MOLECULAR AND CELLULAR BIOLOGY, Nov. 1989, p. 4872–4881 Vol. 9, No. 11
0270-7306/89/114872-10$02.00/0
Copyright © 1989, American Society for Microbiology

A Specific 31-Nucleotide Domain of U1 RNA Directly Interacts with the 70K Small Nuclear Ribonucleoprotein Component

CHARLES C. QUERY, REX C. BENTLEY,† AND JACK D. KEENE*

Department of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710

Received 28 February 1989/Accepted 26 July 1989

We have defined the nucleotide sequence of a protein-binding domain within U1 RNA that specifically recognizes and binds both to a U1 small nuclear ribonucleoprotein component (the 70K protein) and to the previously defined RNA-binding domain of the 70K protein. We have investigated direct interactions between purified U1 RNA and 70K protein by reconstitution in vitro. Thirty-one nucleotides of U1 RNA, corresponding to stem-loop I, were required for this interaction. Nucleotides at the 5′ end of U1 RNA that are involved in base pairing with the 5′ splice site of pre-mRNA were not required for binding. In contrast to other reports, these findings demonstrate that a specific domain of U1 RNA can bind directly to the 70K protein independently of any other snRNP-associated proteins.

The 52-kilodalton U1 RNA-associated protein (the 70K protein) is a component of the U1 small nuclear ribonucleoprotein (snRNP) complex. Several snRNPs have been implicated in the removal of intervening sequences during precursor mRNA splicing (see references 11, 25, and 41 for reviews). The U1 snRNP, in particular, has been shown to interact with the 5′ splice site, at least in part by base pairing of this site with the 5′-terminal 10 nucleotides (nt) of U1 RNA (1, 5, 18, 29, 45). The human U1 snRNP is composed of 165 nt of U1 RNA and two classes of proteins: U1-specific proteins (70K, A, and C) and the Sm complex (consisting of six U snRNP-common proteins) (9, 22). Although these proteins probably play important structural roles in snRNPs and at least some are required for the interaction of U1 snRNP with 5′ splice sites (29), none have been assigned specific functions.

Previous studies of the RNA-binding properties of the U1 snRNP-specific proteins have involved indirect methods including immunoprecipitation of nuclease-treated cell extracts (34) and microinjection of RNA into *Xenopus* oocytes (12). These studies suggested that the 70K protein, as well as the A protein and perhaps the C protein, associates with the 5′ half of U1 RNA, in particular, stem-loop I. In addition, weak requirements were implicated for regions of stem-loops II, III, and IV (12, 34). More recently, Hamm et al., using a protein sequestration assay to study U1 RNA-protein associations, argued that only the C protein directly contacts stem-loop I U1 RNA (13). These authors also suggested that the 70K and A proteins associate with U1 RNA through interactions with the C protein. Although such in vivo studies have the advantage of examining the snRNP in the presence of all potential structural proteins, the resulting complexity makes it difficult to elucidate the specific RNA-protein interactions within the snRNP. None of the methodologies used in previous studies provided direct evidence of whether U1 RNA directly contacts the 70K, A, or C protein.

We have attempted to reduce the complexity of the analysis of the U1 snRNP structure by determining which regions of U1 RNA and of the 70K protein are necessary to form a specific association. By in vitro reconstitution of purified human U1 RNA and 70K protein, we have investigated the direct interaction of these two components of the U1 snRNP. Using various binding methods, we maintained the specificity of the U1 RNA-70K reconstitution through progressive deletion of both components. We previously defined the RNA-binding domain of the 70K protein required for this interaction (35). In this study, we report that both the 70K protein and the RNA-binding domain alone are directly recognized by 31 nt of U1 RNA that make up the 5′-most stem-loop structure. The A, C, and Sm proteins were not required for this recognition and binding. Further deletion of the 31-nt structure resulted in reduced efficiency of reconstitution, thus defining the minimal RNA domain needed for recognition of the 70K protein.

MATERIALS AND METHODS

Enzymes, antisera, host strains, and vectors. Enzymes were purchased from Bethesda Research Laboratories, Inc.; United States Biochemical Corp.; and New England BioLabs, Inc.

Autoimmune antisera were obtained from the Duke University Medical Center Fluorescent Antinuclear Antibody Laboratory. Anti-β-galactosidase monoclonal antibody was purchased from Promega Biotec.

Escherichia coli Y1089, Y1090, and JM83 were obtained from the American Type Culture Collection, Rockville, Md. Strain NM522 and the pGEM-3Zf(+) single-stranded system were purchased from Promega Biotec.

DNA sequencing. Mutated U1 DNAs were sequenced by dideoxynucleotide chain termination (40) with a modified T7 DNA polymerase (44) from the Sequenase system (U.S. Biochemical Corp.). Oligonucleotides complementary to the SP6 and T7 promoters and specific for U1 sequences were synthesized on an Applied Biosystems model 380A DNA synthesizer and purified by high-pressure liquid chromatography.

In vitro transcription. U1, β-globin, hY1, and pGEM3 RNAs were transcribed with SP6 (Promega Biotec) or T7 (U.S. Biochemical Corp.) RNA polymerase as described previously (27), in the presence of 50 μCi of [α-^{32}P]UTP (ICN Pharmaceuticals Inc.). The β-globin (HβΔ6) template was a gift from Robin Reed and Tom Maniatis, Harvard University, and was linearized with *Bam*HI prior to transcription. Immunoglobulin (p381) template was obtained

* Corresponding author.
† Present address: Department of Pathology, Duke University Medical Center, Durham, NC 27710.

from Yair Argon, Duke University, and was linearized with *Hin*dIII. hY1 template was prepared by Susan Deutscher, Duke University, and produces the exact 112-nt RNA transcript after linearization with *Dra*I. pGEM3 RNA was made from pGEM-3Zf(+) linearized with *Pvu*II and transcribed with T7 RNA polymerase, producing a transcript of 269 nt.

Oligonucleotide-directed mutagenesis. Mutagenesis was performed as specified by the supplier (Amersham Corp.), with pGEM-3Zf(+) vectors.

Purification of recombinant fusion protein. The 70K-LacZ fusion protein (35), the 35-216–LacZ fusion protein (36), and the fusion protein isolation procedure (36) were described previously. All fusion proteins required 4 M urea for solubilization.

U1 RNA constructs. A genomic clone of human U1 DNA was a gift from Nouria Hernandez, Cold Spring Harbor Laboratory. A *Bam*HI-*Hin*dIII fragment was cloned between the pSP64 *Eco*RI-*Hin*dIII sites next to the SP6 promoter. 3′ deletions were produced by cleavage of the template at various restriction sites prior to transcription. 5′ deletion mutants were produced by using *Sau*3AI-*Hin*dIII, *Mnl*I-*Hin*dIII, *Fok*I-*Hin*dIII, and *Taq*I-*Hin*dIII fragments, which were subcloned into a blunted pSP64 *Eco*RI site. These constructs represented nucleotides 28 to 164, 48 to 164, 93 to 164, and 117 to 164 of U1 RNA, respectively. It should be noted that these deletions, excepting those involving the *Sau*3AI site, were selected to remove precisely one or more predicted RNA stem-loops (2, 30) to maintain the best possible predicted secondary and tertiary structure of the remaining molecule. Predicted secondary structure was analyzed by using the University of Wisconsin Genetics Computer Group Sequence Analysis Software (7), the FOLD program (49), and the free-energy parameters of Freier et al. (10). The 5′ deletion mutants were truncated by using *Rsa*I; the resulting U1 RNA fragments contained an additional 28 nt of genomic U1 at the 3′ end, accounting for the slower-than-expected mobility of some species (see Fig. 2, lanes 8 to 11). Identical binding results were obtained with the full-length U1 RNA with and without these extra sequences.

