Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 747–751, February 1988
Biochemistry

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

(mRNA splicing cofactor/RNA–protein interactions/antibody-mediated nuclease protection/oligodeoxynucleotide-directed RNase H cleavage)

JEFFREY R. PATTON AND THORU PEDERSON

Cell Biology Group, Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545

Communicated by Mahlon Hoagland, October 9, 1987 (received for review August 10, 1987)

ABSTRACT **The U1 small nuclear ribonucleoprotein (snRNP) particle, a cofactor in mRNA splicing, contains nine proteins, six of which are also present in other U snRNPs and three of which are specific to the U1 snRNP. Here we have used a reconstituted human U1 snRNP together with snRNP monoclonal antibodies to define the RNA binding sites of one of the U1 snRNP-specific proteins. When Sm monoclonal antibody (specific for the B′, B, and D proteins of U snRNPs) was bound to U1 snRNPs prior to micrococcal nuclease digestion, the same ≈24 nucleotide fragment of U1 RNA (corresponding to nucleotides 120–143 and termed the "Sm domain") was protected as when no antibody was bound prior to digestion. In contrast, when RNP monoclonal antibody, which reacts with the U1 snRNP-specific M_r 70,000 protein, was bound, additional U1 RNA regions were protected against nuclease digestion. This phenomenon, which we term "antibody-mediated nuclease protection," was exploited to map the position of the M_r 70,000 protein to stem–loop I of U1 RNA. However, there were also sites of M_r 70,000 protein interaction with more 3′-ward regions of U1 RNA, particularly the Sm domain. This indicates that in the three-dimensional structure of the U1 snRNP, the RNP and Sm antigens are in contact with each other. The proximity of the M_r 70,000 protein's RNA binding site (stem–loop I) to the functionally important 5′ end of U1 RNA suggests that this protein may be involved in the recognition of, or stabilization of base pairing with, pre-mRNA 5′ splice sites.**

The U1 small nuclear (sn) ribonucleoprotein (RNP) particle is a cofactor in mRNA splicing (1). U1 RNA is complexed with nine proteins (2), three of which, termed the M_r 70,000, A, and C proteins (2–8), are specific to the U1 snRNP. Human autoantibodies and murine monoclonal antibodies react with snRNPs (9, 10). In particular, a murine monoclonal antibody, termed RNP, reacts specifically with the M_r 70,000 U1 snRNP protein (11), whereas a second murine monoclonal antibody, termed Sm, reacts with three proteins that are present in all the U snRNPs (7).

To understand how the U1 snRNP participates in mRNA splicing, it will be important to know its structure. We have recently reconstituted the U1 snRNP by incubating SP6-transcribed human U1 RNA in a HeLa cell S100 fraction under specific conditions (12). The reconstituted particle possesses properties of the native U1 snRNP, including sedimentation coefficient, buoyant density, reactivity with Sm and RNP antibodies, and Mg^{2+}-dependent RNP conformation as probed by nuclease digestion (12). This latter property of both the reconstituted and native U1 snRNPs reflects a switch of particle conformation, in which additional regions of the U1 RNA become nuclease-protected by particle proteins at higher Mg^{2+} concentrations (13).

In the course of investigating the use of antibodies against U1 snRNP to select RNP fragments after nuclease digestion, we discovered that, if added before digestion, some (but, importantly, not all) antibodies mediate additional nuclease protection in the U1 snRNP. Because one of these antibodies reacts with a specific epitope, we have been able to exploit this phenomenon of "antibody-mediated nuclease protection" to locate a specific protein on the U1 RNA. A brief account of this work has been presented.*

