The 3' Splice Site of Pre-Messenger RNA Is Recognized by a Small Nuclear Ribonucleoprotein

Benoit Chabot, Douglas L. Black
David M. LeMaster, Joan A. Steitz

Three regions in mammalian messenger RNA precursors (pre-mRNA's) are important for the specificity and efficiency of RNA splicing. Initially, sequence analyses and genetic studies identified the 5' and 3' splice sites as regions which contain sequences that are both conserved and essential for the process (1).

11) and by the use of antibodies to U1 snRNP's (12). Binding analyses performed during the course of in vitro splicing of a globin and an adenovirus pre-mRNA have shown that U1 snRNP's associate with the 5' splice site and U2 snRNP's with the intron region where the RNA branch will be formed (10, 13).

Abstract. *A component present in splicing extracts selectively binds the 3' splice site of a precursor messenger RNA (pre-mRNA) transcript of a human beta-globin gene. Since this component can be immunoprecipitated by either autoantibodies of the Sm class or antibodies specifically directed against trimethylguanosine, it is a small nuclear ribonucleoprotein (snRNP). Its interaction with the 3' splice site occurs rapidly even at 0°C, does not require adenosine triphosphate, and is altered by certain mutations in the 3' splice site region. Binding is surprisingly insensitive to treatment of the extract with micrococcal nuclease. The U5 particle is the only abundant Sm snRNP with a capped 5' end that is equally resistant to micrococcal nuclease. This suggests that, in addition to the U1 and U2 snRNP's, U5 snRNP's participate in pre-mRNA splicing.*

More recently, characterization of the intermediates generated during in vivo and in vitro splicing have revealed the presence of a third important region. This site contains a specific adenosine residue, located between 22 and 37 nucleotides (nt) upstream from the 3' splice site, where the 5' end of the intron forms a 2', 5' -phosphodiester bond to create an unusual branched RNA structure (2–7), often designated a lariat.

Recognition of the 5' splice site and the branch point have been shown to involve small nuclear ribonucleoproteins (snRNP's) which react with autoantibodies called anti-Sm (8). This group of RNA-protein complexes in mammalian cells includes abundant particles containing U1, U2, U5, or U4 and U6 snRNA's and an as yet unknown number of low abundance particles. That U1 and U2 snRNP's are essential components of the splicing machinery has been demonstrated by site-directed cleavage of the RNA moiety of these two snRNP's (9–

Meanwhile, the identity of the component (or components) responsible for recognition of the 3' splice site has remained elusive. The consensus sequence derived for 3' splice sites consists of a pyrimidine-rich stretch of about ten nucleotides followed by NCAG/G (where N is any nucleotide and the slash indicates the position of the splice junction) (1). Deletion of the entire 3' splice site region abolishes splicing (5, 6) while mutations affecting only the conserved AG dinucleotide allow formation of the branched intermediate but block intron excision and the final exon ligation step (5, 6, 14). On the basis of limited complementarity to the 3' splice site consensus sequence, both U1 and U2 snRNA's have been proposed to participate in 3' splice site recognition (15, 16). Experimentally, partially purified U1 snRNP's have been observed to bind to an artificial RNA transcript carrying a 3' splice site but at a much lower efficiency than to one that included a 5' splice site (17).

In another study, however, purified U1 snRNP's recognized the 5' but not the 3' splice site of a mouse beta-globin pre-mRNA transcript (18). These observations have left unresolved the question of whether U1 snRNP's or other splicing components are involved in specific recognition of the 3' splice site region.

Using several different antibodies to snRNP's, we have investigated the identity of a component that recognizes the 3' splice site region of wild-type and mutant human beta-globin transcripts in an in vitro splicing system. We show that some Sm snRNP binds specifically to the 3' splice site and present evidence that the U5 snRNP is the most likely candidate.

An Sm snRNP is involved in 3' splice site binding. Our first hints that the 3' splice site might be recognized by an Sm snRNP came from T1 ribonuclease (RNase) protection experiments performed with antibodies directed against either the Sm determinants on the snRNP proteins or the 2,2,7-trimethylguanosine (TMG) cap structure found on all the abundant U RNA's except U6 (8). These experiments were carried out in splicing extracts to which a labeled human beta-globin transcript that contained the first intron and flanking exons had been added (HβΔ6) (19, 20). Shortly after the addition of this substrate to the splicing mixture at 0°C, antibodies and T1 RNase were added. Fractionation of the immunoprecipitated material on 15 percent polyacrylamide-8M urea gels (Fig. 1A) revealed several discrete RNA fragments when either antibody to Sm (anti-Sm) (lanes 3 and 8) or antibody to TMG (anti-TMG) (lane 5) (21) was used.

