylation reaction in the absence of all oligoribonucleotides. Lanes 2 and 3 show the outcome of a reaction in the presence of 1 or 10 μg (~150 or 1500 pmoles), respectively, of the USE motif oligoribonucleotide. Polyadenylation was effectively inhibited in the presence of excess amounts of this RNA motif, suggesting that a factor important for efficient polyadenylation is titrated away. In the presence of analogous amounts of a nonspecific oligoribonucleotide, polyadenylation proceeded as normal (lanes 4,5).

We have also found that the USE oligoribonucleotide inhibits not only the polyadenylation reaction but also in vitro cleavage as measured in the presence of cordycepin (not shown).

The data from Figure 4 suggested that RRM1, the U1 RBD of U1 snRNP-A protein, could also interact with the USE motifs. However, because RRM1 mediates a very stable binding to U1 RNA (Lutz-Freyermuth et al. 1990) it is unlikely that this site would be available to function in the polyadenylation reaction. In addition, our data (Fig. 2) suggest that U1 snRNP-A protein binds the USE motif by use of RRM2 whereas RRM1 may keep the protein tethered to the U1 snRNP. Hence, we would predict that an oligoribonucleotide representing the RRM1-binding site on U1 RNA (which is essentially a mutant binding site for RRM2) would not affect the polyadenylation reaction. Such an oligoribonucleotide was prepared (see Materials and methods) and found to bind specifically to RRM1 and not to RRM2 (data not shown) similar to the results with full-length U1 RNA (Fig. 4). Figure 5B shows that 1 or 10 μg (~150 and 1500 pmoles) of the U1 oligoribonucleotide (lanes 4 and 5) had little or no effect on polyadenylation, whereas the USE motif oligoribonucleotide (lanes 2 and 3) again showed significant inhibition.

In Figure 5C we show that a mutant oligoribonucleotide that has a mutation in the USE (mUSE) causes significantly less inhibition of in vitro polyadenylation compared with the USE oligoribonucleotide. The wild-type oligoribonucleotide sequence was GCUUUAUUUGUAAACCAGUAUA, whereas the mUSE sequence was GCUUUCUCUCUAAACCAGUAUA. The mUSE oligoribonucleotide showed reduced binding to U1 snRNP-A protein (not shown). Lane 1 shows the normal in vitro polyadenylation reaction. Lane 2 shows the expected inhibitory effect of the wild-type USE oligoribonucleotide, here using 0.5 μg (~75 pmoles), which is a lower amount of oligoribonucleotide than that used in Figure 5, A and B. Lane 3 shows that an equivalent amount of the mUSE
had a much reduced inhibitory effect. Lane 4 again shows that the nonspecific oligoribonucleotide had little to no inhibitory effect. Hence, a mutation in the USE that reduced U1 snRNP-A protein binding also affected the ability of the oligoribonucleotide to inhibit in vitro polyadenylation of the SV40 late polyadenylation substrate.

**Inhibition of polyadenylation by anti-U1 snRNP A protein antibodies**

To further authenticate the role of the U1 snRNP-A protein in polyadenylation, rabbit polyclonal antibodies were raised to bacterially expressed U1 snRNP-A protein. We have characterized this antibody and found that it can efficiently immunoprecipitate U1 snRNPs, purified, bacterially produced U1 snRNP-A protein and in vitro-translated U1 snRNP-A protein; in addition, it specifically detects U1 snRNP-A protein on Western blots. Furthermore, we have determined that the cross-linking of USE-containing RNA to the RRMs eliminates the ability of the antibody to precipitate the protein. Hence, we believe the major, or only, epitopes are in the RRMs.

The U1 snRNP-A protein-specific antibodies and a control antibody (prebleed) were purified by HPLC [see Materials and methods]. Figure 6 shows the effects of the purified antibodies on in vitro polyadenylation reactions [Fig. 6]. Polyadenylation was inhibited by antibodies to the U1 snRNP-A protein [lane 3], whereas an identical amount of the control antibody [prebleed] had minimal inhibition.

**Figure 6.** Polyadenylation is inhibited by antibodies to the U1 snRNP-A protein. (Lane 1) In vitro polyadenylation reaction in the absence of all antibodies but including 5 μl of dH2O; (lane 2) reaction plus 5 μl of purified rabbit prebleed; (lane 3) reaction plus 5 μl of purified rabbit anti-A antiserum.
effect on the reaction (lane 2). Previous work has indicated that antibodies to U1 snRNPs inhibited polyadenylation, but the actual protein specificity of these antibodies was unclear (Moore and Sharp 1984).