MATERIALS AND METHODS

Reconstitution of U1 snRNP. The U1 RNA gene clone pHU1 (12) was linearized with *Bam*HI and was transcribed with SP6 RNA polymerase in the presence of 1 mM m^7GpppG, 50 μM GTP, 250 μM each UTP, ATP, and CTP, and [α-^{32}P]GTP (4 mCi/ml; 1 Ci = 37 GBq). Unincorporated nucleotides were removed by gel filtration (Biogel P-60). U1 snRNP was reconstituted in a HeLa S100 fraction (12). Reactions contained 0.5 mM ATP, 20 mM creatine phosphate, 3.2 mM $MgCl_2$, RNasin (400 units per ml), oligodeoxynucleotide complementary to the SP6 polylinker region (40 μg/ml), and 5 ng of ^{32}P-labeled U1 RNA (specific activity 1.0–1.3 × 10^8 cpm per μg), and 60% (vol/vol) S100 extract. The reconstituted U1 snRNP particle was purified on linear 10–30% glycerol gradients (12).

Micrococcal Nuclease Digestion in the Presence of snRNP Antibodies. Reconstituted U1 snRNP in 50 mM NH_4Cl/20 mM Tris·HCl, pH 7.5/5 mM 2-mercaptoethanol/50 μM phenylmethylsulfonyl fluoride were incubated with antibodies or nonimmune IgG (both in 0.15 M NaCl/10 mM sodium phosphate buffer, pH 7.2) at 4°C for 30 min, followed by digestion with micrococcal nuclease (5000 units per ml) at 37°C for 30 min in the presence of 1 mM $CaCl_2$. The antigen–antibody complexes were selected by binding to protein A-Sepharose as described (14), and the RNA was isolated from the bound and unbound fractions by proteinase K digestion, phenol/chloroform extraction, and ethanol precipitation. (Particles not preincubated with antibodies or nonimmune IgG were nuclease digested and directly deproteinized, and the protein A-Sepharose step was omitted.)

Sequence Analysis of Protected RNA. RNA fragments from antibody-bound, nuclease-digested U1 snRNP were electro-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RNP, ribonucleoprotein; snRNP, small nuclear RNP; nt, nucleotide(s).
*RNA Processing Meeting, May 13–17, 1987, Cold Spring Harbor, NY.

phoresed on 10% polyacrylamide/8.3 M urea gels under denaturing conditions. RNA bands were visualized by autoradiography, eluted, and incubated with specific complementary oligodeoxynucleotides (see Fig. 2*A*) in the presence of a HeLa nuclear extract as a source of RNase H activity (15). The products were recovered by proteinase K digestion, phenol/chloroform extraction, and ethanol precipitation and were subjected to electrophoresis as described above. In other experiments, the initial RNA bands were eluted and digested with RNase T1 and reelectrophoresed.

Immunoblot Analysis of Antibodies. Enriched snRNP preparations from HeLa cells (4) were precipitated with 0.4% (vol/vol) HCl in acetone. The proteins were electrophoresed on 15% polyacrylamide gels (16), transferred to nitrocellulose (17), and stained for total protein and antibody reactivity using the method of King *et al.* (18). The Sm monoclonal antibody reacted only with the B′ and B proteins; the RNP monoclonal antibody reacted only with a M_r 70,000 protein. Nonimmune murine IgG did not react with any proteins in the blots.

RESULTS

During the course of nuclease digestion experiments on U1 snRNP, followed by retrieval of RNP fragments with different snRNP antibodies, we made an unexpected observation. In some cases, binding of a particular antibody to the U1 snRNP *before* digestion led to the protection of additional U1 RNA sequences. In the present report, we describe this phenomenon of antibody-mediated nuclease protection and employ it to map the location of the M_r 70,000 protein in the U1 snRNP.

