MOLECULAR AND CELLULAR BIOLOGY, Dec. 1990, p. 6397–6407
0270-7306/90/126397-11$02.00/0

Vol. 10, No. 12

A Uridylate Tract Mediates Efficient Heterogeneous Nuclear Ribonucleoprotein C Protein-RNA Cross-Linking and Functionally Substitutes for the Downstream Element of the Polyadenylation Signal

JEFFREY WILUSZ[1]* AND THOMAS SHENK[2]

Department of Microbiology and Molecular Genetics, New Jersey Medical School, University of Medicine and Dentistry of New Jersey, Newark, New Jersey 07103,[1] and Department of Biology, Howard Hughes Medical Institute, Princeton University, Princeton, New Jersey 08544

Received 1 June 1990/Accepted 11 September 1990

Every RNA added to an in vitro polyadenylation extract became stably associated with both the heterogeneous nuclear ribonucleoprotein (hnRNP) A and C proteins, as assayed by immunoprecipitation analysis using specific monoclonal antibodies. UV-cross-linking analysis, however, which assays the specific spatial relationship of certain amino acids and RNA bases, indicated that the hnRNP C proteins, but not the A proteins, were associated with downstream sequences of the simian virus 40 late polyadenylation signal in a sequence-mediated manner. A tract of five consecutive uridylate residues was required for this interaction. The insertion of a five-base U tract into a pGEM4 polylinker-derived transcript was sufficient to direct sequence-specific cross-linking of the C proteins to RNA. Finally, the five-base uridylate tract restored efficient in vitro processing to several independent poly(A) signals in which it substituted for downstream element sequences. The role of the downstream element in polyadenylation efficiency, therefore, may be mediated by sequence-directed alignment or phasing of an hnRNP complex.

Precursors to mRNA are found in the nucleus associated with proteins to form heterogeneous nuclear ribonucleoprotein (hnRNP) complexes (reviewed in references 18 and 31). The formation of hnRNP particles occurs immediately after transcription (19, 21, 37). Many RNA processing events, including splicing and polyadenylation, most likely occur within these complexes and not on naked RNA. Experiments using hnRNP-specific antibodies to inhibit RNA splicing (11, 55), UV-cross-linking analysis (42, 70), and immunoprecipitation of nuclease-resistant RNA segments (61) clearly document an association of hnRNP proteins with splicing and polyadenylation substrate RNAs in vitro. The hnRNP particle, therefore, plays some role in pre-mRNA metabolism.

Purified 30S–40S hnRNP particles consist of six major proteins and 20 to 30 auxiliary proteins (6, 8, 10, 15, 20, 49, 68, 69). The A1/A2, B1/B2, and C1/C2 species comprise the major hnRNP proteins. Approximately 25 additional proteins are found associated with immunopurified hnRNP particles prepared with antibodies specific for several independent hnRNP proteins (10, 19, 47) and by two-dimensional gel electrophoresis (69). Many individual hnRNP proteins possess inherent nucleic acid-binding properties (12, 30, 47, 60).

The A1 protein is a basic, evolutionarily conserved protein with an amino-terminal domain that contains two RNP consensus sequences (1, 41, 50) and a glycine-rich COOH-terminal region (12, 13, 30). The A1 protein binds nucleic acids cooperatively and may possess helix-destabilizing properties. The hnRNP C proteins are evolutionarily conserved, antigenically related, acidic phosphoproteins that also contain an RNP consensus sequence as well as a putative nucleoside triphosphate-binding site (9, 19, 29, 40, 48, 62). Numerous multimeric complexes of the A and C proteins, as well as the other hnRNP proteins, have been described (4, 32). C proteins, for example, are normally found in 3C1:1C2 tetramers, with at least three of these units being present per isolated monoparticle (4). Finally, the C1 and C2 proteins, along with the A1 polypeptide and most of the hnRNA, are located in the peripheral portion of intact hnRNP particles (32).

The role of hnRNP particles in RNA metabolism remains to be elucidated. A multicomponent complex of 30 proteins, some of which possess inherent, somewhat selective RNA-binding capacities, clearly has the potential to interact with RNA in many different ways. The fact that several hnRNP proteins show homopolymer-binding preferences (60) and cooperativity (12) leads to the possibility that hnRNP particles may interact with RNA in a specific fashion. Additional lines of biochemical (2, 44, 46, 58, 59, 67) and microscopic (5, 7, 33, 45, 66) evidence support the notion of a nonrandom association of hnRNP particles with RNA. The complex array of RNA-binding proteins of the hnRNP particle could, perhaps, allow it to preferentially recognize and interact with specific RNA sequences. Since hnRNP particles contain many independent RNA-binding proteins, however, this RNA interaction may appear to be essentially random over the general population.

We have recently shown that hnRNP C proteins can be cross-linked by UV light to sequences downstream of several independent polyadenylation sites more efficiently than to random RNA sequences (70). This downstream region contains an element that is not absolutely required for polyadenylation but plays a role in the efficiency of the reaction in vivo and in vitro (14, 16, 23, 26, 27, 34–36, 38, 52–54). The sequence requirements for this element are unclear, as comparison of several independent downstream elements reveals only a U- or GU-rich consensus, and

* Corresponding author.

experiments using fine deletion mapping indicate that the element may be reiterated (73). The efficient cross-linking of C proteins with downstream sequences (70) leads to the possibility that an hnRNP complex, with its multifaceted RNA-binding capacity, may interact with the downstream element in a sequence-mediated fashion and influence the processing event.

In this report, we characterize the association of hnRNP proteins with the simian virus 40 late (SVL) polyadenylation signal. While immunoprecipitation evidence indicated that the hnRNP A and C proteins were associated with all RNAs tested, UV cross-linking studies clearly demonstrated a sequence dependence in the C protein-RNA association. The fact that C proteins were associated with but not efficiently cross-linked to all RNAs, as well as the observation that other hnRNP proteins were quantitatively associated with all RNAs, leads us to propose that the sequence-directed cross-linking of C proteins with RNA substrates is the result of a specific alignment of proteins within the hnRNP complex with respect to the RNA bases. A tract of at least five consecutive uridylate residues was required for the cross-linkable alignment of C proteins with respect to the RNA bases. This sequence requirement was further supported by cross-linking experiments performed by using a short, polylinker-derived RNA containing a UUUUU substitution. Finally, the polylinker-derived sequence containing the UUUUU tract was able to functionally replace downstream element sequences, mediating efficient processing with three independent polyadenylation signals. We propose that the interaction between the downstream sequence element and C proteins represents a specific alignment or phasing of the substrate RNA with an hnRNP complex and that this alignment allows the enzymatic events of polyadenylation to occur efficiently.

MATERIALS AND METHODS

RNAs and plasmids. All RNAs were transcribed in vitro by using SP6 polymerase (39) and [^{32}P]UTP so that, on the average, 1 in every 10 uridylate residues would be radioactively labeled. RNAs were purified from 5% acrylamide–7 M urea gels prior to use.

Transcripts were derived as follows. pSVL contains the 241-bp *Bam*HI-*Bcl*I fragment of simian virus 40 (sequence positions 2533 to 2770) inserted into the *Bam*HI site of pSP65 (Promega). Transcription of *Dra*I-, *Hpa*I-, or *Fnu*4HI-linearized templates yields RNAs of 224 bases (SVL), 163 bases (SVL1), or 149 bases (SVL2), respectively. pSVL3 contains the simian virus 40-specific *Alu*I-*Hin*dIII fragment of pSVL inserted into pGEM4 DNA (Promega) digested with *Hin*cII and *Hin*dIII. Transcription of *Dra*I-cut template gives a 122-base RNA (SVL3). pSVL6 was constructed by ligation of a synthetic oligonucleotide into *Bam*HI-*Hin*cII-cleaved pSVL3 DNA as previously described (70). Transcription of *Dra*I-cut DNA gives a 111-base RNA (SVL6). pSVL7 was produced by inserting the *Hin*cII-to-*Xba*I fragment of pSVL into pGEM4 DNA cut with *Sma*I and *Xba*I. Transcription of *Dra*I-cleaved template gives an 89-base RNA (SVL7). The *Bsm*I-to-*Eco*RI fragment of pSVL7 was removed, and the remaining large segment was blunted with Klenow polymerase and recircularized to produce pSVL8. Transcription of *Dra*I-cut DNA gives a 61-base RNA (SVL8). To make pSVL9, the synthetic oligonucleotide 5′-AGCTTACCTG AAACATAAATGAATGCA-3′, along with its complement, was inserted into pSVL3 DNA cleaved with *Bsm*I and *Hin*dIII. Transcription of *Pvu*II-cut template gives a 144-base RNA (SVL9). pSVL10 was constructed by removing the *Hin*cII fragment from pSVL3 and recircularizing the plasmid. Transcription of *Pvu*II-linearized template gives a 121-base RNA (SVL10). pC220 contains a 1.38-kbp *Eco*RI-to-*Hin*dIII fragment of bacteriophage lambda (sequences 26104 to 27479) cloned into the same sites of pSP64. Transcription of *Sca*I-cut DNA yields a 220-base control RNA.