Discussion

U1 snRNP and polyadenylation

The critical role of snRNPs in splicing has been well established (for review, see Lührmann et al. 1990; Moore et al. 1993); however, a role for snRNPs, or snRNP proteins, in polyadenylation has not been as well defined. Despite evidence obtained with purified components that no snRNP is absolutely required for cleavage and polyadenylation (Wahle and Keller 1992), several lines of evidence suggest a role, or roles, for the U1 snRNP in polyadenylation. First, studies have shown that anti-Sm and anti-U1 snRNP antibodies inhibit polyadenylation (Wahle and Keller 1992); secondly, recent experiments with the SV40 late polyadenylation signal suggest that a direct interaction between U1 RNA and a binding site in the precursor RNA affects polyadenylation efficiency (Wasserman and Steitz 1993). Finally, immunoprecipitations have suggested an association between the U1 snRNP and poly(A) polymerase (Raju and Jacob 1988).

In previous in vitro polyadenylation studies highly purified basal polyadenylation factors were shown to accurately polyadenylate specific RNA substrates. In these studies it was shown that snRNPs were not components of the purified polyadenylation factors (Takagaki et al. 1989; Bienroth et al. 1991; Murthy and Manley 1992). However, these data do not rule out the possibility that snRNPs affect the overall function of the polyadenylation apparatus under normal condition in the nucleus or in crude HeLa cell extracts. Our data provide the first evidence for the involvement of a snRNP protein in polyadenylation. We show that the U1 snRNP-A protein, when tethered to the U1 snRNP, binds to the motifs of the upstream polyadenylation efficiency elements of the SV40 late RNAs. This binding is accomplished through utilization of a second RRM on the U1 snRNP-A protein (Fig. 3). Hence, the U1 snRNP-A protein is able to interact with two RNAs simultaneously: [1] the U1 RNA by utilizing the U1-specific RRM1 (Lutz-Freyermuth et al. 1990), and [2] the upstream element of the SV40 late polyadenylation signal by use of RRM2. The significance of this interaction in polyadenylation is indicated by our finding that the specific disruption of the interaction between RRM2 and the motifs of the SV40 upstream element inhibits SV40 late polyadenylation in vitro (Fig. 5A,C). Our UV cross-linking data (Fig. 2, lane 5) suggested that, along with U1 snRNP-A protein, a number of other proteins, not identifiable as snRNP proteins, cross-linked to the wild-type SV40 RNA substrate, and were immunoprecipitated with the anti-Sm antibody. Linker substitution of the USEs (triple mutant) eliminated not only the cross-linking of the U1 snRNP-A protein but also most of the other proteins. Not all of these proteins can be eliminated as nonspecific in the negative control by use of nonspecific antibody (lane 2). Hence, it is possible that other proteins bind to the USEs. However, the wild-type and triple mutant substrates that we used contained the entire polyadenylation signal (Fig. 1) and could bind other polyadenylation factors at the AAUAAA or downstream elements. The fact that mutation of the USEs eliminated most binding may suggest that a larger interactive complex is formed, and stabilized, when the entire polyadenylation signal (USEs–AAUAAA–DSEs) is intact. If this is the case, then it is expected that significant protein–protein interactions are involved.

Positive and negative controls by U1 snRNP-A protein

Purified U1 snRNP-A protein, untethered to U1 RNA, has been shown to inhibit polyadenylation of its own pre-mRNA in vitro (Boelens et al. 1993). This is apparently accomplished by binding of U1 snRNP-A protein to sites on its own pre-mRNA that lie upstream of the AAUAAA. These sites are identical to the U1 snRNP-A protein binding site on U1 RNA (Fig. 1), and are bound by RRM1 and not by RRM2. The availability of significant amounts of free RRM1 would be expected to occur only when U1 snRNP-A protein is in excess of U1 RNA. It is known that the binding of RRM1 to U1 RNA is quite stable (Lutz-Freyermuth et al. 1990), and that there is probably little free U1 snRNP-A protein in cells (Boelens et al. 1993). Hence, an excess of U1 snRNP-A protein may be a situation the cell avoids. In agreement, it has been suggested that U1 snRNP-A protein specifically autoregulates its own synthesis post-transcriptionally via the mechanism of the RRM1 (of free U1 snRNP-A protein) binding to specific recognition sequences upstream of the AAUAAA in its own RNA (Boelens et al. 1993). The free U1 snRNP-A protein bound to the RNA appears to inhibit polyadenylation through interaction with the poly(A) polymerase (Gunderson et al. 1994).