When ^{32}P-labeled U1 snRNP was incubated with Sm monoclonal antibody and subsequently digested with micrococcal nuclease (5000 units per ml), the predominant RNA selected by protein A-Sepharose was a ≈24-nucleotide (nt) species (Fig. 1). RNase T1 and T2 digestion data (not shown) revealed that this ≈24-nt fragment contains nt 120–143 of U1 RNA, corresponding to the previously described "Sm domain" (13, 19, 20). Since this sequence is the same as the one protected when U1 snRNP is digested without prior antibody binding (13, 19, 20), it follows that Sm monoclonal antibody does not mediate appreciable nuclease protection. (We note minor bands in the lane labeled Sm, the upper two of which may correspond to bands 1 and 3 seen with RNP antibody; however, with Sm antibody these are recovered substoichiometrically with respect to the ≈24-nt fragment.) In contrast, when RNP monoclonal antibody was used, a different pattern of resistant fragments appeared (Fig. 1, bands 1–4), which contain three major RNA species longer than, and one shorter than, the Sm domain fragment. The RNP antibody nuclease-resistant RNA not bound to protein A-Sepharose contained the Sm domain RNA. When nonimmune IgG was incubated with U1 snRNP prior to digestion, no RNA was recovered in the protein A-Sepharose bound fraction, and only the expected Sm domain RNA was present in the nonbound fraction (Fig. 1). It follows that RNP antibody mediates protection against nuclease digestion of U1 RNA sequences that, with Sm or no antibody, are nuclease sensitive. It has been suggested that RNP antibody can react with RNA. When RNP antibody was incubated with U1 RNA and digested with nuclease, no resistant RNA was observed in either the bound or nonbound fraction (data not shown).

We used a series of complementary oligodeoxynucleotides (Fig. 2*A*) to orient the locations of the RNP antibody nuclease-resistant fragments within the U1 RNA molecule. The RNA fragments were eluted from the gel and incubated in the presence of oligodeoxynucleotides and RNase H (see *Materials and Methods*). As a control, the protected ≈24-nt

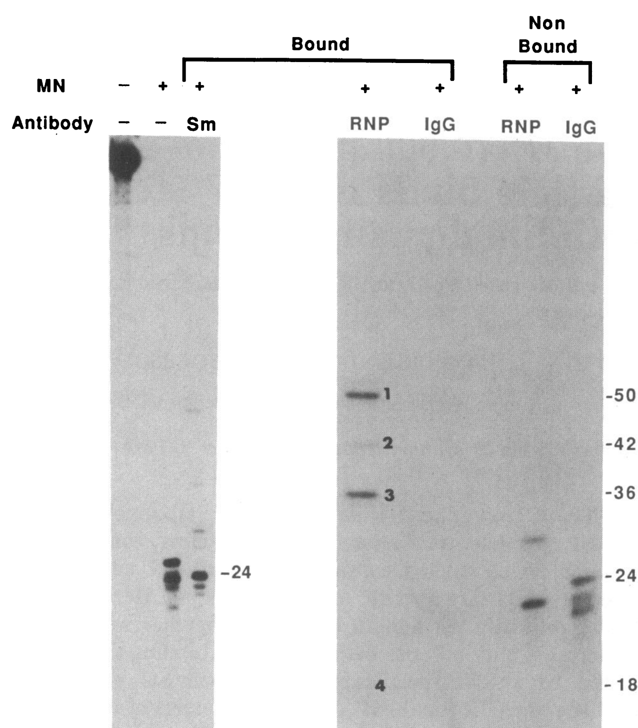
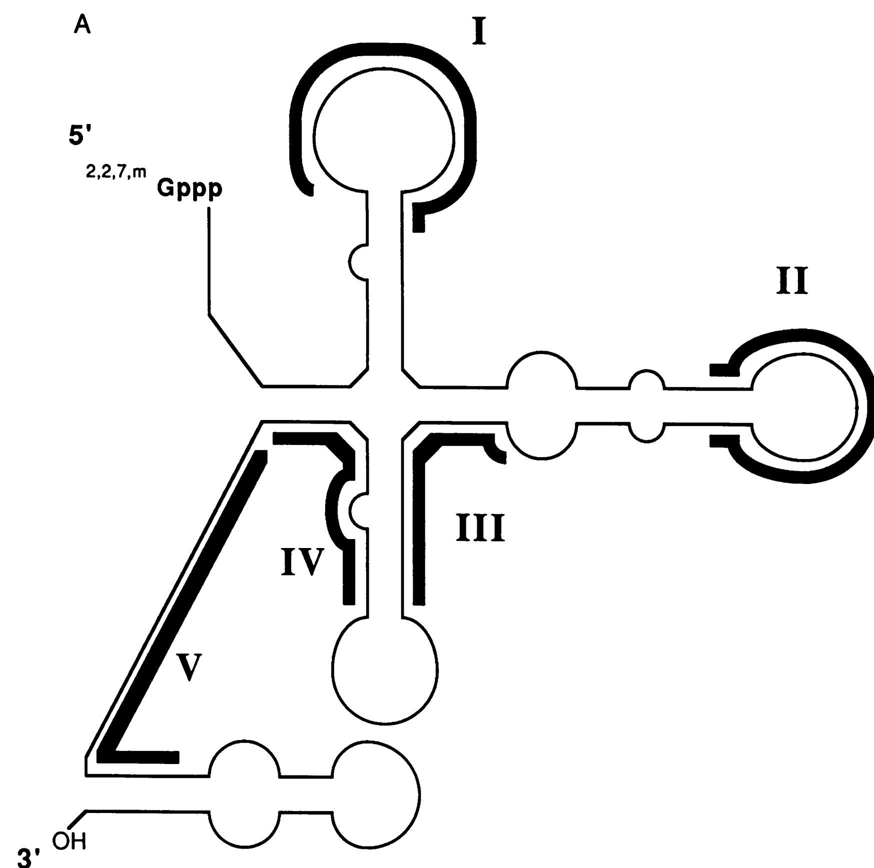