Derivatives of SVL3 RNA containing point mutations in the +50 to +55 uridylate tract were constructed by cloning appropriate synthetic oligonucleotides containing an equal mixture of G or T residues at positions +50 through +52 into *Bsm*I-*Hin*dIII-cut DNA. Transcription of *Dra*I-cut templates yields a 122-base RNA with the indicated point mutations. All mutations derived with synthetic oligonucleotides were confirmed by chemical sequence analysis (data not shown).

Transcription of *Hin*dIII-linearized pGEM4 DNA gives a 65-base RNA (GEM4). p4U5 was constructed by inserting the oligonucleotide 5′-TCGACAAAAAG-3′ and its appropriate complement into pGEM4 DNA cut with *Bam*HI and *Sal*I. Transcription of *Hin*dIII-cut template yields a 64-base RNA (4U5).

Polyadenylation signals containing substitutions for downstream sequences were constructed as follows. pSVL-GEM was prepared by inserting the *Sma*I-to-*Hin*dIII fragment of pGEM4 into *Bsm*I-*Hin*dIII-cleaved pSVL. Transcription of *Hin*dIII-linearized template yields a 214-base RNA (SVL-GEM). pSVL-4U5 was made by inserting the *Sma*I-to-*Hin*dIII fragment of p4U5 into *Bsm*I-*Hin*dIII-cut pSVL. Transcription of *Hin*dIII-linearized DNA gives a 213-base RNA (SVL-4U5). pSVE was prepared by inserting the *Hpa*I-to-*Bam*HI fragment of pSVL into *Sma*I-*Bam*HI-cut pGEM4. Transcription of the *Bam*HI-linearized template gives a 162-base RNA (SVE). pSVE-GEM was prepared by inserting the *Sma*I-to-*Hin*dIII fragment of pGEM4 into pSVE that had been cleaved with *Bsm*I (and then blunt ended with Klenow polymerase) and *Hin*dIII. Transcription of *Hin*dIII-linearized DNA gives a 143-base RNA (SVE-GEM). pSVE-4U5 was made in a similar fashion by inserting the *Sma*I-to-*Hin*dIII fragment of p4U5 into the *Bsm*I and *Hin*dIII sites of pSVE. Transcription of *Hin*dIII-cut DNA gives a 142-base RNA (SVE-4U5).

Polyadenylation and cleavage assays. In vitro processing assays were performed by using the system of Moore and Sharp (43) as previously described (71). Comparisons of processing efficiency were performed at times during which product was accumulating linearly to ensure valid conclusions.

UV-cross-linking and label transfer analysis. Nuclear salt wash extracts were prepared from HeLa spinner cells grown in 10% horse serum as described previously (17). RNAs were incubated with HeLa nuclear extract under in vitro polyadenylation conditions favoring the cleavage reaction (3% polyvinyl alcohol, 1 mM EDTA, 40 to 60% extract) for 10 min. These conditions were chosen to ensure the stability of the added substrate RNAs. Equal amounts or equal counts per minute of RNAs labeled with UTP of the same specific activity were used to ensure equivalent concentrations of ^{32}P-labeled U residues for valid comparisons. Reaction mixtures were irradiated on ice, using a 15-W germicidal light (71). Then 10 μg of RNase A was added, and the mixture was incubated at 37°C for 15 min. An equal volume of 2× protein sample loading buffer was added; reaction mixtures were boiled for 5 min and loaded onto a 10% acrylamide gel containing sodium dodecyl sulfate (SDS).

Antibody production. Monoclonal antibodies to the hnRNP A1 protein were prepared by using purified rat A1

protein produced in *Escherichia coli* (12), kindly provided by B. Merrill and K. Williams. BALB/c mice were immunized by using two intraperitoneal injections of 250 μg of A1 protein 14 days apart, followed by a 100-μg boost prior to fusion. Cell fusions and selections were performed as described previously (25). Positive hybridoma supernatants were identified by using a nitrocellulose dot blot assay (28) and alkaline phosphatase-conjugated rabbit anti-mouse immunoglobulin G (Sigma). Hybridomas were cloned twice by limiting dilution and grown in RPMI medium supplemented with 10% fetal bovine serum and glutamine for antibody production. The monoclonal antibody used in this study (1A1) recognized both A1 and A2 proteins by Western immunoblot analysis as well as by immunoprecipitation analysis (65; data not shown).

Antibodies to the hnRNP C proteins were produced in a similar manner, using 50 μg of affinity-purified, UV-cross-linked material as the antigen source. The affinity-purified antigen was generated as follows. Biotinylated SVL3 RNAs were prepared by inclusion of biotinylated UTP (Sigma Chemical Co.) in the in vitro SP6 transcription reaction at a ratio of one biotinylated UTP molecule to five UTP molecules. This resulted in SVL3 RNA with one to two biotinylated residues per molecule. Biotinylated SVL3 RNA was incubated with the HeLa nuclear extract, and UV cross-linking was performed as described above except on a larger scale. Reaction mixtures were then adjusted to 0.1% SDS and 400 mM NaCl, boiled for 5 min, centrifuged to eliminate precipitated proteins, and incubated with streptavidin-agarose (Bethesda Research Laboratories) at room temperature for 30 min, using a batch technique. The agarose beads were pelleted at 3,000 × *g* and washed eight times with a buffer containing 25 mM Tris (pH 7.6), 400 mM NaCl, and 0.1% SDS. Proteins were eluted from the beads by treatment with RNase A (10 μg) at 37°C for 30 min and concentrated by precipitation with organic solvents prior to injection into mice. The monoclonal antibody used in this study (1B12) specifically immunoprecipitated both the C1 and C2 proteins (Fig. 1B).

Immunoprecipitation. After UV cross-linking and RNase digestion, 200 μl of buffer A (25 mM Tris [pH 7.6], 75 mM NaCl, 0.05% Nonidet P-40) was added, and precipitated material was pelleted by centrifugation. The reaction mixture was transferred to a fresh tube, and 100 μl of monoclonal antibody supernatant was added. After incubation on ice for 2 h, 0.5 μl of rabbit anti-mouse immunoglobulin G (Cappel) was added, and the mixture was incubated on ice for 30 min. Then 50 μl of a 10% solution of Formalin-fixed *Staphylococcus aureus* cells (Pansorbin; Calbiochem) was added, and this mixture was incubated for an additional 10 min. Immune complexes were pelleted and washed four times with buffer A. Pellets were resuspended in 25 μl of sample loading buffer and heated at 100°C for 5 min prior to analysis on a 10% acrylamide gel containing 0.1% SDS.

For immunoprecipitation analysis of RNA, samples were incubated in the polyadenylation system and precleared with Pansorbin prior to addition of antibody. After immunoprecipitation, washed pellets were resuspended in 300 μl of buffer A, 5 μg of tRNA was added, and samples were extracted with phenol-chloroform-isoamyl alcohol. RNAs released into the aqueous phase were concentrated with ethanol and analyzed on a 5% polyacrylamide gel containing 7 M urea.

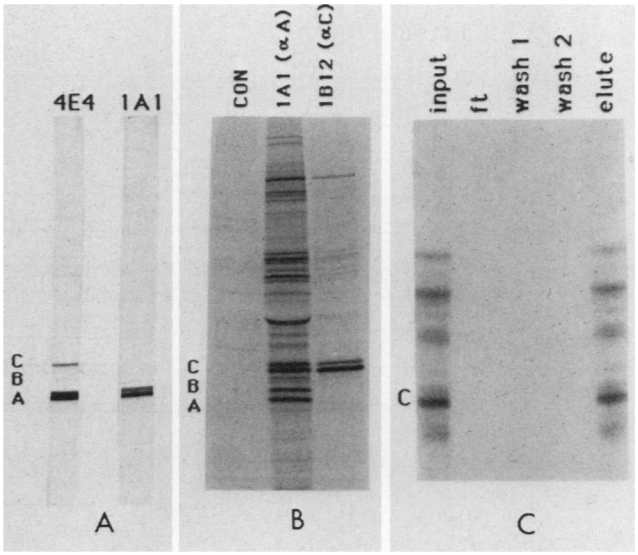

FIG. 1. Production and characterization of hnRNP protein-specific monoclonal antibodies. (A) Western analysis of HeLa cell nuclear proteins, using representatives of the two classes of antibodies obtained by using recombinant A1 protein. Positions of the major hnRNP proteins are indicated. (B) Immunoprecipitation analysis of total HeLa cell proteins labeled with [^{35}S]methionine. Monoclonal antibodies used are indicated above the lanes. Lane CON refers to preimmune mouse serum. Precipitated proteins were analyzed on a 10% acrylamide gel containing 0.1% SDS. (C) Demonstration of the biotin-streptavidin affinity purification technique used to purify the hnRNP C proteins for antibody production. Proteins were cross-linked to ^{32}P-labeled, biotinylated SVL3 RNA. Reactions were then incubated with streptavidin-agarose, washed with a high-salt buffer, and eluted with RNase. Cross-linked proteins present in each fraction were analyzed on a 10% acrylamide gel containing SDS. ft, Flowthrough.