On the other hand, our data suggest a positive role for U1 snRNP-A protein in affecting polyadenylation, in this case through interaction with an upstream polyadenylation efficiency element. This effect would occur under conditions where RRM1 is occupied on U1 RNA and tethers U1 snRNP-A protein to the U1 snRNP. Hence, positive and negative actions of U1 snRNP-A protein may be regulated by the availability of RRM1 controlled by whether or not the protein is tethered to U1 snRNP.

U1 snRNP-A protein and last exon definition

The studies of Berget and co-workers have suggested an exon definition model for splicing regulation. In this model it is proposed that a complex of U2 and U1 snRNPs interacts with the 3′ splice site via binding of the U2 snRNP (Robberson et al. 1990; Grabowski et al. 1991; Kuo et al. 1991; Reich et al. 1992). The exon is defined when the snRNP complex extends across the exon to the 5′-splice site where the U1 snRNP binds (Robberson et al. 1990). Last exon definition would have
to be accomplished by a different mechanism because there is no 5′-splice site to be used for definition. In this regard, data have shown a link between splicing and polyadenylation in vitro [Niwa and Berget 1991] and in vivo [Chiu et al. 1991; Nesic et al. 1993]. The presence of a 3′-splice site increases the efficiency of utilization of the SV40 late polyadenylation signal [Niwa et al. 1990; Chiu et al. 1991]. Such findings support the suggestion that the last exon is defined by the 3′-splice site and some feature of the polyadenylation signal [Niwa et al. 1990; Niwa and Berget 1991]. Our present data also support this concept and suggest a possible mechanism. As shown in Figure 7, the U2/U1 snRNP complex may interact with the 3′ splice site as is proposed for internal exons. However, the subsequent definition of the final exon may be accomplished through the interaction of the U1 snRNP-A protein (tethered to U1 snRNP) with the upstream element of the polyadenylation signal. Our data suggest that this interaction greatly stimulates the polyadenylation process.

As mentioned above, Wassarman and Steitz [1993] have shown recently that the Ul snRNP can cross-link, via Ul RNA, to a site in last exon of SV40 late RNA. The site of this RNA–RNA interaction is 52 nucleotides upstream of the U1 snRNP-A protein binding motifs in the upstream element [see Fig. 7]. An effect of this interaction on polyadenylation can be explained by the second model shown in Figure 7 where the U1 snRNP is bound in close proximity to the USE, thus allowing the interaction between the USE and the U1 snRNP-A protein. The question that arises is whether the two mechanisms depicted in Figure 7 are both utilized. We believe that they are and that the latter mechanism, which uses direct U1 snRNP binding, may be a special condition. In the case of the SV40 19S and 16S classes of late mRNA the last exons are quite large, 2116 and 1211 nucleotides, respectively. Indeed, some of the 19S class of the late mRNA is polyadenylated but unspliced. Hence, the insertion of the U1 interaction site near the USE may be a means to circumvent normal last exon definition, which may be inefficient or impossible because of the large size of the last exon. Such a mechanism may be even more important when the precursor RNAs are being synthesized in high amounts and must be processed very efficiently, as is the case for the SV40 late RNA at late times in the lytic infection. This is not meant to infer that the utilization of U1 snRNP-A protein in polyadenylation is unique to papovaviruses. The small DNA viruses have long served as accurate models for eukaryotic gene expression mechanisms, especially RNA processing. They consistently utilize, or manipulate, existing cellular mechanisms for the expression of their genes. Hence, we suspect that the specific utilization of U1 snRNP-A protein in SV40 late RNA polyadenylation which we have described is a viral application of an analogous process used in the synthesis of many, if not all, spliced and polyadenylated cellular RNAs.