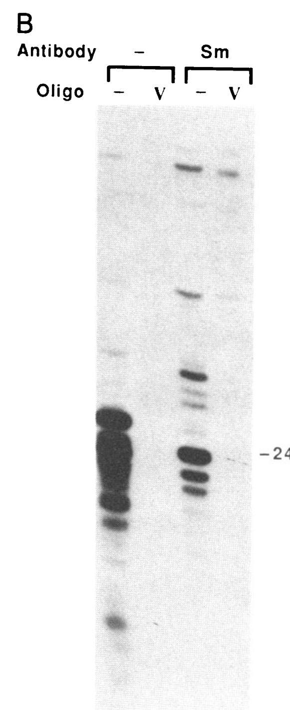
FIG. 1. RNP monoclonal antibody mediates nuclease protection of the U1 snRNP. The reconstituted U1 snRNP (12) was treated with Sm or RNP monoclonal antibody, mouse IgG, or buffer and then was digested with micrococcal nuclease (5000 units per ml). The RNP fragments containing bound antibody were selected by protein A-Sepharose. The material recovered (Bound), as well as the material flowing through protein A-Sepharose (Non Bound), was deproteinized and electrophoresed on a 10% polyacrylamide/urea gel. The approximate sizes (in nt) of some of the bands (estimated by comparison with DNA size markers) are indicated. The numbers (1–4) next to the bands in the bound lane labeled RNP denote RNA fragments excised from the gel for subsequent experiments. In the cases in which particles were not first incubated with antibody, they were digested with nuclease and directly deproteinized, omitting the protein A-Sepharose step. MN, micrococcal nuclease.

fragment obtained after digestion in either the presence or absence of Sm antibody (Fig. 1) was incubated with oligomer V, which is complementary to the Sm domain (Fig. 2*A*). As can be seen in Fig. 2*B*, this RNA was completely digested when oligomer V was present.