RESULTS

Production of monoclonal antibodies specific for the hnRNP A and C proteins. The hnRNP C proteins are part of a multicomponent complex of approximately 30 proteins referred to here as an hnRNP monoparticle. We have previously correlated the presence of sequences downstream of several polyadenylation or cleavage sites with the ability of the RNA to be efficiently cross-linked by UV light to the hnRNP C proteins (70). To characterize this interaction by additional criteria, as well as to investigate the role of other components of the hnRNP monoparticle in the interaction with polyadenylation substrate RNAs, we produced monoclonal antibodies to several hnRNP-specific polypeptides.

Twenty monoclonal antibodies to A1 protein were generated by using a recombinantly produced rat A1 protein. These antibodies fell into two groups on the basis of Western blotting to HeLa nuclear proteins (Fig. 1A). The first class, represented by 4E4, recognized all six of the major hnRNP proteins in Western blot analysis but failed to work in immunoprecipitation analyses (data not shown). The second class, represented by 1A1, recognized both A1 and A2 proteins by Western analysis and by immunoprecipitation (Fig. 1B). These proteins share considerable similarity (76 to 85%) in their RNA-binding domains (9). Immunoprecipitations with the 1A1 monoclonal antibody consistently coprecipitated other proteins, many of which comigrated with major components of the hnRNP monoparticle (Fig. 1B). The extensive similarity between the hnRNP A2 and B1

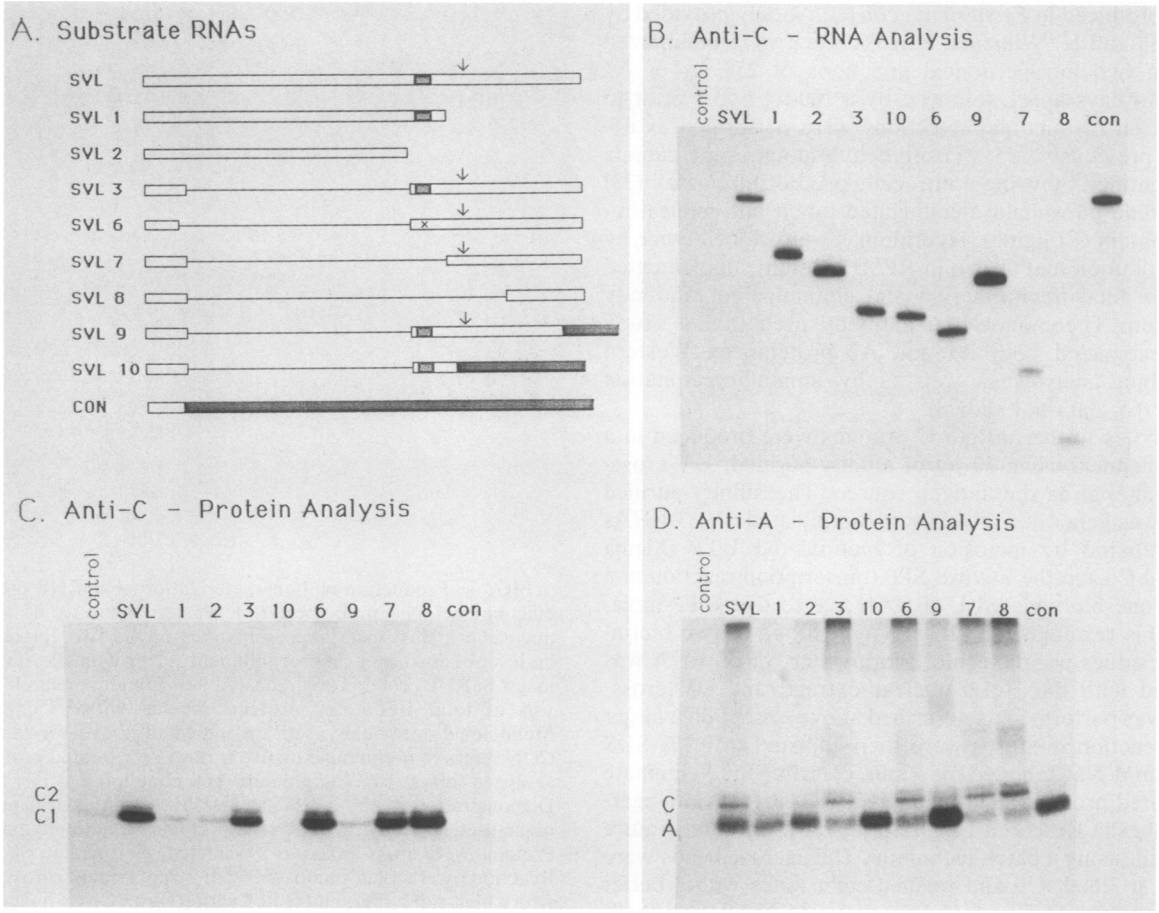

FIG. 2. Analysis of RNAs associated with hnRNP A and C proteins by coimmunoprecipitation and UV-cross-linking assays. (A) Diagrams of the SVL transcript and its derivatives. Symbols: ▩, AAUAAA motif; ↓, cleavage site; ■, non-SVL-specific sequence. SVL, SVL3, SVL6, SVL7, and SVL 8 all contain the +30 to +55 RNA segment discussed in the text. (B) RNA analysis. Radiolabeled RNAs were incubated in the in vitro polyadenylation system and immunoprecipitated by using the 1B12 C protein-specific monoclonal antibody. Coprecipitated RNAs were analyzed in a 5% acrylamide gel containing 7 M urea. (C and D) Immunoprecipitation analysis of proteins cross-linked to derivatives of the SVL RNA, using the hnRNP C or A protein-specific monoclonal antibody. Radiolabeled RNAs were incubated in the in vitro system and cross-linked to proteins by UV light. After digestion with RNase A, samples were incubated with a monoclonal antibody specific for the C (C) or A (D) proteins. Immune complexes were precipitated by using Formalin-fixed *S. aureus* cells. Proteins were analyzed in 10% acrylamide gels containing 0.1% SDS. Lanes marked "control" represent immunoprecipitations using a preimmune mouse serum and the SVL3 RNA; lanes marked "con" contained lambda-derived control RNA (see panel A).

proteins, as well as the possibility that the 1A1 antibody recognizes features of the RNP consensus domain, may have contributed to this complex immunoprecipitation pattern. We have not, however, detected any cross-reaction of the 1A1 antibody with the hnRNP C protein (data not shown).

The source of C proteins for antibody production was generated by using an affinity purification approach (Fig. 1C) that takes advantage of the covalent nature of the UV-induced RNA-protein cross-link and the strength of the biotin-streptavidin interaction. After UV cross-linking using biotinylated SVL-3 RNA (Fig. 2A), reaction mixtures were adjusted to 400 mM NaCl and 0.1% SDS to disrupt most of the noncovalent RNA-protein interactions. Extracts were immediately boiled to inhibit RNases. The proteins cross-linked to biotinylated transcripts were selectively precipitated by association with streptavidin-agarose, and then cross-linked proteins were specifically eluted from the matrix by RNase digestion. They could be purified further by gel electrophoresis or, in the case presented here, used directly for monoclonal antibody production. The selectivity of this approach is illustrated in Fig. 1C. Although overall yields are low, this method offers a rapid alternative to conventional purification of RNA-binding proteins identified by UV-cross-linking analysis. The C protein-specific antibody obtained through this procedure immunoprecipitated both C1 and C2 proteins (Fig. 1B) but did not work well in Western analyses (data not shown).

hnRNP C proteins, but not hnRNP A proteins, exhibit sequence-mediated interactions with a polyadenylation substrate RNA. To characterize the association of hnRNP proteins with the SVL polyadenylation signal, a series of deletion derivatives ranging in size from 63 to 224 bases was constructed (Fig. 2A). The combination of these overlapping deletions was chosen so that sequence-specific interactions could be localized to a specific RNA segment.