Materials and methods
Antisera
Anti-g10 epitope tag (anti-T7 Tag) antibodies were purchased from Novagen (Madison, WI). Anti-Sm monoclonal antibodies (Y12) were a gift of Dr. Joan Steitz (Yale University, New Haven, CT). Rabbit polyclonal antibodies were made to purified Ul snRNP-A by Cocalico Biologicals (Reamstown, PA), and were tested for specificity by Western blotting and immunoprecipitation [data not shown].

In vitro transcription of RNA substrates
RNA transcripts for UV cross-linking, solution binding, and in vitro polyadenylation reactions were synthesized by use of either SP6 or T7 RNA polymerase according to the supplier (Promega) in the presence of 50 μCi of [α^32P]UTP (Amersham). RNAs were gel-purified by overnight crush elution in high salt buffer [0.4 M NaCl, 10 mM Tris at pH 7.5, 0.1% SDS] prior to use in reactions. The human U1 RNA template was described previously [Lutz-Freyermuth and Keene 1989]; the SV40 substrates used were described by Schek et al. [1992].

In vitro transcription and translation
U1 snRNP-A protein constructs are described by Lutz-Freyermuth et al. [1990]. Coupled in vitro transcription/translation reactions (TNT) were performed in rabbit reticulocyte lysates as described by the supplier (Promega) in the presence of [35S]methionine for 1 hr at 30°C. Translation products were then used in RNA-binding assays.

UV cross-linking reactions and protein immunoprecipitations
UV cross-linking was performed as described previously [Pino et al. 1989]. Briefly, reactions contained 60% [vol/vol] HeLa nuclear extract [see below], 1 mM ATP [Pharmacia], 20 mM phosphocreatine [Sigma], 2.6% polyvinyl alcohol, and ~5×10^5 cpm of [35S]labeled RNA in a reaction volume of 25 μl. Reactions were incubated at 30°C for 10 min, 5 μg of Escherichia coli tRNA was added, and treatment with UV light ensued for 15 min at 4°C. A 15 W germicidal lamp [Sylvania G15T8] was used.

Figure 7. A model. The U1 snRNP-A protein can interact with the SV40 USE while tethered to the U1 snRNP, which is also interactive with the U2 snRNP at the 3′ splice site. This protein–RNA interaction with the USE may also occur if U1 RNA/ SV40 RNA base-pairing is taking place at the U1 site [Wassarman and Steitz 1993]. [p(A)] Site of cleavage and poly(A) addition; [DSE] downstream element.
HeLa cell nuclear extracts were prepared as described previously [Lutz et al. 1990]. Immunoprecipitations were performed as described previously (Moore 1990). In vitro polyadenylation reactions were performed as described above for UV cross-linking, except $1 \times 10^5$ cpm of substrate RNA was used, and the reactions were allowed to incubate at 37°C for 1 hr [Scheck et al. 1992]. Reaction products were extracted as above and analyzed on 40-cm 5% (29:1) polyacrylamide gels containing 7 M urea.

Solution binding of proteins to RNAs

In vitro-translated proteins and $^{32}$P-labeled in vitro-transcribed RNAs were mixed, allowed to bind for 15 min at 30°C, and immunoprecipitated as described [Lutz et al. 1990]. Coprecipitated RNAs were extracted with phenol/chloroform/isoamyl alcohol, precipitated with ethanol, and analyzed on 10% (18:1) polyacrylamide gels containing 8.3 M urea.

Nuclear extracts and in vitro polyadenylation reactions

HeLa cell nuclear extracts were prepared as described previously [Moore 1990]. In vitro polyadenylation reactions were performed as described above for UV cross-linking, except $1 \times 10^5$ cpm of substrate RNA was used, and the reactions were allowed to incubate at 37°C for 1 hr [Scheck et al. 1992]. Reaction products were extracted as above and analyzed on 40-cm 5% (29:1) polyacrylamide gels containing 7 M urea.

Oligoribonucleotides

Oligoribonucleotides were synthesized on an Applied Biosystems DNA synthesizer, model 380A, in the DNA Synthesis Facility of the Wistar Institute (Philadelphia, PA). An oligoribonucleotide representing the third USE motif of the SV40 late polyadenylation signal (Fig. 1), called USE oligo, had the sequence GCUUUAUUGUAACCAUUA. A mutant USE (mUSE) had the sequence GCUUUCUCUCUAACCAUUA. A nonspecific oligoribonucleotide was synthesized and had the sequence UAG-GUAACGUAGGGCU. An oligoribonucleotide representing the binding site on U1 RNA of the U1 snRNP-A protein was also synthesized (U1 oligo), and had the sequence UUAUCC-AUUUGCACUCGGGAUGU.