To make a U1 transcript without additional 5′ or 3′ sequences, we inserted the complete U1 *Bam*HI-*Hin*dIII sequence into the pGEM-3Zf(+) *Eco*RI site, next to the T7 promoter, and used oligonucleotide-directed mutagenesis to remove all bases, except one G, between the promoter start site and the first base of the U1 RNA coding sequence. To make the exact 3′ end of U1 RNA, we inserted an *Apa*I site next to the exact 3′ end of the U1 RNA coding sequence. This construct, NEU1, was digested with *Apa*I, blunted with Klenow fragment of DNA polymerase (U.S. Biochemical Corp.), and transcribed with T7 DNA polymerase to produce U1 RNA authentic in sequence, except for the trimethylated 5′ cap and other modified bases.

Internal deletions of NEU1 were made by oligonucleotide-directed mutagenesis. In brief, these replaced bases 28 to 38 with UUG (U1ΔL1), replaced bases 16 to 48 with UU (U1ΔSL1), and removed base 22 (U1ΔU_{22}). The template for U1SL1 (nt 16 to 48) was created by deleting bases 1 to 15 and changing base 49 to a G, thus creating a *Sma*I restriction site; digestion of the template with *Sma*I and subsequent transcription produces an RNA consisting of U1 nt 16 to 48. All RNA transcripts used for binding were transcribed in the presence of [α-^{32}P]UTP as described previously (27), purified on 5% acrylamide–8.3 M urea gels, phenol-chloroform extracted, and concentrated by ethanol precipitation.

RNA transcripts synthesized by using T7 polymerase and single-stranded template have been reported previously to contain an additional, non-template-encoded base at the 3′ end (28). Therefore, terminal RNA sequences were determined by using the two-dimensional wandering-spot method and end-group analysis as described previously (17). The U1 RNA transcript containing nt 1 to 48 was synthesized from template digested with *Mnl*I (New England BioLabs), which should provide U1 DNA base pairs 1 to 47 as the template. The 3′ end of this transcript contained up to two additional non-template-encoded bases. Transcripts ending in nucleotides 47, 48, and 49 were in the ratio 45:45:10. Nucleotide 47 was 100% C (which corresponds to nt 47 of U1 RNA); nt 48 was 50% C, 26% A, 12% G, and 12% U; and nt 49 was 41% C, 31% A, 14% G, 14% U. Thus, this transcript (see Fig. 2, lanes 4) contained heterogeneous 3′ ends, of which 27.5% corresponded to nt 1 to 48 of authentic U1 RNA. These ratios, however, probably depend upon the exact conditions of transcription and may vary significantly from one transcription of this template to another.

Alkali-digested RNA ladders. U1 RNA was transcribed by using T7 RNA polymerase as described above, dephosphorylated with calf intestinal alkaline phosphatase (Boehringer Mannheim Biochemicals) (24), and ^{32}P labeled at the 5′ end with polynucleotide kinase (Bethesda Research Laboratories) and [γ-^{32}P]ATP. RNA was 3′ end labeled by the terminal addition of [^{32}P]pCp by using T4 RNA ligase (Pharmacia Fine Chemicals) as described previously (17). Labeled RNAs were purified on a 70-cm 8.3 M urea–10% polyacrylamide gel prior to degradation because of heterogeneity of bases at the 3′ end. Partial alkali ladders were made by incubation of RNA in 33 mM sodium bicarbonate (pH 9.0) at 90°C for 10 min. RNA was partially digested with RNase T1 in 20 mM sodium citrate (pH 5.0)–1 mM EDTA–7 M urea at 50°C for 15 min. Alkali ladders were incubated with 70K-LacZ or 35-216–LacZ fusion proteins and immunoprecipitated with anti-β-galactosidase antibody as described below, and the coprecipitated RNA was analyzed by 10% polyacrylamide–8.3 M urea gel electrophoresis and autoradiography.

RNA binding. 70K-LacZ and control fusion proteins were incubated with RNA in a mixture of NET-2 (100 mM NaCl, 50 mM Tris [pH 7.4], 0.05% Nonidet P-40), 4 M urea, 250 U of RNasin (Promega Biotec) per ml, 0.6 mM vanadium ribonucleoside complex, and 1 mg of *E. coli* tRNA (Sigma Chemical Co.) per ml at 4°C for 10 min. These assays required 4 M urea to maintain the solubility of the LacZ fusion proteins. Immunoprecipitations were performed essentially as previously described (21, 22). 70K-LacZ fusion proteins were immunoprecipitated with anti-β-galactosidase monoclonal antibody in a mixture of NET-2, 4 M urea, 0.4 mM vanadium ribonucleoside complex, and Pansorbin (Calbiochem-Behring). Before being added, the Pansorbin was washed twice with NET-1 (100 mM NaCl, 50 mM Tris [pH 7.4], 0.5% Nonidet P-40) and twice with the buffer used for immunoprecipitation and incubated with 200 μg of tRNA per ml for 10 min at 4°C. The relative efficiencies of U1 RNA transcript binding to 70K-LacZ fusion protein were determined by Cerenkov counting of the coprecipitated RNA. Coprecipitated RNA was analyzed by electrophoresis through a 5% polyacrylamide–8.3 M urea gel followed by autoradiography.

Mobility shift assays. 70K sequences were transcribed with T7 RNA polymerase (27), and the resulting RNA was translated in vitro (20) in rabbit reticulocyte lysates as described by the supplier (Promega Biotec). Translation products representing amino acids 92 to 216 (35), without

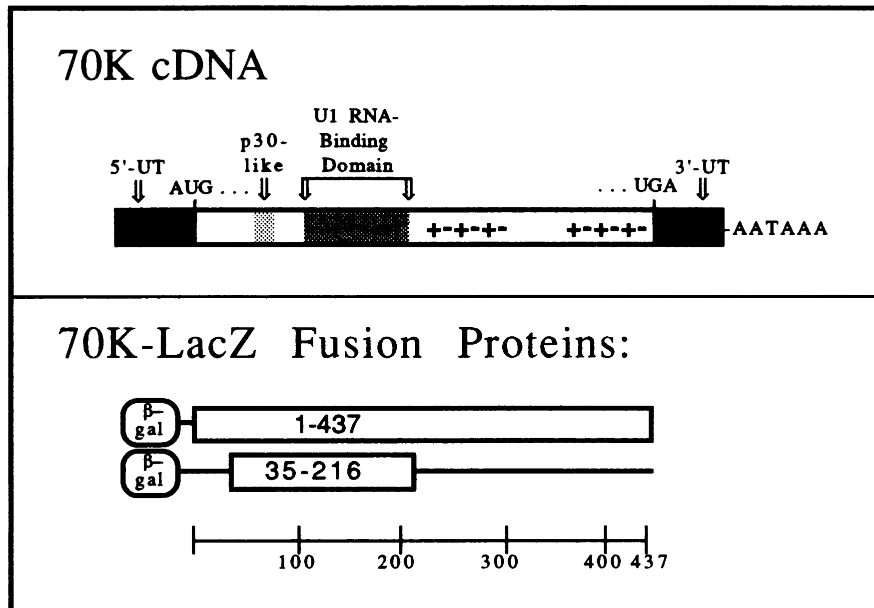

FIG. 1. Schematic of 70K cDNA and 70K-LacZ fusion proteins used for analyses of the U1 RNA binding site. (Top) Human 70K cDNA, indicating three regions of the protein: a region of sequence similarity to $p30^{gag}$ protein of type C retroviruses (36), the U1 RNA-binding domain, and a highly charged (+ and −) carboxyl half of the protein. (Bottom) Two LacZ fusion proteins used for RNA-binding studies. The first contains all 437 amino acids of the 70K protein. The second contains amino acids 35 to 216, encompassing the U1 RNA-binding domain (35). The scale is in amino acid residues.