The SVL derivatives were incubated with HeLa nuclear extract in the in vitro polyadenylation system in the presence of EDTA, followed by analysis of RNAs coimmunoprecipitated by using the hnRNP-specific monoclonal antibodies described above. All RNAs tested, including the nonspecific lambda-derived transcript, were coprecipitated with use of the C (Fig. 2B) or A (data not shown) protein-specific

antibody. The SVL7 and SVL8 transcripts were coprecipitated somewhat less efficiently in this experiment than were the other transcripts, perhaps because of their relatively small size. We suspect that both hnRNP proteins were associated with individual RNA molecules for two reasons. First, all transcripts were nearly quantitatively coimmunoprecipitated with use of either the A or C protein-specific antibody. Second, each RNA added to the nuclear extract became associated with a large complex, sedimenting at 20S to 25S in a sucrose gradient (data not shown). Thus, it is likely that the hnRNP A and C proteins, and most likely a form of hnRNP monoparticle, assembled on each RNA, regardless of its sequence. Since the exact nature of an in vivo hnRNP particle is still unclear (18), for convenience we will subsequently refer to the complex that forms on these RNAs as an hnRNP complex.

The ability of the various SVL derivatives to interact with the hnRNP A and C proteins was further analyzed by UV-cross-linking analysis. This procedure assesses not simply RNA binding; the efficiency of cross-linking also depends on the proximity of amino acids within the bound protein to the RNA bases (56).

Equimolar amounts of RNAs radiolabeled with UTP of the same specific activity were incubated in the in vitro polyadenylation system for 10 min to allow specific complex formation. Mixtures were irradiated with UV light to cross-link closely associated RNA-binding proteins, followed by RNase digestion. Prior to gel electrophoresis, cross-linked proteins were immunoprecipitated by using the hnRNP A or C protein-specific antibodies.

The hnRNP A and C proteins were selectively cross-linked to different SVL deletion derivatives (Fig. 2C and D). A comparison of the derivatives efficiently cross-linked to the C proteins (SVL3, SVL6, SVL7, and SVL8; Fig. 2C) showed that the interaction was indeed sequence mediated and required the +30 to +55 sequence block described previously (70). The hnRNP A proteins were efficiently cross-linked to most of the SVL RNA derivatives (Fig. 2D), but the pattern of efficient cross-linking did not implicate any specific RNA segment in the interaction. Curiously, specific cross-linking of C proteins to SVL derivatives of 125 bases or less resulted in a decrease in the nonspecific cross-linking to A proteins. In addition, the lambda-derived control RNA was very efficiently cross-linked to the hnRNP A proteins but not to the C proteins. The A proteins would apparently cross-link nonspecifically to RNAs incubated in the nuclear extract unless they were specifically prohibited from doing so by a competing polypeptide.

We conclude that the hnRNP C proteins, but not the A proteins, show a sequence-mediated association with the SVL polyadenylation signal as assayed by UV cross-linking. Since the results of our coimmunoprecipitation studies (Fig. 2B) indicated that both proteins were associated with all RNAs tested, it is the alignment or phasing of the hnRNP C proteins (as components of an hnRNP complex) with respect to the RNA bases which must be sequence directed. The lack of sequence specificity observed by immunoprecipitation analysis, therefore, may be due to the fact that the hnRNP proteins exist as constituents of hnRNP complexes. While the complex as a whole exhibits nonspecific RNA binding, the alignment of the individual proteins of the hnRNP complex with respect to the RNA bases or the overall phasing of hnRNP complexes appears to be influenced by specific sequences.

A stretch of five uridylate residues permits efficient cross-linking of hnRNP C proteins. Previous work by Swanson and Dreyfuss (60) has shown that C proteins bind tightly to a poly(U) agarose matrix, implicating U-rich sequences in efficient C protein-RNA interactions. Our earlier mutational analyses of the uridylate stretches in the +30 to +55 downstream element of the SVL polyadenylation signal were consistent with this suggestion (70). To further delineate the sequence requirements for the cross-linkable C protein interaction with this RNA segment, a series of single-point mutations in the six-base uridylate stretch from +50 to +55 of the SVL transcript was constructed (Fig. 3A). These mutants contain U to G transversions which result in a progressive reduction in the number of consecutive uridylate residues found in this region of the transcript. Results of cross-linking analyses performed with these mutants are shown in Fig. 3B. Shortening of the U tract to five bases by a U to G change at position +50 had no detectable effect on C protein-SVL RNA cross-linking. A transversion at position +51, which leaves only four consecutive U residues, reduced cross-linking efficiency by about fivefold. An interruption of the tract at position +52, leaving only three consecutive U residues, reduced C protein-RNA cross-linking to background levels. Finally, these point mutations had no effect on the cross-linking of other proteins to the SVL transcript (Fig. 3B). The slight reduction in cross-linking to larger material in the gel is simply a result of the reduced cross-linking to C proteins because these bands are generated by nonspecific covalent modifications and protein-protein cross-links during UV irradiation.

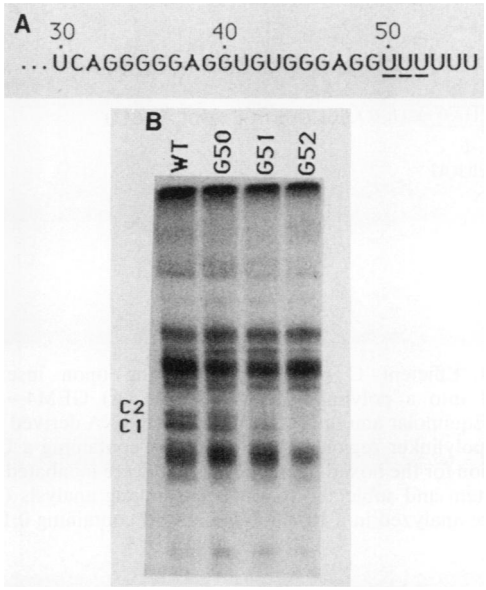

FIG. 3. Effects of point mutations in the +50 to +55 uridylate tract of the SVL3 RNA on C protein-RNA cross-linking. Equimolar amounts of radiolabeled SVL3 RNA or transcripts containing the indicated U to G transversions in the SVL downstream sequence (A) were incubated in the in vitro polyadenylation system and subjected to UV-cross-linking analysis (B). Labeled proteins were electrophoresed in a 10% acrylamide gel containing 0.1% SDS. The identity of the C protein-specific bands was confirmed by using monoclonal antibodies (data not shown). WT, wild type.

A stretch of five or more uridylate residues, in the context of SVL-specific downstream sequences, was therefore required for an efficient, cross-linkable interaction of the hnRNP C protein. The observation that a four-base uridylate stretch located from +15 to +18 relative to the SVL cleavage site (see Fig. 5A) had no effect on the interaction of C

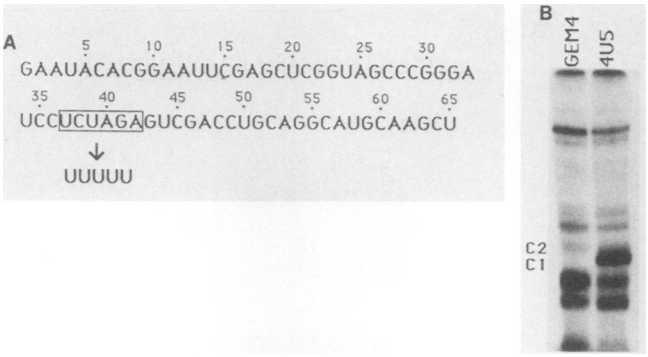

FIG. 4. Efficient C protein cross-linking upon insertion of UUUUU into a polylinker-derived RNA. (A) GEM4→4U5 sequence. Equimolar amounts of a radiolabeled RNA derived from the pGEM4 polylinker region (GEM4) and one containing a UUUUU substitution for the boxed nucleotides (4U5) were incubated in the in vitro system and subjected to UV-cross-linking analysis (B). Proteins were analyzed in a 10% acrylamide gel containing 0.1% SDS.