Bacterially produced U1 snRNP-A protein

U1 snRNP-A protein was made in the pET system (Novagen) by IPTG induction of transformed E. coli BL21(DE3) cultures. The cells were resuspended in 50 mM NaCl and 20 mM Tris-HCl (pH 8.0) and were lysed by lysosome treatment and freeze/thaw. The protein was then precipitated by addition of polyethyleneimine (PEI) to a final concentration of 1% PEI. The PEI pellet was washed twice with 1 mM NaCl, 20 mM Tris-Cl (pH 8.0), 2 mM EDTA to release the protein. The resulting supernatant was precipitated by 30% ammonium sulfate, and this precipitate was dialyzed against 100 mM NaCl, 20 mM Tris (pH 8.0). The dialyze was applied to a 1.5x10-cm Bio-Rex 70 column, and the column was developed with a salt gradient of 100-600 mM NaCl. Fractions were assayed by SDS-PAGE, and the peak fractions were used for immunization.

HPLC purification of antibodies

Rabbit polyclonal antibodies were purified by a two-step method. First, the antibodies were subjected to a 40% ammonium sulfate precipitation. Then, the precipitate was extensively dialyzed against 10 mM potassium phosphate buffer (pH 6.8) (buffer A), followed by HPLC purification on a 25 x 1-cm Hydapore-AX column (Rainin) as per the manufacturer's instructions. Briefly, the column was equilibrated with buffer A, the sample was loaded, and the sample was eluted with a 10-60% buffer B gradient (B = 300 mM potassium phosphate at pH 6.8). One-milliliter fractions were collected. Fractions 25-27 (part of the IgG peak) were pooled and concentrated approximately three-fold by use of Microcon-10 concentration units (Amicon) The final protein concentrations for both specific and control antibody samples were ~0.3 mg/ml of IgG. In addition, their profiles on silver-stained SDS-PAGE gels were identical.

Acknowledgments

We wish to thank Jack Keene for assistance in locating U1 snRNP-A constructs, Bill Wunner for preparing oligoribonucleotides, Matthias Görlich for aid in antibody purification, Steve White and Barbara Golden for the protein purification scheme, Pat O'Connor, Mike Malim, and Sherri Adams for reading the manuscript, and the members of the Alwine laboratory for helpful discussions. Cheers to all.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bandziulis, R.I., M.S. Swanson, and G. Dreyfuss. 1989. RNA-binding proteins as developmental regulators. Genes & Dev. 3: 431-437.

Bienroth, S., E. Wahle, C. Suter-Cazzolara, and W. Keller. 1991. Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. J. Biol. Chem. 266: 19768-19776.

Boelens, W.C., E.J.R. Jansen, W.J. van Venrooij, R. Stripecke, I.W. Mattaj, and S.I. Gunderson. 1993. The human snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. Cell 72: 881-892.

Brown, P.H., L.S. Tiley, and B.R. Cullen. 1991. Efficient polyadenylation within the human immunodeficiency virus type 1 long terminal repeat requires flanking U3-specific sequences. J. Virol. 65: 3340-3343.

Carswell, S. and J.C. Alwine. 1989. Efficiency of utilization of the Simian Virus 40 late polyadenylation site: Effects of upstream sequences. Mol. Cell. Biol. 9: 4248-4258.

Chiojou, H.C., C. Dabrowski, and J.C. Alwine. 1991. Simian virus 40 late mRNA leader sequences involved in augmenting mRNA accumulation via multiple mechanisms, including increased polyadenylation efficiency. J. Virol. 65: 6677-6685.

Cole, C.N. and T.P. Stacy. 1985. Identification of sequences in the herpes simplex virus thymidine kinase gene required for efficient processing and polyadenylation efficiency. Mol. Cell. Biol. 5: 2104-2113.

Conway, L. and M. Wickens. 1985. A sequence downstream of AUUAAA is required for formation of simian virus 40 late mRNA in 3' termini in frog oocytes. Proc. Natl. Acad. Sci. 82: 3949-3953.

DeZazzo, J.D. and M.J. Imperiale. 1989. Sequences upstream of AUUAAA influence poly[A] site selection in a complex transcriptional unit. Mol. Cell. Biol. 9: 4951-4961.