further purification from the in vitro translation reaction, were incubated with RNA in a mixture of KNET buffer [20 mM KCl, 80 mM NaCl, 2 mM ethylene glycol-bis(β-aminoethyl ether)-*N,N,N,N'*,-tetraacetic acid (EGTA), 50 mM Tris (pH 7.4), 0.05% Nonidet P-40], 1 mM $MgCl_2$, 2.5% polyvinyl alcohol, 1 mM dithiothreitol, 100 U of RNasin per ml, 5 μg of poly(A) RNA (Sigma) per ml, 0.4 mg of *E. coli* tRNA per ml, and 0.4 mM vanadium ribonucleoside complex at 37°C for 20 min to achieve equilibration. In vitro-transcribed RNAs were included at a final concentration of 40 μg/ml. The products were added to one-quarter volume of loading buffer (KNET buffer containing 20% glycerol, 1 mg of xylene cyanol FF per ml, and 1 mg of bromophenol blue per ml) for immediate electrophoresis on a native 5.5% polyacrylamide–90 mM Tris borate (pH 8.3) gel as described previously (35). Native gels were rinsed in 50 mM Tris (pH 10.0) for 10 min to deacylate [^{35}S]methionyl tRNA, soaked in 0.5 M sodium salicylate for 20 min, dried, and fluorographed.

RESULTS

U1 RNA-specific binding to the 70K protein. We have previously demonstrated that recombinant 70K protein could interact with in vivo-labeled U1 RNA (35). Both the 70K-LacZ fusion protein and 70K in vitro translation products selectively bound U1 RNA from total HeLa cell RNA. Thus, the recombinantly produced 70K-LacZ fusion protein and the in vitro-translated 70K protein interact specifically with U1 RNA. Using a series of deletion mutants of the 70K protein, we delimited the U1 RNA-binding domain to amino acids 92 to 216. For further analyses of this interaction, we used two LacZ fusion protein constructs, one containing the entire 70K sequence and one containing principally the U1 RNA-binding domain (Fig. 1).

Sequence-specific binding within the U1 RNA molecule. To investigate the regions of U1 RNA that are required for in vitro reconstitution with the 70K protein, we constructed a series of U1 RNA species by progressive deletion of the U1 DNA from the 3' end or the 5' end, or both (see Materials and Methods). Most of these deletions were selected to remove precisely one or more predicted RNA stem-loops (2, 30) to maintain the best possible predicted secondary and tertiary structure of the remaining molecule. Transcripts of these U1 RNA constructs (Fig. 2A) were produced in vitro and assayed for the ability to bind to the 70K-LacZ fusion protein. The fusion protein, solubilized in 4 M urea, was incubated with the RNA transcript and immunoprecipitated with anti-β-galactosidase monoclonal antibody, and the coprecipitated RNA was analyzed. As regions were deleted from the 3' end (Fig. 2B, lanes 1 to 6), nt 1 to 164, 1 to 117, 1 to 94, and 1 to 48 coprecipitated with the 70K-LacZ fusion protein (Fig. 2B, lanes 1 to 4); however, nt 1 to 32 and 1 to 17 did not bind (Fig. 2B, lanes 5 and 6). Thus, binding to the 70K-LacZ protein was observed until stem-loop I was interrupted (see Fig. 6). Similarly, deletions from the 5' end showed that there was no detectable binding to any transcripts lacking stem-loop I sequences (Fig. 2B, lanes 9 to 11). In addition, U1 RNA fragments representing double or single stem-loops, but lacking stem-loop I sequences (Fig. 2B, lanes 13, 15, and 16), did not bind. Transcripts were not coprecipitated when either the antibody or the fusion protein was omitted from the reaction (Fig. 2B, lanes C1 and C2 for full-length U1 RNA; data for other transcripts not shown). Other fusion proteins (La-LacZ, Ro-LacZ, and ribosomal proteins P0-LacZ, P1-LacZ and P2-LacZ) and β-galactosidase alone also were incubated with U1 RNA, and no coprecipitated RNA was detected (data not shown).

In many repetitions of the experiment, the transcript of nt 1 to 48 appeared to interact with approximately 25 to 50% of the efficiency of the intact U1 RNA (Fig. 2B, lane 4); however, as described below, the presence of nt 48 was critical for efficient interaction with the 70K protein. We