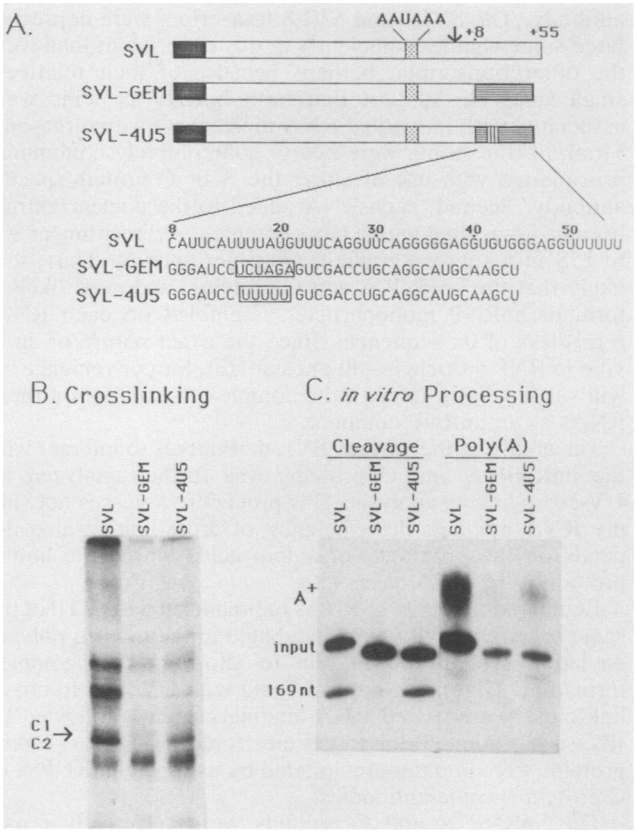

FIG. 5. Demonstration that UUUUU can substitute for the downstream element of the SVL poly(A) signal in mediating processing efficiency. (A) Downstream sequences of the SVL poly(A) signal were replaced by polylinker-specific sequences (SVL-GEM) or polylinker-derived sequences containing the UUUUU substitution (SVL-4U5). (B) Equimolar amounts of RNAs were incubated in the in vitro polyadenylation system and subjected to UV-cross-linking analysis. (C) Equimolar amounts of the SVL RNA downstream derivatives were incubated in the in vitro polyadenylation system in the presence of AMP[CH2]PP and EDTA (to assay cleavage only) or with ATP and phosphocreatine (to assay the complete processing reaction). RNA products of the reaction were analyzed on a 5% acrylamide gel containing 7 M urea.

protein with the RNA (70) is consistent with this conclusion. Finally, the addition of plasmid-specific sequences 3′ to the +50 to +55 uridylate tract had no effect on C protein-RNA cross-linking (data not shown). This finding makes it unlikely that the 3′ location of the U tract in the SVL transcript has a major influence on efficient C protein cross-linking.

To test whether this observation could be generalized to non-poly(A) signal-containing RNAs, we replaced the *Xba*I site in the pGEM4 polylinker region with a five-base uridylate stretch by oligonucleotide mutagenesis (GEM4 versus 4U5; Fig. 4A). This results in a runoff transcript from *Hin*dIII-linearized DNA of 64 nucleotides with only a 5% increase in U content (to 23%) compared with the pGEM4-derived 65-nucleotide transcript. UV-cross-linking analysis was performed by using equal amounts of GEM4 and 4U5 transcripts labeled with UTP of the same specific activity (Fig. 4B). The substitution of UCUAGA with UUUUU had a dramatic effect on the cross-linking of C proteins to the RNA, while cross-linking to other proteins remained unchanged. Furthermore, both RNAs were nearly quantitatively precipitated by using either the anti-C or anti-A monoclonal antibody (data not shown). We conclude that in at least two unrelated contexts, a UUUUU tract is sufficient to align the binding of the hnRNP C proteins (present as part of an hnRNP complex) to RNA in a cross-linkable manner. While we cannot rule out the possibility that other elements contribute to the apparent hnRNP phasing that we observe, it is unlikely that such elements would be present in the polylinker-derived 4U5 transcript.

The C-protein alignment element UUUUU can efficiently substitute for the downstream element of several independent polyadenylation signals. We have shown previously that the regions downstream of six different poly(A) cleavage sites play some role in efficient cross-linking of hnRNP C proteins (70), and it is well established that downstream domains contribute to the efficiency of polyadenylation. This fact raised the possibility that the polyadenylation reaction might be modulated by sequence-mediated alignment or phasing of hnRNP complexes, perhaps at elements such as UUUUU.

Our previous study showed that while a 25-base deletion in the SVL polyadenylation signal (+30 to +55) resulted in a fourfold decrease in in vitro processing efficiency, point mutations in the two uridylate tracts in this segment had no effect on polyadenylation efficiency, even though they dramatically decreased cross-linking to the hnRNP C proteins (70). Since the downstream element of the SVL polyadenylation signal is known to be reiterated (35, 52, 73), it remains possible that the sequence-directed alignment of hnRNP proteins mediates polyadenylation efficiency through downstream element interactions. Other RNA-binding proteins of the hnRNP complex could, perhaps, still maintain an optimal alignment or phasing through a different sequence, e.g., the GU-rich sequence located just upstream from the UUUUU tract in this 25-base segment.

This model, postulating a role for hnRNP complex alignment or phasing in the function of the downstream element, was first tested in the context of the SVL polyadenylation substrate. The entire downstream domain of this RNA (+8 to +55 relative to the cleavage site) was substituted with a *Sma*I to *Hin*dIII fragment containing either the pGEM4 polylinker (SVL-GEM; Fig. 5A) or the variant containing the UCUAGA to UUUUU substitution (SVL-4U5; Fig. 5A). Such a deletion or substitution would effectively remove the entire downstream element and therefore allow a definitive test of the role of hnRNP complex alignment or phasing on

polyadenylation efficiency. These RNAs were first assayed for their protein associations by UV cross-linking (Fig. 5B). As expected, SVL-GEM RNA was cross-linked to C proteins with much lower efficiency than was wild-type SVL RNA. SVL-4U5 RNA was cross-linked to C proteins very efficiently. The variant RNAs were next tested for their ability to be processed in cell-free polyadenylation extracts (Fig. 5C). Both SVL-GEM and SVL-4U5 RNAs were unstable when added to nuclear extracts in the presence of ATP and Mg^{2+}. This phenomenon has been noted previously for SVL derivatives lacking the downstream domain (70). Nevertheless, it was clear that the SVL-4U5 RNA was polyadenylated with greater efficiency than was SVL-GEM RNA [Fig. 5C, poly(A)]. In fact, the same proportions of wild-type SVL and SVL-4U5 RNAs were polyadenylated. The ability of the UUUUU sequence to substitute for the downstream element of SVL was confirmed by monitoring the endonucleolytic cleavage reaction. The SVL RNA derivatives were stable in the cleavage reaction, which was performed in the presence of EDTA. SVL-GEM RNA was cleaved at a fivefold-reduced efficiency to produce a 169-nucleotide 5′ fragment (the small 3′ fragment was run off the end of the gel), while the SVL-4U5 RNA was cleaved as efficiently as was normal SVL substrate RNA (Fig. 5C).

The ability of the UUUUU tract to substitute for the downstream element in a second substrate RNA was tested by using the simian virus 40 early (SVE) polyadenylation signal. SVE RNA was chosen for two reasons. First, the efficiency with which this signal is processed has been shown to be strongly influenced by its downstream domain (26, 27). Second, its downstream domain is required for efficient cross-linking to the hnRNP C proteins (70), but it does not include a UUUUU tract (Fig. 6A). C protein cross-linking to the wild-type SVE RNA (Fig. 6B) was not as efficient as that observed with the SVL transcript (Fig. 5B). The level of cross-linking to the hnRNP C proteins was, however, clearly influenced by the presence of the SVE downstream sequences (Fig. 6B). The SVE downstream domain (−3 to +52 relative to the cleavage site) was replaced with either GEM- or 4U5-derived sequences (Fig. 6A). SVE-4U5 RNA, but not SVE-GEM RNA, served as an efficient substrate for both C-protein cross-linking (Fig. 6B) and in vitro polyadenylation (Fig. 6C).

Finally, we also replaced the downstream domain of the adenovirus type 5 L3 polyadenylation signal (+10 to +67 relative to the cleavage site) with either the GEM- or 4U5-derived segment. Similar levels of recovery in the efficiency of polyadenylation were observed with the L3-4U5 RNA compared with those of the L3-GEM transcript (data not shown).

We conclude that the UUUUU element, which apparently is involved in hnRNP complex phasing or the alignment of the proteins in an hnRNP particle and permits efficient cross-linking to the hnRNP C proteins, was also able to substitute for the downstream element of several independent polyadenylation signals to mediate efficient in vitro processing.