DeZazzo, J.D., J.E. Kilpatrick, and M.J. Imperiale. 1991. Involvement of long terminal repeat U3 sequences overlapping the transcriptional control region in human immunodeficiency virus type 1 mRNA 3' end formation. Mol. Cell. Biol. 11: 1624-1630.

Fresco, L.D., D.S. Harper, and J.D. Keene. 1991. Leucine periodicity of U2 small nuclear ribonucleoprotein particle [snRNP]...
A' protein is implicated in snRNP assembly via protein-protein interactions. *Mol. Cell. Biol.* **11**: 1578–1589.

Gil, A. and N.J. Proudfoot. 1984. A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3’ end formation. *Nature* **312**: 473–474.

——. 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA formation. *Cell* **49**: 399–406.

Grabowski, P.J., F.H. Nasim, H.-C. Kuo, and R. Burch. 1991. Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.* **11**: 5919–5928.

Gunderson, S.I., K.G. Murthy, and J.L. Manley. 1992. Characterization of the AAUAAA motif of mammalian mRNA 3' polyadenylation sites. *Methods Enzymol.* **200**: 410–418.

Kenan, D.J., C.C. Query, and J.D. Keene. 1991. RNA recognition motif: Towards identifying determinants of specificity. *Trends Biochem. Sci.* **16**: 214–220.

Kuo, H.-C., F.H. Nasim, and P.J. Grabowski. 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. *Science* **251**: 1045–1050.

Lerner, E.A., M.R. Lerner, C.A. Janeway, and J.A. Steitz. 1981. Monoclonal antibodies to nucleic-acid containing cellular constituents: Probes for molecular biology and autoimmune disease. *Proc. Natl. Acad. Sci.* **78**: 2737–2741.

Lührmann, R., B. Kastner, and M. Bach. 1990. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochim. Biophys. Acta* **1087**: 265–292.

Lutz-Freyermuth, C., M.J. Imperiale, H. Ali, and J.R. Nevins. 1984. Requirement of a downstream sequence for generation of a poly(A) addition site. *Cell* **37**: 993–999.

Lutz-Freyermuth, C., C.C. Query, and J.D. Keene. 1990. Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small nuclear ribonucleoprotein complex binds with high affinity to stem-loop II of U1 RNA. *Proc. Natl. Acad. Sci.* **87**: 6393–6397.

Manley, J.L. 1988. Polyadenylation of mRNA precursors. *Biochim. Biophys. Acta* **950**: 1–12.

McDevitt, M.A., M.J. Imperiale, H. Ali, and J.R. Nevins. 1984. Sequences capable of restoring poly(A) site function define two distinct downstream elements. *EMBO J.* **5**: 2907–2913.

Moore, C.L. 1990. Preparation of mammalian extracts active in polyadenylation. *Methods Enzymol.* **181**: 49–74.

Moore, C.L. and P.A. Sharp. 1984. Site-specific polyadenylation in a cell-free reaction. *Cell* **36**: 581–591.

Moore, M.J., C.C. Query, and P.A. Sharp. 1993. Splicing of precursors to messenger RNAs by the spliceosome. In *The RNA world* (ed. R.F. Gesteland and J.F. Atkins), pp. 303–358. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Murthy, K.G. and J.L. Manley. 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.* **267**: 14804–14811.

Nesic, D., J. Cheng, and L.E. Maquat. 1993. Sequences within the last intron function in RNA 3’ end formation in cultured cells. *Mol. Cell. Biol.* **13**: 3359–3369.

Niwa, M. and S.M. Berget. 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes & Dev.* **5**: 2086–2095.

Niwa, M., S.D. Rose, and S.M. Berget. 1990. In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes & Dev.* **4**: 1552–1559.

Pandey, N.B., N. Chodchoy, R.-J. Liu, and W.F. Marzluff. 1990. Introns in histone genes alter the distribution of 3’ ends. *Nucleic Acids Res.* **18**: 3161–3170.

Pinol-Roma, S., S.A. Adam, Y.D. Choi, and G. Dreyfuss. 1989. Ultraviolet-induced cross-linking of RNA to protein in vivo. * Methods Enzymol.* **180**, pp. 410–418.

Query, C.C., R.C. Bentley, and J.D. Keene. 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* **57**: 89–101.