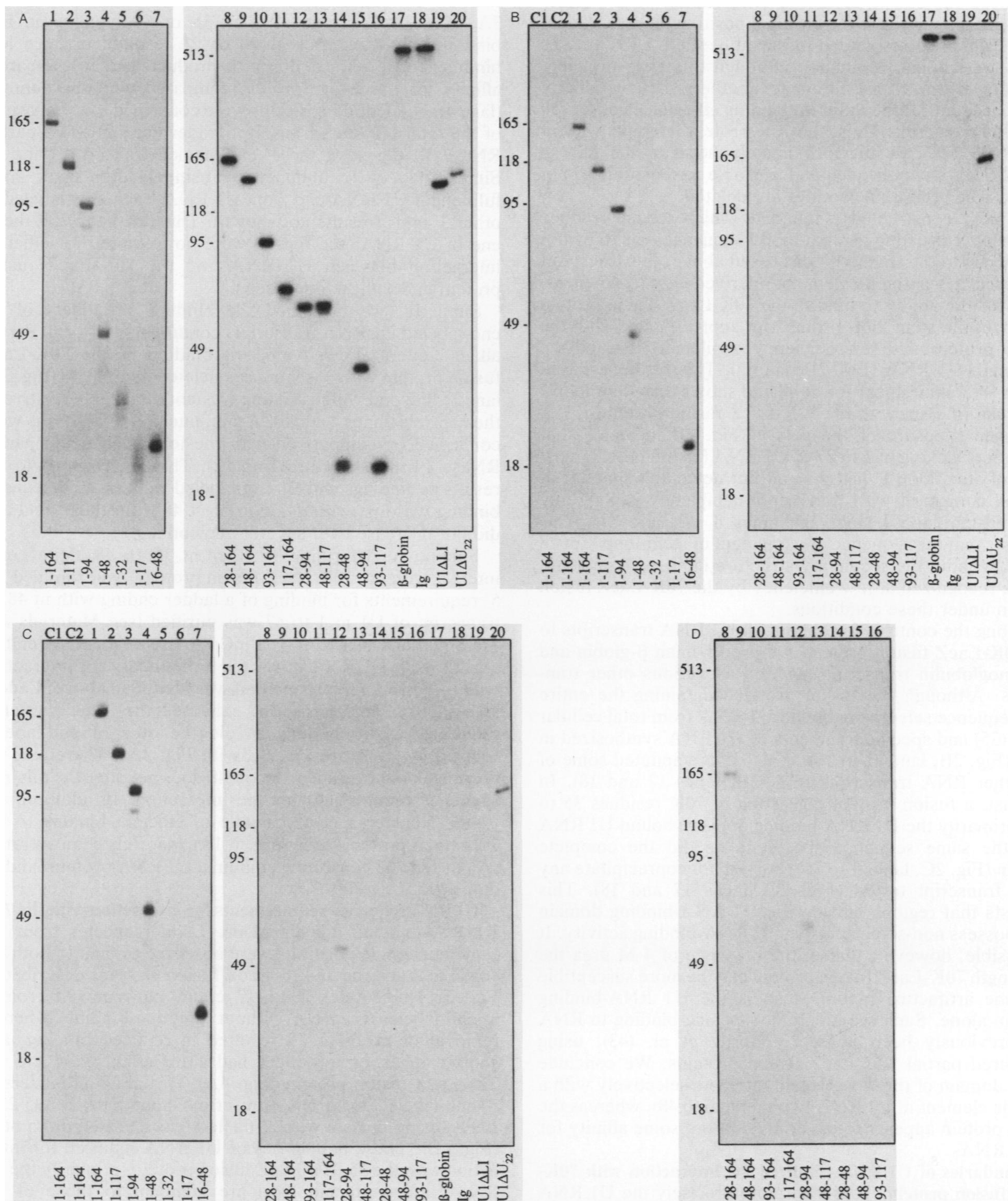

FIG. 2. Interaction of U1 RNA transcripts with 70K-LacZ fusion protein. (A) Various deleted transcripts of U1 RNA were synthesized in vitro and analyzed by electrophoresis on a denaturing 5% polyacrylamide–8.3 M urea gel. The nucleotides contained in each transcript are numbered from the 5′ end of U1 RNA (as shown in Fig. 6) and are shown below the lanes. Transcripts in lanes 8 to 11 migrate more slowly than might be expected, as described in Materials and Methods. (B) Gel-purified transcripts as shown in panel A were incubated with the full-length 70K-LacZ fusion protein and immunoprecipitated with monoclonal anti-β-galactosidase antibody. The coprecipitated RNA was analyzed as in panel A. In lane C1, antibody was omitted from the immunoprecipitation; in lane C2, fusion protein was omitted. (C) Transcripts as shown in panel A were incubated with a LacZ fusion protein containing amino acids 35 to 216 of the 70K protein (Fig. 1) and analyzed as in panel B. (D) A 10-fold-longer exposure of lanes 8 to 16 from panel B. Ig, Immunoglobulin.

sequenced the 3′ end of this in vitro transcript and found heterogeneous bases at the terminal position, approximately 28% of which corresponded to the authentic U1 RNA bases (see Materials and Methods), although this ratio probably varies from one transcription to another; such variability would account for the different binding efficiencies observed with this transcript. Thus, the 70K protein interacts with nt 1 to 48 of U1 RNA with approximately the same efficiency as with the entire molecule, and a stable association is not formed when these nucleotides are deleted.

We next constructed a template which would produce transcripts consisting of stem-loop I sequences (nt 16 to 48). In contrast to the transcript discussed above, this transcript contained the authentic template-derived nt 48. When assayed for the ability to bind to the 70K-LacZ fusion protein as above, the stem-loop I transcript coprecipitated with the fusion protein with an efficiency similar to that of the full-length U1 RNA (Fig. 2B, lane 7). This further demonstrates the requirement for nt 48 and shows that stem-loop I alone can interact with the 70K-LacZ fusion protein.

Longer exposure of the gels in Fig. 2B indicated that transcripts lacking nt 1 to 27 of U1 RNA (lacking the 5′-most 12 nt of stem-loop I) had a small but detectable amount of binding compared with transcripts completely lacking any part of stem-loop I (Fig. 2D, lanes 8, 12, and 14). This suggests that the sequence of a fragment of stem-loop I (nt 28 to 48), without the secondary structure imposed by a stem, can be recognized at low efficiency by the 70K-LacZ fusion protein under these conditions.

Among the controls for binding of U1 RNA transcripts to the 70K-LacZ fusion protein, we used human β-globin and immunoglobulin transcripts as well as various other transcripts. Although the fusion protein containing the entire 70K sequence selectively bound U1 RNA from total cellular RNA (35) and specific fragments of U1 RNA synthesized in vitro (Fig. 2B, lanes 1 to 16), it also coprecipitated some of the other RNA transcripts (Fig. 2B, lanes 17 and 18). In contrast, a fusion protein consisting of 70K residues 35 to 216 (primarily the U1 RNA-binding domain) bound U1 RNA with the same sequence specificity as did the complete protein (Fig. 2C, lanes 1 to 16), but did not coprecipitate any other transcript tested (Fig. 2C, lanes 17 and 18). This suggests that regions outside the U1 RNA-binding domain may possess non-sequence-specific RNA-binding activity. It is possible, however, that in the presence of 4 M urea the full-length 70K-LacZ fusion protein may be more susceptible to some artifactual binding than is the U1 RNA-binding domain alone. Such seemingly nonspecific binding to RNA has previously been noted by Spritz et al. (43), using denatured partial 70K-LacZ fusion proteins. We conclude that a domain of the 70K protein interacts selectively with a specific element in U1 RNA (within nt 16 to 48), whereas the intact protein appears to have, in addition, some affinity for other RNAs.

Boundaries of U1 RNA required for interaction with 70K-LacZ fusion protein. To define more precisely the U1 RNA sequences necessary for interaction with the 70K protein, we examined the binding of terminally labeled U1 RNA fragments to 70K-LacZ fusion proteins. A U1 RNA transcript consisting of nt 1 to 94 was 3′ end labeled, partially digested with alkali, incubated with 70K-LacZ fusion proteins, and immunoprecipitated with anti-β-galactosidase antibody. Such ladders of RNA fragments allow a comparison of the binding efficiencies of RNAs differing by a single nucleotide. Analysis of the coprecipitated RNA indicated approximately equal binding to each ladder species when up to 17 bases were progressively removed from the 5′ end (Fig. 3A, lane 3). Cleavage after nt 18 or 19 resulted in less binding; cleavage after nt 20 or 21 resulted in even less binding; and all shorter digestion products had little apparent affinity for the fusion protein compared with the controls (lane 4). Nucleotide positions were confirmed by comparison of the total U1 RNA transcript ladder (lane 2) and a partial RNase T1 digest of the 3′-end-labeled transcript (lane 1). Similar data were obtained by using ladders made from full-length U1 RNA and from U1 RNA fragments labeled at other 3′ ends (results not shown). Thus, nt 1 to 17 at the 5′ end of U1 RNA (see Fig. 6) were not required for efficient interaction between U1 RNA and the 70K-LacZ fusion protein under these conditions.

Similarly, we examined the binding of ladders of 5′-end-labeled U1 RNA transcripts containing nt 1 to 84. Partial alkali ladders showed efficient binding to the 70K-LacZ fusion protein with 3′ sequences deleted up to nt 48 (Fig. 3B, lane 3), beyond which binding was not detectable relative to the controls (lanes 4 and 5). Again, nucleotide positions were confirmed by comparison with the total ladder and partial RNase T1 digests (lanes 1 and 2). Thus, in contrast to the results of bindings with 3′-end-labeled ladders, all detectable binding was lost with deletion of nt 48 from the 3′ end and did not decrease over several nucleotides.

Finally, to determine whether nt 18 to 48 alone could interact with the 70K-LacZ fusion protein, we examined the 5′ requirements for binding of a ladder ending with nt 48. A transcript of U1 nt 1 to 47 was purified (see Materials and Methods) and ligated to [^{32}P]pCp to create an RNA ending exactly with U1 RNA nt 48, the 3′ boundary for interaction with the 70K-LacZ protein that was identified above. Ladder species from this transcript exhibited the same sequence requirements for binding as did the other 3′-end-labeled transcripts (compare Fig. 3C with Fig. 3A). Therefore, the presence of nucleotides 3′ to nt 48 does not alter the minimal sequence requirement for the interaction. In addition, the smallest ladder species capable of efficient binding to the 70K-LacZ protein represents nt 18 to 48, thus demonstrating that nt 18 to 48 constitute a minimal U1 RNA protein-binding domain.

U1 RNA sequence requirements for interaction with the 70K U1 RNA-binding domain alone. Ladder species from the same nucleotide positions were observed to bind to both the 70K-LacZ and the 35-216–LacZ fusion proteins described in Fig. 1. Nucleotides 1 to 17 could be removed from a 3′-end-labeled U1 RNA without affecting binding, whereas removal of nt 18 or 19 resulted in reduced binding, and shorter digestion products had little affinity for the 35-216–LacZ fusion protein (Fig. 4A). Partial alkali ladders of 5′-end-labeled U1 RNA transcripts bound to the 35-216–LacZ fusion protein with 3′ nucleotides digested up to nt 48 (Fig. 4B). These boundaries of U1 RNA required for interaction with the 35-216–LacZ fusion protein (primarily the U1 RNA-binding domain [35]) are identical to those of the complete 70K-LacZ protein.