The same RNase H method was then used to identify the protected RNA fragments obtained after digestion of U1 snRNP to which RNP antibody had first been bound. As shown in Fig. 2*C*, protected RNA band 1 (see Fig. 1) was cleaved only with oligomer I, and band 4 was cleaved only with oligomer III. In contrast, bands 2 and 3 were cleaved with either oligomer IV or V. None of the bands were cleaved by oligomer II, which is complementary to the second loop of U1 RNA. It is of interest that the slowest migrating minor RNA band observed with Sm antibody is not oligomer V-sensitive (arrow, Fig. 2*B*), suggesting that it might be the same RNA fragment as band 1 seen with RNP antibody.

From the lengths of the protected RNAs as determined by their mobility in denaturing gels (Figs. 1 and 2*C*) and the fact that none of the sequences complementary to the oligonucleotides appear more than once in U1 RNA, the oligodeoxynucleotide/RNase H cleavage results lead to the map shown in Fig. 3. The sequence composition of bands 1 and 3 (see Fig. 1) was confirmed by RNase T1 digestion (data not shown). The widths of the stippled bars (Fig. 3) indicate the relative amounts of the RNA sequences recovered after





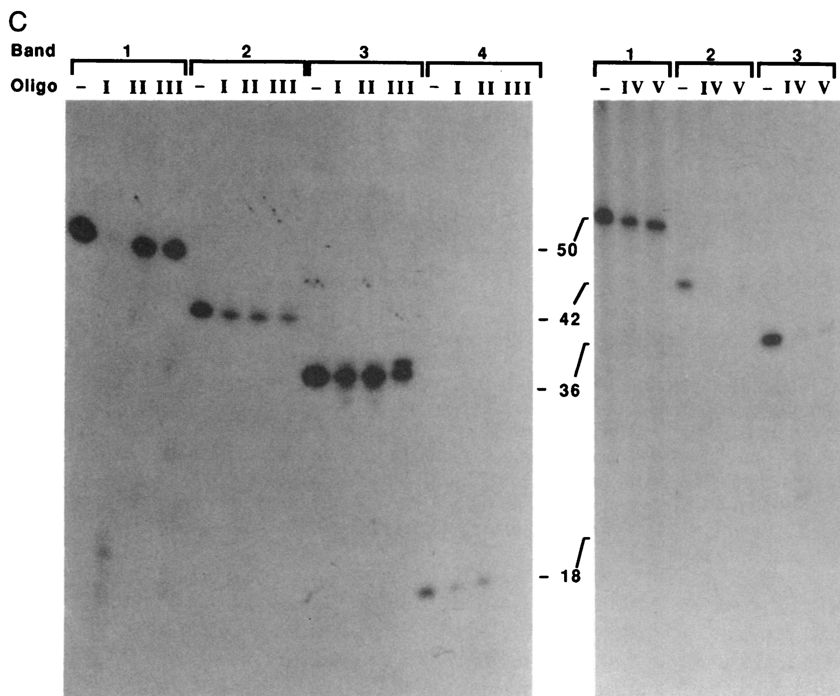


FIG. 2. RNase H mapping of protected fragments. (*A*) Oligodeoxynucleotides used to map the location of the protected U1 RNA fragments. Oligomer I is complementary to nucleotides 29–42; II, 64–77; III, 87–100; IV, 110–123; V, 124–143. (*B*) The fragments protected from nuclease digestion of the U1 snRNP with or without prior incubation with Sm antibody (see Fig. 1) were incubated with RNase H and oligomer V (V). The RNA products were isolated and electrophoresed along with an equivalent amount of input RNA not exposed to the oligomer (−). (*C*) The U1 RNA fragments protected from nuclease digestion after pretreatment of U1 snRNP with RNP antibody (bands 1–4 in Fig. 1) were incubated with the indicated oligomer in the presence of RNase H. The RNA products were isolated and electrophoresed along with an equivalent amount of input RNA not exposed to the oligomer (−). Oligo, Oligomer. The approximate sizes (in nt) of some of the bands are indicated.

nuclease digestion of RNP antibody-bound U1 snRNP. In particular, the fragments containing nt 10–59 and 111–143 (corresponding to bands 1 and 3 in Fig. 1) were recovered at a 1:1 stoichiometry.