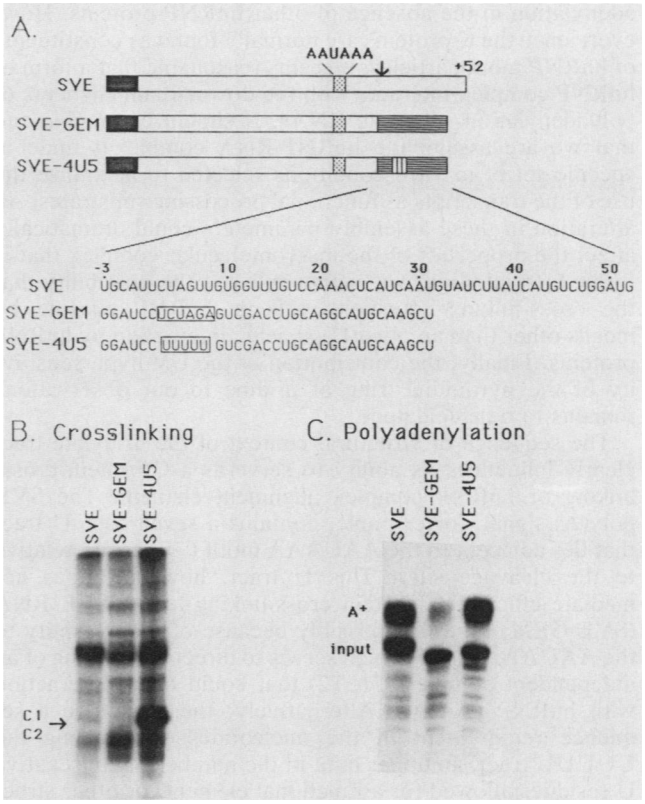

FIG. 6. Demonstration that UUUUU can substitute for the downstream element of the SVE poly(A) signal in mediating processing efficiency. (A) Downstream sequences of the SVE polyadenylation signal were replaced with polylinker-specific sequences (SVE-GEM) or polylinker-derived sequences containing the indicated UUUUU substitution (SVE-4U5). (B) Equimolar amounts of RNAs were incubated in the in vitro polyadenylation system and subjected to UV-cross-linking analysis. (C) SVE RNA downstream derivatives were incubated in the complete in vitro polyadenylation system. RNA products of the processing reaction were analyzed on a 5% acrylamide gel containing 7 M urea.

DISCUSSION

The data presented above describe several findings which extend previous studies of C protein-RNA interactions. We describe a one-step affinity approach for purifying proteins identified by UV-cross-linking analysis. Two new monoclonal antibodies reagents are described that proved useful in these studies and will clearly have future applications in the study of hnRNP proteins. Evidence is presented that hnRNP proteins in addition to the C proteins are stably associated with polyadenylation substrate RNAs. We have extended previous observations of the properties of C protein-RNA interactions based on chromatography on homopolymer affinity columns by fine mapping the minimal requirements for efficient C protein-SVL RNA interaction. Finally, we present a positive correlation between C protein cross-linking and efficient in vitro RNA processing of our model substrates.

Two main conclusions can be drawn from this study. First, a stretch of five consecutive uridylate residues, present either in the normal context of the downstream domain of the SVL polyadenylation signal or in a randomly selected polylinker-derived sequence, can mediate efficient UV cross-linking of hnRNP C proteins to RNA (Fig. 3 and 4). Results of coimmunoprecipitation studies (Fig. 2) suggest that this interaction most likely reflects an hnRNP complex-RNA interaction rather than simply a C protein-RNA interaction. Second, this UUUUU tract, embedded in the polylinker sequence, can substitute for the downstream element in several different polyadenylation substrate RNAs (Fig. 5 and 6). We cannot rule out the possibility that C proteins interact with the downstream element and influence poly-

adenylation in the absence of other hnRNP proteins. However, since the C proteins are normally found as constituents of hnRNP monoparticles, it seems reasonable that a form of hnRNP complex interacts with the downstream elements of polyadenylation substrate RNAs. It should be emphasized that we are assembling hnRNP-RNA complexes under a specific set of in vitro conditions selected to maximize the use of the transcripts as functional processing substrates. An alteration in these assembly parameters could dramatically affect the properties of the macromolecular complex that is formed. In addition, we cannot rule out the possibility that the cross-linking specificity reflects hnRNP phasing by means other than an oligo(U)-specific interaction by hnRNP proteins. Finally, the contribution of the UV hypersensitivity of the pyrimidine ring of uridine to our observations remains to be elucidated.

The sequence or structural context of the uridylate tract clearly influences its ability to serve as a C protein cross-linking or hnRNP complex alignment element. The SVE poly(A) signal, for example, contains a seven-base U tract that lies adjacent to the AAUAAA motif (−7 to −13 relative to the cleavage site). This U tract, however, does not mediate efficient C protein cross-linking to the SVE RNA (SVE-GEM; Fig. 6B), possibly because of its proximity to the AAUAAA motif, which serves to direct the binding of an independent complex (71, 72) that could block interaction with hnRNP proteins. Alternatively, there may be a sequence requirement in the nucleotides surrounding the UUUUU tract, an upper limit in the number of consecutive U residues allowed for a functional element, or other structural influences yet to be determined.

It is clear that sequences in addition to UUUUU can mediate the cross-linkable alignment of C proteins. The SVE downstream domain does not contain a UUUUU tract (Fig. 6A), but it is cross-linked to C proteins (Fig. 6B). The adenovirus L3 and E1B poly(A) addition signals also lack the UUUUU element within their downstream domains, but these RNAs are also efficiently cross-linked to C proteins (70). The ability of the UUUUU element to functionally substitute for the SVE (Fig. 6C) and L3 (data not shown) elements, however, suggests a common mechanism for mediation of processing efficiency by downstream elements. Perhaps C proteins can recognize several different sequence elements. It is also possible, since the hnRNP particle consists of numerous independent RNA-binding proteins, that there are multiple mechanisms which rely on different constituents of the hnRNP complex for aligning or phasing it on the downstream domain of the polyadenylation signal in a sequence-directed manner. This proposed flexibility in specific hnRNP particle-RNA interactions could explain the lack of a good consensus sequence for the downstream element as well as its apparently reiterated nature (35, 73).

Several lines of evidence indicate that the hnRNP C protein interactions we have observed are occurring on functional polyadenylation substrate RNA and are not a "dead-end" interaction. First, the SVL polyadenylation substrate RNA, which is very efficiently (>80%) processed in our in vitro system, was nearly quantitatively coimmunoprecipitated with use of hnRNP-specific monoclonal antibodies (Fig. 2). Second, cross-linked C proteins can be detected in the 50S polyadenylation signal recognition complex identified on native gels (43). Third, antibodies to the hnRNP C proteins reduce in vitro polyadenylation efficiency when used in depletion experiments (data not shown). Fourth, antibodies to the hnRNP C proteins will coimmunoprecipitate the AAUAAA-specific 64-kDa protein (71) if samples are not treated with RNase prior to addition of antibody (data not shown). This finding indicates that the hnRNP C proteins are bound to the same molecule as is the 64-kDa protein, which has been shown to be involved in polyadenylation specificity or stimulation (72). Finally, while C proteins are not necessary for polyadenylation in a purified reconstituted system (72), it should be noted that this fractionated system also does not respond to downstream element influences (52) as is the case with in vivo studies or studies using crude in vitro systems. Gilmartin and Nevins have recently reported a fractionated poly(A) system that does respond to downstream element influences (24). Establishing the relationship of C proteins to processing in that system would be informative. In conclusion, these observations, along with the other data presented above and in our previous work, argue strongly for a role for hnRNP proteins in polyadenylation efficiency. The technically challenging genetic and biochemical experiments to definitively address the role of the hnRNP complex in polyadenylation are now being initiated.

To provide additional support for our proposed role of hnRNP complexes in polyadenylation, we attempted to assay the effect of the addition of poly(U) on the in vitro reaction. The results of these studies were inconclusive because all RNAs, including the lambda-derived control RNA (Fig. 2, con), inhibited the polyadenylation reaction when added at concentrations high enough to abolish the hnRNP C protein (data not shown). This result probably reflects the fact that hnRNP complexes, which contain C proteins, will form on every RNA added to the in vitro system (Fig. 2) in a sequence-independent manner, as well as the many specific and nonspecific associations of hnRNP proteins within the nuclear extract.