These data demonstrate that nt 18 to 48 constitute the minimal U1 RNA domain required for efficient interaction with both the full-length 70K-LacZ fusion protein and the U1 RNA-binding domain, 35-216–LacZ. This sequence represents essentially all of the predicted stem-loop I structure of U1 RNA (see Fig. 6). The possibility of additional sites of protein-RNA contact has not been entirely excluded, but, if additional sites exist, they are neither necessary nor sufficient for a stable interaction between the 70K-LacZ fusion

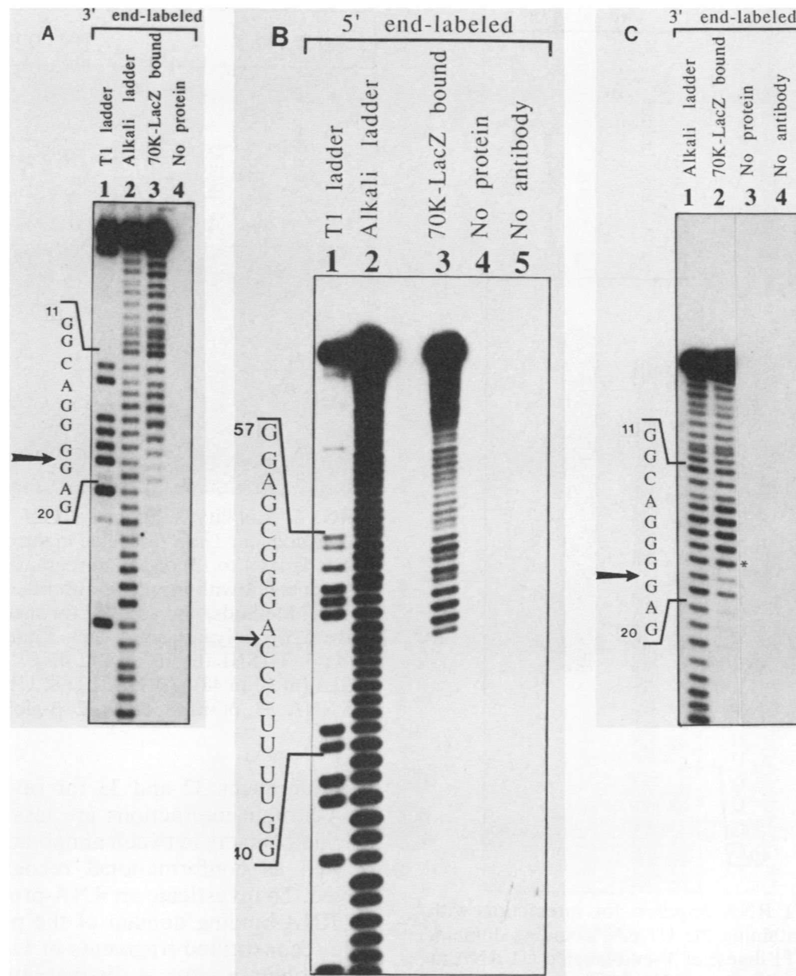

FIG. 3. Boundaries of U1 RNA required for interaction with full-length 70K-LacZ fusion protein. U1 RNA transcripts were end labeled, partially digested with alkali, and bound to the 70K-LacZ fusion protein as in Fig. 2B. RNA was analyzed on denaturing 10% polyacrylamide–8.3 M urea gels. Arrows indicate the point beyond which nucleotides could not be removed while maintaining an efficiency of binding similar to that of the intact U1 RNA: nucleotides above the arrows could be removed without affecting binding, whereas nucleotides below the arrows could not. (A) Lanes: 1, partial RNase T1 digest of 3'-end-labeled U1 RNA nt 1 to 94; 2, total 3'-end-labeled ladder of U1 RNA nt 1 to 94 after binding and immunoprecipitation; 3 and 4, RNA species that were bound and coprecipitated with 70K-LacZ protein (lane 3) or no protein (lane 4). (B) Lanes: 1, partial RNase T1 digest of 5'-end-labeled U1 RNA nt 1 to 84; 2, total 5'-end-labeled ladder of U1 RNA nt 1 to 84 after binding and immunoprecipitation; 3 to 5, RNA species that were bound and coprecipitated with 70K-LacZ protein (lane 3), no protein (lane 4), or no antibody (lane 5). (C) Lanes: 1, total 3'-end-labeled ladder of U1 RNA nt 1 to 48 after binding and immunoprecipitation; 2 to 4, RNA species that were bound and coprecipitated with 70K-LacZ protein (lane 2), no protein (lane 3), or no antibody (lane 4). The asterisk indicates an RNA species consisting of U1 RNA nt 18 to 48 exactly, which is the minimal RNA sequence observed to bind to the 70K-LacZ fusion protein with efficiency similar to that of full-length U1 RNA.

protein and U1 RNA under these conditions of reconstitution.

Interaction of U1 RNA fragments with in vitro-translated 70K protein. As an alternative assay that involves neither a LacZ fusion protein solubilized in 4 M urea nor recognition of an RNA-protein complex by antibody, we used a native gel system for analysis of RNA-protein complexes (35). Various in vitro-transcribed RNAs were assayed for the ability to produce a complex of altered mobility after incubation with in vitro-transcribed and translated ^{35}S-labeled 70K polypeptide representing the U1 RNA-binding domain (amino acids 92 to 216). A transcript consisting of U1 RNA nt 16 to 48 (U1SL1; see Materials and Methods) (Fig. 5, lane 6), as well as full-length U1 RNA (lane 5), created a complex of slower mobility. In contrast, a slower-migrating complex was not observed with U1 RNA mutants that lack stem-loop I (U1ΔSL1, lane 7; U1 nt 48 to 164, lane 9) or with unrelated RNAs (*E. coli* tRNA, lane 4; hY1 RNA, lane 10; pGEM3 RNA, lane 11; β-globin RNA, lane 12). No similar species or shifts in mobility were observed from products of unprogrammed translations (lanes 1 to 3). These results independently demonstrate that U1 RNA stem-loop I is necessary and sufficient to interact with the 70K protein.

Role of a bulged nucleotide and loop I. One of the best-characterized RNA-protein interactions is the binding of bacteriophage R17 coat protein to a stem-loop structure in R17 replicase mRNA (4, 38). This stem structure, similar to stem-loop I of U1 RNA, contains a predicted "bulged" A nucleotide, which is required for the association of R17 coat protein (47). Therefore, we constructed a U1 RNA mutant lacking the bulged U at position 22. This mutant RNA coprecipitated with the 70K-LacZ fusion protein with effi-