The M_r 70,000 protein was the only one detected when RNP antibody was used to probe immunoblots of snRNP (data not shown and ref. 11). Since the RNA fragments recovered when RNP antibody was used contain sequences defined not only by oligomer I but also by oligomers III, IV, and V (Fig. 2 *B* and *C*), it follows that the M_r 70,000 protein interacts not only with stem–loop I but also with the indicated regions of stem III and the Sm domain (Fig. 3).

DISCUSSION

Learning the structure of the U1 snRNP will be an important step in solving the mechanism of mRNA splicing. The gross topography of protein mass on U1 RNA has previously been studied by nuclease digestion experiments (13, 19–21), which defined the binding site for the Sm "core" proteins

FIG. 3. Location of M_r 70,000 protein binding sites. The positions within U1 RNA of bands 1–4 from the RNP monoclonal antibody-mediated nuclease protection experiments are denoted by stippled bars. The thicknesses of the bars signify that bands 1 and 3 were recovered at a 1:1 stoichiometry, whereas bands 2 and 4 were not. The ≈24-nt fragment present after digestion in the presence of Sm (or no) antibody is indicated by the solid black bar. The sequence content of the protected regions was based on the sizes of the RNase H cleavage products and was confirmed by RNase T1 digestion patterns (data not shown). In particular, T1 digestion of band 1 yielded the 6-, 8-, and 9-mer oligonucleotides expected (corresponding to nt 30–35, 22–29, and 43–51, respectively) and did not yield the T1 oligonucleotides corresponding to nt 1–12 and 59–70. T1 digestion of bands 2 and 3 yielded a 7-mer and 9-mer, corresponding to nucleotides 112–118 and 123–131, respectively. T1 digestion of band 4 yielded a 7-mer, corresponding to nucleotides 86–92.

that are found in the major, nucleoplasmic snRNP. Here we have studied the location of one of the U1-specific proteins in the U1 snRNP by exploiting an antibody-mediated, nuclease-protection phenomenon. Using conditions of nuclease digestion in which only the Sm domain of (untreated) U1 snRNP is protected, we found that if a monoclonal antibody to the U1-specific M_r 70,000 protein is bound to the particle before nuclease digestion, additional RNA sequences are protected (Fig. 1).

What does this phenomenon of antibody-mediated nuclease protection reflect? It is not a matter of simply any antibody binding to the U1 snRNP, since Sm monoclonal antibody does not cause significant additional nuclease protection (Fig. 1). It is possible that binding of RNP monoclonal antibody induces a conformational change in the M_r 70,000 protein that, in turn, leads to a tighter binding of this protein to the RNA. Alternatively, it is possible that when bound to the M_r 70,000 protein, the antibody itself binds to adjacent (previously naked) RNA segments. This latter scenario requires invoking the existence of a high-affinity RNA-binding pocket in regions of this antibody other than the antigen-combining site, which seems improbable. Moreover, a direct protection of U1 RNA sequences by the antibody cannot be the case for the interaction of the M_r 70,000 protein with the Sm domain (Fig. 3), because this RNA sequence is already tightly bound by proteins other than the M_r 70,000 one (13, 19, 20).

Bands 1 and 3 (Fig. 3) were recovered at a 1:1 stoichiometry when RNP antibody was used, showing an interaction between the M_r 70,000 protein and other proteins bound to

the Sm domain. Compatible with this is the observation that, when Sm antibody is used, at least two RNA bands are seen in addition to the ≈24-nt Sm domain fragment (Fig. 1), and these comigrate with bands 1 and 3 seen with RNP antibody. In addition, a U1 RNA molecule that lacks nt 1–42 does not form a stable U1 snRNP even though the Sm core region is intact (unpublished results).