Previous work indicated that a deletion of the +30 to +55 segment of the SVL poly(A) signal resulted in a fivefold decrease in polyadenylation efficiency. Interruption of the +50 to +55 U tract in this region, however, did not result in a decrease in processing efficiency (70). This finding is consistent with previous observations that the downstream element of the SVL polyadenylation signal is reiterated (52, 73). We have recently identified a 50-kDa protein that requires a G-rich region within the +30 to +55 segment of the downstream region for efficient cross-linking (and binding) to the SVL RNA (Fig. 5B; unpublished observations). This abundant nuclear protein may represent one of the auxiliary polypeptides of the hnRNP monoparticle (47), and it may provide an alternative or additional mechanism for the sequence-mediated alignment or phasing of hnRNP complexes with RNA. We are now evaluating the relationship of this protein to the hnRNP particle as well as to the efficiency of polyadenylation in vitro. To rule out the possibility that SVL-4U5 and SVE-4U5 RNAs contained a fortuitous binding site for this 50-kDa protein (despite the fact that it was not detected in cross-linking experiments), we removed the three G residues located four bases upstream of the UUUUU element (Fig. 5 and 6). Removal of this G-rich stretch had no effect on the ability of the UUUUU tract to functionally substitute for the SVL downstream element (data not shown).

If hnRNP complexes are involved in efficient polyadenylation, they may act through several mechanisms. They may hold the RNA substrate in an optimal configuration for enzymatic processing. Alternatively, hnRNP complexes may melt out structural conformations that are refractory to AAUAAA-specific complex assembly. The association of helix-destabilizing activity with hnRNP proteins is consis-

tent with this suggestion (12, 30). A third possibility is that the complexes attract specific factors that mediate processing. The peripheral location of the C proteins in the hnRNP particle and their resulting accessibility would support this model (32). Finally, the hnRNP associations could prevent nonspecific RNA-binding proteins from interacting with the substrate RNA and interfering with polyadenylation. This proposal fits well with the observations of Ryner et al. (52) that the SVL downstream element is not required for optimal polyadenylation efficiency in cell-free reactions using partially purified components. It is unlikely that the hnRNP proteins play an enzymatic role in polyadenylation, since the hnRNP C proteins do not copurify with activities required to reconstitute in vitro polyadenylation (72).

The mix of sequence-mediated and nonspecific interactions that we observed between hnRNP proteins and RNAs may explain earlier reports of apparently random versus nonrandom distributions of monoparticles on different RNA molecules (2, 7, 44, 49, 64, 68). The RNA associations of hnRNP particles may depend on RNA sequence, RNA structure, or protein-protein interactions within the particle itself. Possibly, the ability of hnRNP particles to orient themselves in nonspecific (perhaps through the A proteins) and specific (perhaps through the C proteins) fashions with respect to RNA sequence allows flexibility in the function of these particles. Some particles may, for example, bind RNA nonspecifically and participate in general hnRNA packaging (15, 49, 64, 68). Other complexes, such as those aligned by the UUUUU sequence described here, may be positioned to play a more active role in RNA processing efficiency or selectivity. Similar multifunctional roles in viral RNA packaging and metabolism have been suggested for the vesicular stomatitis virus nucleocapsid protein (reviewed in reference 3).

Finally, hnRNP proteins have also been implicated in in vitro splicing reactions by antibody depletion-inhibition experiments (11, 55) and by the specific immunoprecipitation of nuclease-resistant RNA segments (61). The RNA segments protected by the hnRNP proteins in these experiments involved the uridylate-rich polypyrimidine tracts of splicing substrates. Perhaps the hnRNP complexes are the "default" binding activity for this sequence element, in a manner similar to that for the downstream element of the polyadenylation signal. Specific factors such as U2AF (51), Sex Lethal (57), or U5-associated proteins (22, 63) may compete with the hnRNP proteins for these sequence elements and act as specific activators or repressors of RNA splicing. The sequence-mediated alignment or phasing of hnRNP complexes on pre-mRNA may also influence other aspects of nuclear mRNA metabolism, including stability and transport to the cytoplasm.

ACKNOWLEDGMENTS

We thank B. Merrill and K. Williams for a generous gift of recombinant A1 protein, R. Dickinson and M. Flocco for synthesis of oligonucleotides, and M. Marlow for expert technical assistance in the preparation of monoclonal antibodies.

This work was supported by Public Health Service grant CA-38965 to T.S. and biomedical research support grant 2 S07 RR05393 to J.W., both from the National Institutes of Health. J.W. received postdoctoral support from the American Cancer Society (grant PF2706) and Howard Hughes Medical Institute, and T.S. is an American Cancer Society Professor.

LITERATURE CITED

1. **Adam, S. A., T. Nakagawa, M. S. Swanson, and G. Dreyfuss.** 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. Mol. Cell. Biol. **6**:2932–2943.
2. **Augenlicht, L. H.** 1979. Protected nucleotide sequences in nuclear ribonucleoprotein. Biochemistry **18**:3780–3786.
3. **Banerjee, A. K.** 1987. Transcription and replication of rhabdoviruses. Microbiol. Rev. **51**:66–87.
4. **Barnett, S. F., D. L. Friedman, and W. M. LeStourgeon.** 1989. The C proteins of HeLa 40S nuclear ribonucleoprotein particles exist as anisotropic tetramers of (C1)3 C2. Mol. Cell. Biol. **9**:492–498.
5. **Beyer, A. L., A. H. Bouton, and O. L. Miller.** 1981. Correlation of hnRNP structure and nascent transcript cleavage. Cell **26**: 155–165.
6. **Beyer, A. L., M. E. Christensen, B. W. Walker, and W. M. LeStourgeon.** 1977. Identification and characterization of the packaging proteins of core 40S hnRNP particles. Cell **11**:127–138.
7. **Beyer, A. L., O. L. Miller, and S. L. McKnight.** 1980. Ribonucleoprotein structure in nascent hnRNA is nonrandom and sequence-dependent. Cell **20**:75–84.
8. **Billings, P. B., and T. E. Martin.** 1978. Proteins of nuclear ribonucleoprotein subcomplexes. Methods Cell Biol. **17**:349–376.
9. **Burd, C. G., M. S. Swanson, M. Gorlach, and G. Dreyfuss.** 1989. Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. Proc. Natl. Acad. Sci. USA **86**:9788–9792.
10. **Choi, Y. D., and G. Dreyfuss.** 1984. Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. Proc. Natl. Acad. Sci. USA **81**:7471–7475.
11. **Choi, Y. D., P. J. Grabowski, P. A. Sharp, and G. Dreyfuss.** 1986. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. Science **231**:1534–1539.
12. **Cobianchi, F., R. L. Karpel, K. R. Williams, V. Notario, and S. H. Wilson.** 1988. Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1. J. Biol. Chem. **263**:1063–1071.
13. **Cobianchi, F., D. N. SenGupta, B. Z. Zmudzka, and S. H. Wilson.** 1986. Structure of rodent helix-destabilizing protein revealed by cDNA cloning. J. Biol. Chem. **261**:3536–3543.
14. **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. Mol. Cell. Biol. **5**:2104–2113.
15. **Conway, G., J. Wooley, T. Bibring, and W. M. LeStourgeon.** 1988. Ribonucleoproteins package 700 nucleotides of pre-mRNA into a repeating array of regular particles. Mol. Cell. Biol. **8**:2884–2895.
16. **Conway, L., and M. P. Wickens.** 1985. A sequence downstream of AAUAAA is required for formation of simian virus 40 late mRNA 3′ termini in frog oocytes. Proc. Natl. Acad. Sci. USA **82**:3949–3953.
17. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
18. **Dreyfuss, G.** 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Annu. Rev. Cell Biol. **2**:459–498.
19. **Dreyfuss, G., Y. D. Choi, and S. A. Adam.** 1984. Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. Mol. Cell. Biol. **4**:1104–1114.
20. **Dreyfuss, G., M. S. Swanson, and S. Pinol-Roma.** 1988. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trends Biochem. Sci. **13**:86–91.
21. **Economides, I. V., and T. Pederson.** 1983. Structure of nuclear ribonucleoproteins: heterogeneous nuclear RNA is complexed with a major sextet of proteins in vivo. Proc. Natl. Acad. Sci. USA **80**:1599–1602.
22. **Gerke, V., and J. A. Steitz.** 1986. A protein associated with small nuclear ribonucleoprotein particles recognizes the 3′ splice site of pre-messenger RNA. Cell **47**:973–984.