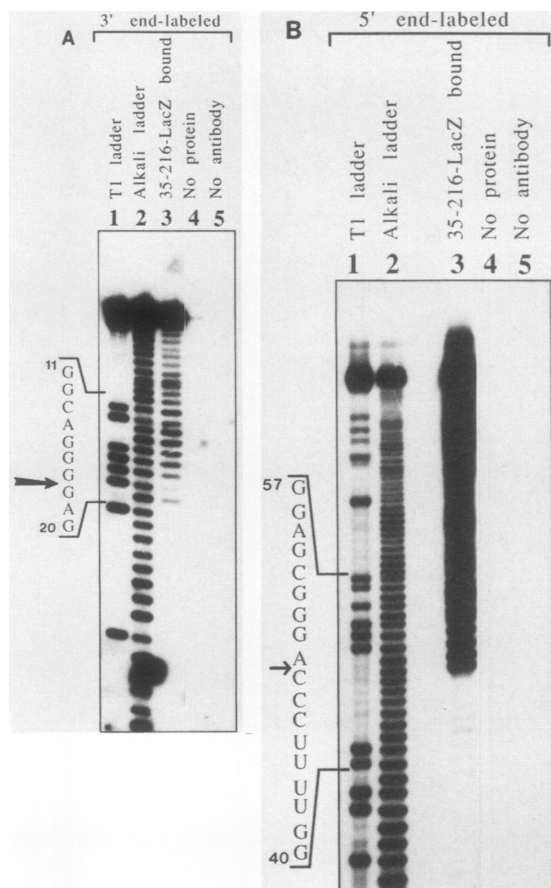

FIG. 4. Boundaries of U1 RNA required for interaction with 70K-LacZ fusion protein containing the U1 RNA-binding domain. (A) Lanes: 1, partial RNase T1 digest of 3'-end-labeled U1 RNA nt 1 to 94; 2, total 3'-end-labeled ladder of U1 RNA nt 1 to 94 after binding and immunoprecipitation; 3 to 5, RNA species that were bound and coprecipitated with 35-216–LacZ protein (lane 3), no protein (lane 4), or no antibody (lane 5). (B) Lanes: 1, partial RNase T1 digest of 5'-end-labeled U1 RNA nt 1 to 84; 2, total 5'-end-labeled ladder of U1 RNA nt 1 to 84 after binding and immunoprecipitation; 3 to 5, RNA species that were bound and coprecipitated with 35-216–LacZ protein (lane 3), no protein (lane 4), or no antibody (lane 5).

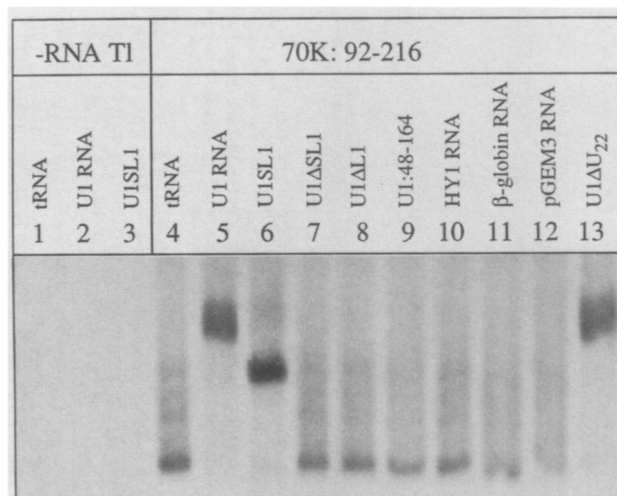

FIG. 5. Mobility shift analysis of the 70K protein U1 RNA-binding domain. The ^{35}S-labeled in vitro-translated 70K polypeptide 92-216 (lanes 4 to 13) or an unprogrammed translation (lanes 1 to 3) was incubated with in vitro-transcribed RNAs (described in Materials and Methods) and assayed for specific complex formation on a native 5.5% polyacrylamide gel. Lanes: 1, *E. coli* tRNA; 2, U1 RNA; 3, U1SL1 (nt 16 to 48); 4, *E. coli* tRNA; 5, U1 RNA; 6, U1SL1 (nt 16 to 48); 7, U1ΔSL1; 8, U1ΔL1; 9, U1 nt 48 to 164; 10, hY1 RNA; 11, pGEM3 RNA; 12, β-globin RNA; 13, U1ΔU$_{22}$ RNA.

ciency equivalent to that of the wild-type U1 RNA transcript (Fig. 2, lanes 20) and created a complex with in vitro-translated 70K protein of slower mobility on native gels (Fig. 5, lane 13). We also have examined the requirement for sequences within the loop of stem-loop I. Phylogenetic comparisons of yeast and mammalian U1 sequences have indicated that stem-loop I is conserved as a secondary structure (without sequence conservation in the stem) but that the sequence of loop I is specifically conserved (18, 42). A U1 RNA mutant lacking the loop I sequence failed to bind to 70K-LacZ fusion proteins (Fig. 2, lanes 19) or to in vitro-translated 70K protein (Fig. 5, lane 8), indicating that either the loop sequence or the stem-loop conformation is necessary for interaction with the 70K protein.

DISCUSSION

Protein-binding sites on RNA. Although chemical probes and crystallography have revealed important contacts between bases in DNA and a number of regulatory proteins (see references 32 and 33 for reviews), details of specific RNA-protein interactions are less well studied. Sequence-specific contacts between amino acids and nucleotide bases, as well as conformational recognition, are probably involved. To investigate an RNA-protein interaction for which the RNA-binding domain of the protein is known (35), we have reconstituted fragments of U1 RNA with recombinant 70K protein in vitro. A distinct portion of U1 RNA consisting of 31 nt from positions 18 to 48 from the 5' end was necessary and sufficient for binding to the protein. From the RNA structural predictions of Zuker and Stiegler (50) and comparisons of U1 RNA from evolutionarily diverse organisms (2, 30), this region has been shown to form a stable stem-loop of RNA (Fig. 6).

Other previously studied examples of RNA structures that interact directly with protein include the RNAs which bind to bacteriophage R17 coat protein (4), ribosomal protein L54 (which binds 5S RNA [15, 16]), and tRNA synthetases (14, 31, 39). In R17, a 21-nt RNA stem-loop was analyzed by saturation mutagenesis for the contribution of each nucleotide to the binding to R17 coat protein (4, 38). In this RNA structure, both a stem with a bulged A residue and the sequence of the loop were essential for the interaction (47). In our study, the unpaired bases composing loop I of U1 RNA were required; however, a predicted bulged U in the stem was not required for binding to the 70K protein. Thus, bulged or unpaired bases within a stem of RNA may not be generally required for protein recognition.

Stem-loop I of U1 RNA, where 70K protein binds, has also been predicted to form in the corresponding RNA of divergent species, including *Saccharomyces cerevisiae* (19, 42). The sequence of the loop is more conserved among species than is the sequence of the stem, although a stem of similar size is present in each case. Recent studies have shown that the human 70K protein can also recognize the yeast U1 RNA counterpart, snR19 (C. C. Query, C. Guthrie, and J. D. Keene, unpublished results), suggesting that the

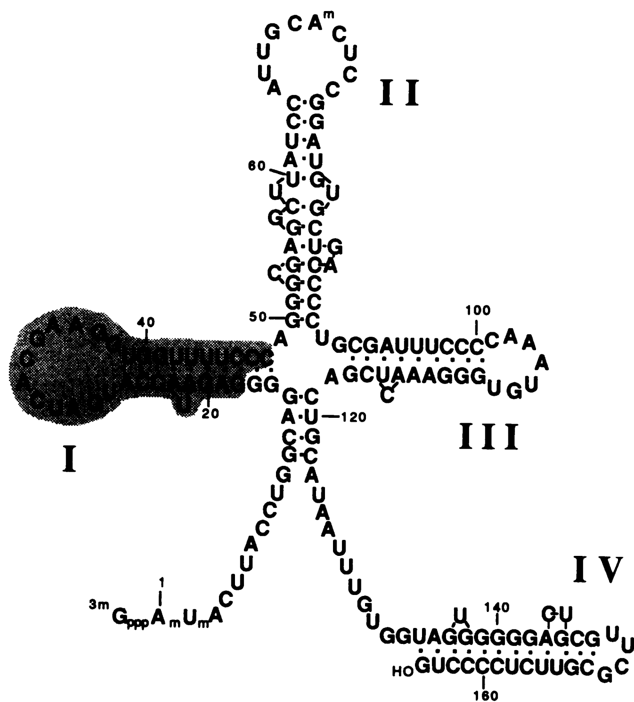

FIG. 6. Domain of U1 RNA that recognizes the 70K U1 snRNP protein. A proposed secondary structure of human U1 RNA showing stem-loops I, II, III and IV (2, 30). Stem-loops II, III, and IV could be eliminated without affecting the binding to the 70K-LacZ fusion protein, but nt 18 to 48 of stem-loop I were necessary and sufficient for binding.