Subsequent analyses of these fragments after a second round of T1 RNase digestion both by gel fractionation (Fig. 1B) and by two-dimensional fractionation (fingerprint) and secondary analyses demonstrated that one of them (band B) was a 19-nt oligomer (UCUAUUUU-CCCACCCUUAG) that maps to the 3' end of the intron (Fig. 2). Two larger fragments containing the 3' splice site (B' and B'', 27 and 37 nt, respectively) were also detected, but in lower yields than the B species (Fig. 1A, lanes 3, 5, and 8). The fourth prominent fragment (band A) immunoprecipitated by anti-Sm and anti-TMG (Fig. 1A, lanes 3, 5, and 8) was a 15-nt protected region (CAG/GU-UGGUAUCAAG) encompassing the 5' splice site. [The use of an $[\alpha^{-32}P]CTP$

(cytidine triphosphate)–labeled transcript (Fig. 1) overemphasizes the efficiency of 3' splice site binding since band B contains seven CTP-labeled residues while band A contains only one.]

The splicing extract used for the experiment shown in Fig. 1, lanes 1 to 6, contained exogenous adenosine triphosphate (ATP), which is required for splicing to proceed (19, 22). However, even when no ATP was added to the splicing mixture, snRNP binding to the 3' as well as the 5' splice site occurred (Fig. 1A, lanes 7 and 8). Binding to the 3' splice site was also observed later in the reaction (Fig. 1A, lane 4). The appearance of many of the additional bands in this lane can be ascribed to transcript protection by U1 or U2 snRNP's as reported previously (10) (legend to Fig. 1).

The component involved in 3' splice site binding is neither the U1 nor the U2 snRNP. To ascertain which of the Sm snRNP's might be involved in 3' splice site binding, we first examined pre-mRNA fragments immunoprecipitated by antibodies to U1 and U2 RNP's [anti-(U1)RNP and anti-(U2)RNP] (21). As we anticipated from previous work (10), at zero time in the splicing reaction U1 but not U2 snRNP's bound the 15-nt fragment (band A) containing the 5' splice site (Fig. 1A, lanes 1 and 2). Also visible in both these immunoprecipitates was the 3' splice site, 19-nt fragment (band B), but in barely detectable amounts compared to the anti-Sm or anti-TMG signal (compare lanes 1 and 2 with lanes 3 and 5). This extremely low level of 3' splice site binding was not noted in our previous study (10) because the transcript in that case was labeled with [α-^{32}P]GTP rather than [α-^{32}P]CTP. These observations raised the possibility that U1 or U2 snRNP's (or both) might be involved in 3' splice site binding, but that the antibodies somehow interfered with the snRNP's continued interaction with the 3' splice site or that the essential epitopes might become sequestered upon snRNP binding to the 3' splice site region.

We therefore analyzed fragments immunoprecipitated from splicing extracts in which U1 or U2 RNA had been selectively degraded by prior incubation with RNase H and an appropriate complementary oligodeoxynucleotide (10). The ability of the U1 snRNP to bind the 5' splice site (Fig. 3, lanes 1, 3, and 4) was abolished when the 5' end of its RNA was trimmed (Fig. 3, lanes 5, 7, and 8). By contrast, the immunoprecipitability of the 3' splice site fragment with anti-Sm (Fig. 3, lanes 3, 7, and 11) or anti-TMG (lanes 4, 8, and 12) did not change

after digestion of the 5' end of either U1 or U2 snRNA. The completeness of selective degradation of each targeted snRNA was verified by examining the small RNA's on gels and by analysis of splicing activity (not shown); in the experiment in Fig. 3, more than 95 percent of U1 or U2 snRNA was degraded, and in both cases less than 5 percent of the splicing activity remained. In addition, we found that the RNase H–degraded forms of U1 and U2 snRNP's remained immunoprecipitable by anti-Sm but not by anti-TMG. We conclude that the snRNP involved in binding the 3' splice site region cannot be either U1 or U2.

Binding is affected by mutations at the 3' splice site. If binding to the 3' splice site is physiologically meaningful, we would expect it to respond to mutations