23. **Gil, A., and N. J. Proudfoot.** 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. Cell **49**:399–406.
24. **Gilmartin, G. M., and J. R. Nevins.** 1989. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. Genes Dev. **3**:2180–2189.
25. **Harlow, E., and D. Lane.** 1988. Antibodies: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Hart, R. P., M. A. McDevitt, and J. R. Nevins.** 1985. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and early simian virus 40 polyadenylation sites. Mol. Cell. Biol. **5**:2975–2983.
27. **Hart, R. P., M. A. McDevitt, and J. R. Nevins.** 1985. Poly(A) site cleavage in HeLa nuclear extract is dependent on downstream sequences. Cell **43**:677–683.
28. **Hawkes, R., E. Niday, and J. Gordon.** 1982. A dot-immunobinding assay for monoclonal and other antibodies. Anal. Biochem. **119**:142–147.
29. **Holcomb, E. R., and D. L. Friedman.** 1984. Phosphorylation of the C-proteins of HeLa cell hnRNP particles. J. Biol. Chem. **259**:31–40.
30. **Kumar, A., K. R. Williams, and W. Szer.** 1986. Purification and domain structure of core hnRNP proteins A1 and A2 and their relationship to single-stranded DNA-binding proteins. J. Biol. Chem. **261**:11266–11273.
31. **LeStourgeon, W. M., L. Lothstein, B. W. Walker, and A. L. Beyer.** 1981. The composition and general topology of RNA and protein in monomer 40S ribonucleoprotein particles, p. 49–87. *In* H. Busch (ed.), The cell nucleus, vol. 9. Academic Press, Inc., New York.
32. **Lothstein, L., H. P. Arenstorf, S. Y. Chung, B. W. Walker, J. C. Wooley, and W. M. LeStourgeon.** 1985. General organization of protein in HeLa 40S nuclear ribonucleoprotein particles. J. Cell Biol. **100**:1570–1581.
33. **Malcolm, D. B., and J. Sommerville.** 1977. The structure of nuclear ribonucleoprotein of amphibian oocytes. J. Cell Sci. **24**:143–165.
34. **Mason, P. J., J. A. Elkington, M. M. Lloyd, M. B. Jones, and J. G. Williams.** 1986. Mutations downstream of the polyadenylation site of a xenopus beta-globin mRNA affect the position but not the efficiency of 3' processing. Cell **46**:263–270.
35. **McDevitt, M. A., R. P. Hart, W. W. Wong, and J. R. Nevins.** 1986. Sequences capable of restoring poly(A) site function define two distinct downstream elements. EMBO J. **5**:2907–2913.
36. **McDevitt, M. A., 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.
37. **McKnight, S. L., and O. L. Miller.** 1976. Ultrastructural patterns of RNA synthesis during early embryogenesis of drosophila melanogaster. Cell **8**:305–319.
38. **McLauchlin, J., D. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from the AAUAAA signal is required for efficient formation of mRNA 3' termini. Nucleic Acids Res. **13**:1347–1367.
39. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
40. **Merrill, B. M., S. F. Barnett, W. M. LeStourgeon, and K. R. Williams.** 1989. Primary structure differences between proteins C1 and C2 of HeLa 40S nuclear ribonucleoprotein particles. Nucleic Acids Res. **17**:8441–8449.
41. **Merrill, B. M., M. B. LoPresti, K. L. Stone, and K. R. Williams.** 1986. High pressure liquid chromatography purification of UP1 and UP2, two related single-stranded nucleic acid-binding proteins from calf thymus. J. Biol. Chem. **261**:878–883.
42. **Moore, C. L., J. Chen, and J. Whoriskey.** 1988. Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding. EMBO J. **7**:3159–3169.
43. **Moore, C. L., and P. A. Sharp.** 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell **41**:845–855.
44. **Ohlsson, R. I., C. van Eekelen, and L. Phillipson.** 1982. Nonrandom localization of ribonucleoprotein (RNP) structures within an adenovirus mRNA precursor. Nucleic Acids Res. **10**:3053–3068.
45. **Osheim, Y. N., O. L. Miller, and A. L. Beyer.** 1985. RNP particles at splice site junction sequences on drosophila transcripts. Cell **43**:143–151.
46. **Patton, J. R., and C. B. Chae.** 1985. Specific regions of the intervening sequences of beta-globin RNA are resistant to nuclease in 50S heterogeneous nuclear RNA-protein complexes. Proc. Natl. Acad. Sci. USA **82**:8414–8418.
47. **Pinol-Roma, S., Y. D. Choi, M. J. Matunis, and G. Dreyfuss.** 1988. Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. Genes Dev. **2**:215–227.
48. **Preugschat, F., and B. Wold.** 1988. Isolation and characterization of a xenopus laevis C protein cDNA: structure and expression of a heterogeneous nuclear ribonucleoprotein core protein. Proc. Natl. Acad. Sci. USA **85**:9669–9673.
49. **Pullman, J. M., and T. E. Martin.** 1983. Reconstitution of nucleoprotein complexes with mammalian heterogeneous nuclear ribonucleoprotein (hnRNP) core proteins. J. Cell Biol. **97**:99–111.
50. **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.
51. **Ruskin, B., P. D. Zamore, and M. R. Green.** 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell **52**:207–219.
52. **Ryner, L. C., Y. Takagaki, and J. L. Manley.** 1989. Sequences downstream of AAUAAA signals affect pre-RNA cleavage and polyadenylation in vitro both directly and indirectly. Mol. Cell. Biol. **9**:1759–1771.
53. **Sadofsky, 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.
54. **Sadofsky, 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.
55. **Sierkowska, H., W. Szer, P. J. Furdon, and R. Kole.** 1986. Antibodies to hnRNP core proteins inhibit in vitro splicing of human beta-globin pre-mRNA. Nucleic Acids Res. **14**:5241–5254.
56. **Smith, K. C.** 1976. The radiation-induced addition of proteins and other molecules to nucleic acids, p. 187–218. *In* S. Y. Wang (ed.), Photochemistry and photobiology of nucleic acids, vol. 2. Academic Press, Inc., New York.
57. **Sosnowski, B. A., J. M. Belote, and M. McKeown.** 1989. Sex specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. Cell **58**:449–459.
58. **Steitz, J. A., and R. Kamen.** 1981. Arrangement of 30S heterogeneous nuclear ribonucleoprotein on polyoma virus late nuclear transcripts. Mol. Cell. Biol. **1**:21–34.
59. **Stevenin, J., R. Gattoni, G. Divilliers, and M. Jacob.** 1979. Mild nuclease treatment as a probe for a non-random distribution of adenovirus-specific RNA sequences and of cellular RNA in nuclear ribonucleoprotein fibrils. J. Mol. Biol. **155**:185–205.
60. **Swanson, M. S., and G. Dreyfuss.** 1988. Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. Mol. Cell. Biol. **8**:2237–2241.
61. **Swanson, M. S., and G. Dreyfuss.** 1988. RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns. EMBO J. **7**:3519–3529.
62. **Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss.** 1987. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. Mol. Cell. Biol. **7**:1731–1739.

63. **Tazi, J., C. Alibert, J. Temsamani, I. Reveillaud, G. Cathala, C. Brunei, and P. Jeanteur.** 1986. A protein that specifically recognizes the 3′ splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein. Cell **47**:755–766.
64. **Thomas, J. O., S. K. Glowacka, and W. Szer.** 1983. Structure of complexes between a major protein of heterogeneous nuclear ribonucleoprotein particles and polyribonucleotides. J. Mol. Biol. **171**:439–455.
65. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
66. **Tsanev, R. G., and L. P. Djondjurov.** 1982. Ultrastructure of free ribonucleotide complexes in spread mammalian nuclei. J. Cell Biol. **94**:662–666.
67. **Van Eekelen, C., R. Ohlsson, L. Philipson, E. Mariman, R. Van Beek, and W. van Venrooij.** 1982. Sequence dependent interaction of hnRNP proteins with late adenoviral transcripts. Nucleic Acids Res. **10**:7115–7131.
68. **Wilk, H. E., G. Angeli, and K. P. Schaefer.** 1983. In vitro reconstitution of 35S ribonucleoprotein complexes. Biochemistry **22**:4592–4600.
69. **Wilk, H. E., H. Werr, D. Friedrich, H. H. Kiltz, and K. P. Schaefer.** 1985. The core proteins of 35S hnRNP complexes: characterization of nine different species. Eur. J. Biochem. **146**:71–81.
70. **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.
71. **Wilusz, J., and T. Shenk.** 1988. A 64 kd nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. Cell **52**:221–228.
72. **Wilusz, J., T. Shenk, Y. Takagaki, and J. L. Manley.** 1990. A multicomponent complex is required for the AAUAAA-dependent cross-linking of a 64-kilodalton protein to polyadenylation substrates. Mol. Cell. Biol. **10**:1244–1248.
73. **Zarkower, D., and M. Wickens.** 1988. A functionally redundant downstream sequence in SV40 late pre-mRNA is required for mRNA 3′-end formation and for assembly of a precleavage complex in vitro. J. Biol. Chem. **263**:5780–5788.