RNA sequence specificity of the interaction lies mostly within the loop I nucleotides.

Conformation of RNA fragments. It is possible that when a portion of U1 RNA is synthesized separately from the rest of the molecule, it assumes a different conformation. For example, the folding of stem-loop I may be affected by sequences in other regions of U1 RNA. Our conclusions are based upon indications that RNA-RNA interactions within a single stem-loop of U1 RNA approximately reflect the natural state in the intact molecule. Free-energy calculations, compensatory evolutionary changes, RNase susceptibility (2, 28), and sites of psoralen cross-linking (3) predict the existence of stem-loops I, II, III, and IV in the intact U1 RNA molecule. RNA secondary structure analysis (49) also strongly supports the formation of each stem-loop of U1 RNA when synthesized as independent units (data not shown). Recent evidence from our laboratory has demonstrated that isolated stem-loop II binds independently to another U1 RNA-associated protein, the 32-kilodalton A protein (23). U1 RNA constructs used in this study that do not bind to the 70K protein, but that retain an intact stem-loop II, will bind to the A protein; similarly, U1 RNA constructs that do not bind to the A protein, but that retain stem-loop I, will bind to the 70K protein (C. Lutz-Freyermuth, C. C. Query, and J. D. Keene, unpublished results). We also have reported that stem-loop II, whether independently formed or in the intact molecule, is required in a folded conformation for recognition by a unique anti-U1 RNA antibody (6). Such findings suggest that these independent stem-loops retain conformations similar to those in the intact U1 RNA molecule, although further experimentation is needed.

U1 snRNP structure. Recent studies by Hamm et al. (13) suggested that only the C protein of U1 snRNPs directly contacts the RNA and that the 70K and A proteins bind primarily through protein-protein interactions. The data supporting these conclusions were obtained from reconstitution assays, in which proteins of a total *Xenopus* oocyte extract were allowed to assemble on mutant or wild-type U1 RNA and were then immunoprecipitated to analyze which proteins were associated with the RNA. Several observations were used to argue that the C protein binds specifically to loop I of U1 RNA and that binding of the 70K and A proteins is C protein dependent and occurs primarily through protein-protein interactions. First, a double-point mutation in loop I abolished binding of all three proteins, suggesting that these proteins may not bind independently. Second, analysis of several mutants in stem I failed to demonstrate binding of 70K or A in the absence of the C protein binding, as determined by using C protein-specific antiserum. Third, the C protein remained bound to U1 snRNPs under high-salt conditions (750 mM NaCl), whereas 70K and A did not bind under these conditions. Our results are in contradiction with the interpretations of Hamm et al. (13), in that we have demonstrated that stem-loop I of U1 RNA can bind to the 70K protein directly and independently of other proteins. This is consistent with the finding of Woppmann et al. (46) that the 70K protein can be cross-linked to U1 RNA. In addition, we have reported that stem-loop II of U1 RNA directly contacts the A protein (23). Thus, we conclude that the C protein is not required as an intermediary for binding of the 70K or A proteins to U1 RNA.

Our results are compatible with those of Hamm et al. on several points, although our interpretations differ. The demonstration here that deletion of the stem-loop I sequences results in loss of 70K binding is consistent with their observation that mutations disrupting stem-loop I decrease 70K binding. Thus, the stem-loop I structure is clearly important for 70K binding. We also have observed salt sensitivity of U1 RNA binding to the 70K protein: 70K-LacZ proteins lost U1 RNA-binding activity when the assays were performed in the presence of ≥300 mM NaCl (data not shown), suggesting a significant electrostatic component to the interaction, although the native 70K protein in the absence of 4 M urea and β-galactosidase is likely to show a different salt sensitivity.

We cannot explain why *Xenopus* C protein (13) and human 70K protein both seem to bind to U1 stem-loop I, but the differing assays and experimental components may be a source of differing results. It is unlikely, however, that *Xenopus* and human 70K proteins associate differently with U1 RNA, given the high conservation of U1 RNA (reviewed in reference 37) and the near identity in amino acid sequence of the *Xenopus* 70K protein (8) with the U1 RNA-binding domain of the human 70K protein (35). In addition, the complex whole-cell-extract reconstitution system of Hamm et al. (12, 13) may detect multicomponent interactions which could augment snRNP stability. The methods used in these and other previous studies may not have allowed discrimination of protein-RNA interactions from protein-protein-RNA interactions (12, 13, 34). We recognize that our assays may not detect all RNA contacts by these proteins. However, the apparent efficiency of association with the 70K protein was the same when either the complete U1 RNA or the 31-nt stem-loop I was used, and the nucleotides required for binding to 70K protein were identical when either the complete 70K protein or just the U1 RNA-binding domain was used. These data suggest that most, if not all, of the favorable free energy of U1 RNA-70K binding must be

provided by the interaction between the 31-nt stem-loop I and the 70K U1 RNA-binding domain. From data available to date, specific RNA-protein contacts that can be definitively deduced are that stem-loop I (nt 18 to 48) of U1 RNA binds to the 70K protein as reported above, stem-loop II (nt 48 to 84) binds to the A protein (23), and the region of the AU_nG motif (nt 126 to 130) binds to the Sm complex (see reference 26 for a review). Although the structure of the U1 snRNP is likely to involve both protein-protein and protein-RNA interactions, it is not yet clear what specific contacts occur between the protein components of the U1 snRNP.

Implications for U1 snRNP-spliceosomal interactions. U1 RNA interacts at 5′ splice junctions of pre-mRNA through RNA-RNA base pairing. Our finding that nt 18 to 48 of U1 RNA are sufficient to recognize the 70K protein under these conditions of reconstitution leaves the 5′ 17 nt of U1 RNA free to interact with pre-mRNA. However, we cannot address contacts in this region by other proteins. It remains possible that the 70K protein also may modify or recognize base-pairing interactions between U1 RNA and pre-mRNA.

Recent genetic analysis of U1 RNA mutations in stem-loop I has suggested that specific nucleotides in loop I are important for U1 RNA stability for splicing efficiency (48). In contrast, stem I is required as a secondary structure only, for those nucleotides examined. These requirements for U1 RNA stability and splicing efficiency are compatible with our finding that an intact stem-loop I is required for binding to the 70K protein, suggesting a role for the 70K protein in stabilizing U1 RNAs and in splicing efficiency.

ACKNOWLEDGMENTS

This work was supported by grants from the Arthritis Foundation and the National Institutes of Health to J.D.K. C.C.Q. was supported by Medical Scientist Training Program grant GM 02121. R.C.B. is the recipient of Physician Scientist Award AI 00909 from the National Institute of Allergy and Infectious Diseases. J.D.K. is a Pew Scholar in the Biomedical Sciences.