Of the 9 or 10 (2, 22) proteins present in the human U1 snRNP, the binding sites of four are now known: the M_r 70,000 protein (this report) and the Sm antibody-reactive B′, B, and D proteins (19). The approximate locations of three other proteins, E, F, and G, can be inferred from the observation that they are recovered together with the B′, B, and D proteins in Sm antibody-reactive antigenic complexes after nuclease digestion of U1 snRNP (data not shown). This is in agreement with the previous conclusion (19, 23) that the proteins B′, B, D, E, F, and G all reside in a single protein complex occupying the Sm domain of U1 RNA. With regard to the U1 snRNP-specific A and C proteins, we find that human autoantibodies that react with both of these proteins mediate nuclease protection of a ≈90-nt long fragment spanning nt 10–99 of the U1 RNA (unpublished results), suggesting that the A and/or C protein-binding sites include stem–loop II and part of stem III. This means that the great majority of U1 RNA is protein-covered in the U1 snRNP (see also ref. 20). Indeed, perhaps only loop III is not bound by protein; this follows from the hypersensitivity of the guanosine at position 107 to S1 nuclease or RNase T1 digestion of U1 snRNP (20, 21).

This extensively protein-covered structure of the U1 snRNP is reminiscent of the signal recognition particle (24–26). The fact that so much of the RNA is protein-covered in the U1 snRNP restricts the possibilities for intermolecular base pairing with other RNAs (notwithstanding the 5′ ends pairing with pre-mRNA 5′ splice sites). It is possible that the snRNPs participating in mRNA splicing interact with each other, and some evidence for snRNP–snRNP interactions in the spliceosome has been presented (27). The fact that almost all of the RNA in the U1 snRNP is complexed with protein reduces the possible sites for base pairing with other snRNP particles to loop III and/or the 3′ half of (an unwound) stem IV. Alternatively, perhaps snRNP–snRNP associations, to the extent that they occur altogether, are mediated by protein–protein interactions or by dynamic protein dissociation from spliceosome-associated snRNPs , thereby uncovering snRNA sequences for possible inter-snRNP base pairing.

The accurate recognition of pre-mRNA 5′ splice sites *in vitro* requires the RNP form of U1 RNA (28). Moreover, the U1 RNA that can be psoralen-crosslinked to pre-mRNA *in situ* is reactive with U1 snRNP antibodies (29), and such antibodies block splicing *in vitro* (30). Thus, it is clear that the RNP form of U1 RNA is the functionally important one. Perhaps the requisite base pairing between the 5′ splice site of pre-mRNA and the 5′ end of U1 RNA (31) requires stabilization by protein (28, 32), just as codon/anti-codon base pairing may be stabilized by ribosomal proteins. Our finding that the M_r 70,000 protein is positioned close to the functional 5′ terminus of U1 RNA, on stem–loop I, raises the possibility that it functions in splice-site targeting and/or binding stabilization. An interesting question is whether a U1 snRNP lacking the M_r 70,000 protein, but complete in all other respects, would be capable of functioning in splicing.

Note Added in Proof. While this paper was in press, Hamm *et al.* (33) reported that the first stem–loop of *Xenopus* U1 RNA is necessary for binding of the M_r 70,000 protein and also presented evidence for an interaction between the M_r 70,000 and Sm proteins.

We are grateful to John Goodchild for oligodeoxynucleotide synthesis in the Worcester Foundation's central facility (supported by Cancer Center Core Grant P30-12708) and to Stephen King for help with the immunoblots. We thank our colleagues Gary Kunkel, Ann Kleinschmidt, Sandra Mayrand, Ronald Patterson, and Nina Pedersen for advice and critical comments during the course of this work. This work was supported by National Institutes of Health grant GM-21595. J.R.P. holds a National Institutes of Health post-doctoral fellowship (GM-11399).