Raju, V.S. and S.T. Jacob. 1988. Association of poly[A] polymerase with U1 RNA. *J. Biol. Chem.* **263**: 11067–11070.

Reich, C.I., R.W. VanHoy, G.L. Porter, and J.A. Wise. 1992. Mutations at the 3’ splice site can be suppressed by compensatory base changes in U1 snRNA. *Cell* **69**: 1159–1169.

Robberson, B.L., G.J. Cote, and S.M. Berget. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**: 84–94.

Russnak, R. 1991. Regulation of polyadenylation in hepatitis B viruses: Stimulation by the upstream activating signal PS 1 is orientation-dependent, distance-dependent, and additive. *Nucleic Acids Res.* **19**: 6449–6456.

Russnak, R. and D. Ganem. 1990. Sequences 5’ to the polyadenylation signal mediate differential poly(A) site use in hepatitis B virus. *Genes & Dev.* **4**: 764–776.

Sadovsky, M. and J.C. Alwine. 1984. Sequences on the 3’ side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. *Mol. Cell. Biol.* **4**: 1460–1468.

Sadovsky, M., S. Connelly, J.L. Manley, and J.C. Alwine. 1985. Identification of a sequence element on the 3’ side of AAUAAA which is necessary for simian virus 40 late mRNA 3’ end processing. *Mol. Cell. Biol.* **5**: 2713–2719.

Sanfourn, H., P. Brodmann, and T. Hohn. 1991. A dissection of the cauliflower mosaic virus polyadenylation signal. *Genes & Dev.* **5**: 141–149.

Schek, N., C. Cooke, and J.C. Alwine. 1992. Definition of the upstream efficiency element of the simian virus 40 late polyadenylation signal by using in vitro analyses. *Mol. Cell. Biol.* **12**: 5386–5393.

Scherly, D., W. Boelens, W.J. van Venrooij, N.A. Dathan, J. Hamm, and I.W. Mattaij. 1989. Identification of the RNA binding segment of the human U1A protein and definition of its binding site on U1 snRNA. *EMBO J.* **8**: 4163–4170.

Takagaki, Y., L.C. Rymer, and J.L. Manley. 1989. Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes & Dev.* **3**: 1711–1724.

Tsai, D.E., D.S. Harper, and J.D. Keene. 1991. U1 snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids Res.* **19**: 4931–4936.

Valsamakis, A., S. Zeichner, S. Carswell, and J.C. Alwine. 1991. The human immunodeficiency virus type 1 polyadenylation signal: A long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. *Proc. Natl. Acad. Sci.* **88**: 2108–2112.

Valsamakis, A., N. Schek, and J.C. Alwine. 1992. Elements upstream of the AAUAAA within the human immunodeficiency virus polyadenylation signal are required for efficient polyadenylation in vitro. *Mol. Cell. Biol.* **12**: 3699–3705.

Wahl, E. and W. Keller. 1992. The biochemistry of 3’-end cleavage and polyadenylation of messenger RNA precursors.
Lutz and Alwine

Annu. Rev. Biochem. 61: 419–440.
Wassarman, K.M. and J.A. Steitz. 1993. Association with terminal exons in pre-mRNAs: A new role for the U1 snRNP? Genes & Dev. 7: 647–659.
Wickens, M. 1990. How the messenger got its tail: Addition of poly[A] in the nucleus. Trends Biochem. Sci. 15: 277–281.
Wilusz, J. and T. Shenk. 1990. A uridylate tract mediates efficient heterogeneous nuclear ribonucleoprotein C protein-RNA cross-linking and functionally substitutes for the downstream element of the polyadenylation signal. Mol. Cell. Biol. 10: 6397–6407.
Wilusz, J., D.I. Feig, and T. Shenk. 1988. The C proteins of heterogeneous nuclear ribonucleoprotein complexes interact with RNA sequences downstream of polyadenylation cleavage sites. Mol. Cell. Biol. 8: 4477–4483.
Zhang, F. and C.N. Cole. 1987. Identification of a complex associated with processing and polyadenylation in vitro of herpes simplex type 1 thymidine kinase precursor RNA. Mol. Cell. Biol. 7: 3277–3286.
Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal.

C S Lutz and J C Alwine

Genes Dev. 1994, 8:
Access the most recent version at doi:10.1101/gad.8.5.576