LITERATURE CITED

1. **Black, D. L., B. Chabot, and J. A. Steitz.** 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell **42**:737–750.
2. **Branlant, C., A. Krol, and J.-P. Ebel.** 1981. The conformation of chicken, rat and human U1A RNAs in solution. Nucleic Acids Res. **9**:841–858.
3. **Calvet, J. P., and J. A. Myers.** 1987. In-vivo secondary structure analysis of the small nuclear RNA U1 using psoralen cross-linking. J. Mol. Biol. **197**:543–553.
4. **Carey, J., V. Cameron, P. L. de Haseth, and O. C. Uhlenbeck.** 1983. Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. Biochemistry **22**:2601–2610.
5. **Chabot, B., and J. A. Steitz.** 1987. Multiple interactions between the splicing substrate and small nuclear ribonucleoproteins in spliceosomes. Mol. Cell. Biol. **7**:281–293.
6. **Deutscher, S. L., and J. D. Keene.** 1988. A sequence-specific conformational epitope on U1 RNA is recognized by a unique autoantibody. Proc. Natl. Acad. Sci. USA **85**:3299–3303.
7. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
8. **Etzerodt, M., R. Vignali, G. Ciliberto, D. Scherly, I. W. Mattaj, and L. Philipson.** 1988. Structure and expression of a *Xenopus* gene encoding a snRNP protein (U1 70K). EMBO J. **7**:4311–4321.
9. **Fisher, D. E., W. H. Reeves, G. E. Conner, G. Blobel, and H. G. Kunkel.** 1984. Pulse labeling of small nuclear ribonucleoproteins *in vivo* reveals distinct patterns of antigen recognition by human autoimmune antibodies. Proc. Natl. Acad. Sci. USA **81**:3185–3189.
10. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for prediction of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.
11. **Guthrie, C., and B. Patterson.** 1988. Spliceosomal snRNAs. Annu. Rev. Genet. **22**:387–419.
12. **Hamm, J., M. Kazmaier, and I. W. Mattaj.** 1987. *In vitro* assembly of U1 snRNPs. EMBO J. **6**:3479–3485.
13. **Hamm, J., V. L. van Santen, R. A. Spritz, and I. W. Mattaj.** 1988. Loop I of U1 small nuclear RNA is the only essential RNA sequence for binding of specific U1 small nuclear ribonucleoprotein particle proteins. Mol. Cell. Biol. **8**:4787–4791.
14. **Hou, Y.-M., and P. Schimmel.** 1988. A simple structural feature is a major determinant of the identity of a transfer RNA. Nature (London) **333**:140–145.
15. **Huber, P. W., and I. G. Wool.** 1984. Nuclease protection analysis of ribonucleoprotein complexes: use of the cytotoxic ribonuclease α-sarcin to determine the binding sites for Escherichia coli ribosomal proteins L5, L18, and L25 on 5S rRNA. Biochemistry **81**:322–326.
16. **Huber, P. W., and I. G. Wool.** 1986. Use of the cytotoxic nuclease α-sarcin to identify the binding site on eukaryotic 5S ribosomal ribonucleic acid for the ribosomal protein L5. J. Biol. Chem. **261**:3002–3005.
17. **Keene, J. D., M. Schubert, R. A. Lazzarini, and M. Rosenberg.** 1978. Nucleotide sequence homology at the 3′ terminus of RNA from vesicular stomatitis virus and in defective interfering particles. Proc. Natl. Acad. Sci. USA **75**:3225–3229.
18. **Kramer, A., W. Keller, B. Appel, and R. Lührmann.** 1984. The 5′ terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. Cell **38**:299–307.
19. **Kretzner, L., B. C. Rymond, and M. Rosbash.** 1987. S. cerevisiae U1 RNA is large and has limited primary sequence homology to metazoan U1 snRNA. Cell **50**:593–602.
20. **Krieg, P. A., and D. A. Melton.** 1984. Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs. Nucleic Acids Res. **12**:7057–7070.
21. **Kurilla, M. G., and J. D. Keene.** 1983. The leader RNA of vesicular stomatitis virus is bound by a cellular protein reactive with anti-La lupus antibodies. Cell **34**:837–845.
22. **Lerner, M. R., and J. A. Steitz.** 1979. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc. Natl. Acad. Sci. USA **76**:5495–5499.
23. **Lutz-Freyermuth, C., and J. D. Keene.** 1989. The U1 RNA-binding site of the U1 small nuclear ribonucleoprotein (snRNP)-associated A protein suggests a similarity with U2 snRNPs. Mol. Cell. Biol. **9**:2975–2982.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Maniatis, T., and R. Reed.** 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. Nature (London) **325**:673–678.
26. **Mattaj, I. W.** 1988. U snRNP assembly and transport, p. 100–114. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, New York.
27. **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.
28. **Milligan, J. F., D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck.** 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. **15**:8783–8798.
29. **Mount, S. M., I. Pettersson, M. Hinterberger, A. Karmas, and J. A. Steitz.** 1983. The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site *in vitro*. Cell **33**:509–518.
30. **Mount, S. M., and J. A. Steitz.** 1981. Sequences of U1 RNA from Drosophila melanogaster: implications for U1 secondary

structure and possible involvement in splicing. Nucleic Acids Res. **9**:6351–6368.
31. **Normanly, J., R. C. Ogden, S. J. Horvath, and J. Abelson.** 1986. Changing the identity of a transfer RNA. Nature (London) **321**:213–219.
32. **Ollis, D. L., and S. W. White.** 1987. Structural basis of protein-nucleic acid interactions. Chem. Rev. **87**:981–995.
33. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
34. **Patton, J. R., and T. Pederson.** 1988. The M_r 70,000 protein of the U1 small nuclear ribonucleoprotein particle binds to the 5′ stem-loop of U1 RNA and interacts with Sm domain proteins. Proc. Natl. Acad. Sci. USA **85**:747–751.
35. **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.
36. **Query, C. C., and J. D. Keene.** 1987. A human autoimmune protein associated with U1 RNA contains a region of homology that is cross-reactive with retroviral $p30^{gag}$ antigen. Cell **51**: 211–220.
37. **Reddy, R., and H. Busch.** 1988. Small nuclear RNAs: RNA sequences, structure, and modifications, p. 1–37. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, New York.
38. **Romaniuk, P. J., P. Lowary, H.-N. Wu, G. Stormo, and O. C. Uhlenbeck.** 1987. RNA binding site of R17 coat protein. Biochemistry **26**:1563–1568.
39. **Sampson, J. R., A. B. DiRenzo, L. S. Behlen, and O. C. Uhlenbeck.** 1989. Nucleotides in yeast $tRNA^{Phe}$ required for its specific recognition by its cognate synthetase. Science **243**: 1363–1366.
40. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
41. **Sharp, P. A.** 1987. Splicing of messenger RNA precursors. Science **235**:766–771.
42. **Siliciano, P. G., M. H. Jones, and C. Guthrie.** 1987. Saccharomyces cerevisiae has a U1-like small nuclear RNA with unexpected properties. Science **237**:1484–1487.
43. **Spritz, R. A., K. Strunk, C. S. Surowy, S. O. Hoch, D. E. Barton, and U. Francke.** 1987. The human U1-70K snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing and RNA-binding. Nucleic Acids Res. **15**: 10373–10391.
44. **Tabor, S., and C. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84**:4767–4771.
45. **Theissen, H., M. Etzerodt, R. Reuter, C. Schneider, F. Lottspeich, P. Argos, R. Lührmann, and L. Philipson.** 1986. Cloning of the human cDNA for the U1 RNA-associated 70K protein. EMBO J. **5**:3209–3217.
46. **Woppmann, A., J. Rinke, and R. Lührmann.** 1988. Direct cross-linking of snRNP proteins F and 70K to snRNAs by ultra-violet radiation *in situ*. Nucleic Acids Res. **16**:10985–11004.
47. **Wu, H. N., and O. C. Uhlenbeck.** 1987. Role of a bulged A residue in a specific RNA-protein interaction. Biochemistry **26**:8221–8227.
48. **Yuo, C.-Y., and A. M. Weiner.** 1989. Genetic analysis of the role of human U1 snRNA in mRNA splicing. I. Effect of mutation in the highly conserved stem-loop I of U1. Genes Dev. **3**:697–707.
49. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell **46**:827–835.
50. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9**:133–148.