1. Maniatis, T. & Reed, R. (1987) *Nature (London)* **325,** 673–678.
2. Brunel, C., Sri-Widada, J. & Jeanteur, P. (1985) *Mol. Subcell. Biol.* **9,** 1–52.
3. Kinlaw, C. S., Swartz, S. D. & Berget, S. M. (1982) *Mol. Cell. Biol.* **2,** 1159–1166.
4. Hinterberger, M., Pettersson, I. & Steitz, J. A. (1983) *J. Biol. Chem.* **258,** 2604–2613.
5. Wieben, E. D., Madore, S. J. & Pederson, T. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1217–1220.
6. Kinlaw, C. S, Robberson, B. L. & Berget, S. M. (1983) *J. Biol. Chem.* **258,** 7181–7189.
7. Pettersson, I., Hinterberger, M., Mimori, T., Gottlieb, E. & Steitz, J. A. (1984) *J. Biol. Chem.* **259,** 5907–5914.
8. Billings, P. B. & Hoch, S. O. (1984) *J. Biol. Chem.* **259,** 12850–12856.
9. Lerner, M. R. & Steitz, J. A. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 5495–5499.
10. Lerner, E. A., Lerner, M. R., Janeway, C. A. & Steitz, J. A. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 2737–2741.
11. Billings, P. B., Allen, R. W., Jensen, F. C. & Hoch, S. O. (1982) *J. Immunol.* **128,** 1176–1180.
12. Patton, J. R., Patterson, R. J. & Pederson, T. (1987) *Mol. Cell. Biol.* **7,** 4030–4037.
13. Reveillaud, I., Lelay-Taha, M.-N., Sri-Widada, J., Brunel, C. & Jeanteur, P. (1984) *Mol. Cell. Biol.* **4,** 1890–1899.
14. Wieben, E. D. & Pederson, T. (1982) *Mol. Cell. Biol.* **2,** 914–920.
15. Krainer, A. R. & Maniatis, T. (1985) *Cell* **42,** 725–736.
16. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
17. Otter, T., King, S. M. & Witman, G. B. (1987) *Anal. Biochem.* **162,** 370–377.
18. King, S. M., Otter, T. & Witman, G. B. (1986) *Methods Enzymol.* **134,** 291–306.
19. Liautard, J.-P., Sri-Widada, J., Brunel, C. & Jeanteur, P. (1982) *J. Mol. Biol.* **162,** 623–643.
20. Lin, W.-L. & Pederson, T. (1984) *J. Mol. Biol.* **180,** 947–960.
21. Epstein, P., Reddy, R. & Busch, H. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 1562–1566.
22. Bringmann, P. & Luhrmann, R. (1986) *EMBO J.* **5,** 3509–3516.
23. Fisher, D. E., Conner, G. E., Reeves, W. H., Wisniewolski, R. & Blobel, G. (1985) *Cell* **42,** 751–758.
24. Walter, P. & Blobel, G. (1982) *Nature (London)* **299,** 691–698.
25. Gundelfinger, E. D., Krause, E., Melli, M. & Dobberstein, B. (1983) *Nucleic Acids Res.* **11,** 7363–7374.
26. Walter, P., Gilmore, R. & Blobel, G. (1984) *Cell* **38,** 5–8.
27. Konarska, M. M. & Sharp, P. A. (1987) *Cell* **49,** 763–774.
28. Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A. & Steitz, J. A. (1983) *Cell* **33,** 509–518.
29. Setyono, B. & Pederson, T. (1984) *J. Mol. Biol.* **174,** 285–295.
30. Padgett, R. A., Mount, S. M., Steitz, J. A. & Sharp, P. A. (1983) *Cell* **35,** 101–107.
31. Zhuang, Y. & Weiner, A. M. (1986) *Cell* **46,** 827–835.
32. Tatei, K., Takemura, K., Tanaka, H., Masaki, T. & Ohshima, Y. (1987) *J. Biol. Chem.* **262,** 11667–11674.
33. Hamm, J., Kazmaier, M. & Mattaj, I. W. (1987) *EMBO J.* **6,** 3479–3485.