that alter use of the 3' splice site during in vivo or in vitro splicing (or both). We therefore analyzed Sm snRNP binding to three mutant beta-globin substrates containing alterations at or close to the 3' splice site region (5) (Fig. 2). The β$^{Δ3'}$ mutant has suffered a 25-nt deletion that removes the 3' splice site as well as 2 nt of the second exon (−23 to +2 relative to the 3' splice junction). This deletion abolishes splicing without activating any cryptic 3' splice site in vivo or in vitro (5). In the binding assay (Fig. 4), band B disappeared, and no new splice site fragment was immunoprecipitated by anti-Sm (lane 9). The β$^{AG→GG}$ mutant has a point mutation in the next to the last nucleotide of the intron. This mutation blocks splicing after formation of the lariat intermediate and also lowers the
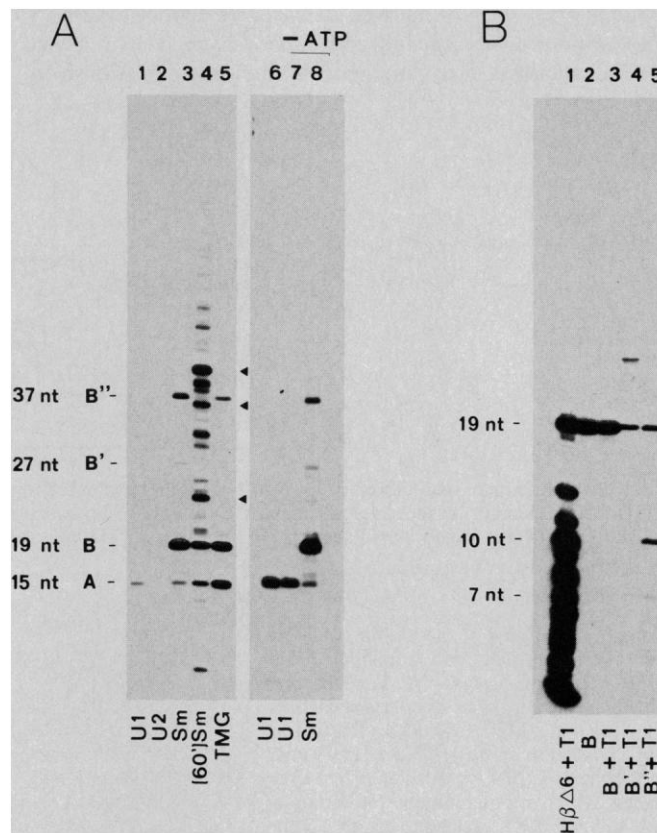

Fig. 1. Binding and protection analyses of a human beta-globin pre-mRNA transcript (HβΔ6) in a splicing extract. (A) [α-^{32}P]CTP-labeled RNA substrate (3 ng, 1 × 10^6 cpm Cerenkov) (20), 10 μl of antiserum, and 5 μl of T1 RNase (29 unit/μl) were added in that order to 24-μl samples of splicing mixture (10, 19) containing 15 μl of nuclear extract, prepared as in (33), and the mixtures were maintained at 0°C for 30 minutes. In a separate experiment (lanes 6 to 8), two aliquots were prepared without the addition of ATP or creatine phosphate (lanes 7 and 8). Immunoprecipitation of T1 resistant fragments was as described (10) with anti-(U1)RNP (lanes 1, 6, and 7), anti-(U2)RNP (lane 2), anti-Sm (lanes 3 and 8) or anti-TMG (lane 5) (21). The sizes of the protected fragments were estimated by comparison with DNA molecular weight markers. The efficiencies with which anti-(U1)RNP, anti-Sm, and anti-TMG precipitate the various splice site fragments differ. Lane 4 corresponds to an assay like that in lane 3 except that the splicing mixture was incubated for 60 minutes at 30°C with the labeled RNA before addition of T1 RNase and anti-Sm. The bands identified by arrowheads have been analyzed previously (10); the lowest arrowhead indicates the C fragment (10), which we had estimated to be 27 nt but now find to be slightly shorter (about 23 nt). This fragment contains the β fragment and GCUG, indicating that at 60 minutes the 3' splice site is protected from T1 RNase. (B) The identity of protected fragments was determined by their migration on a 20 percent polyacrylamide–8M urea gel in TBE buffer (90 mM tris base, 90 mM boric acid, 2.5 mM EDTA) and compared to a complete T1 digest of HβΔ6 (lane 1). B, B', and B" all contain a specific 19-nt fragment that maps to the 3' splice site. This oligomer was the only component of the B fragment (lane 3); B' and B" (lanes 4 and 5) both contained an additional 7-nt fragment and B" an extra 10-nt fragment. The identity of the B fragment was confirmed by two-dimensional nucleotide fractionation and secondary analyses; those of B' and B" were deduced on the basis of gel migration. (The relative intensities of the T1 products making up B' and B" are consistent with their content of C residues.)

efficiency of formation of this intermediate (and of the corresponding excised 5′ exon) relative to the wild-type substrate (5). In the binding assay (Fig. 4, lane 11), immunoprecipitation of the 3′ splice site fragment was strongly reduced; low amounts of an 18-nt fragment (band B*), which lacked the last residue of the wild-type 19-nt fragment after T1 digestion, appeared. The β^{110} mutant is a substrate that has a G to A transition 21 nt upstream of the 3′ splice junction, creating an AG immediately upstream of the pyrimidine stretch. In vitro, this upstream AG is the one used almost exclusively as the 3′ splice site (5). In the binding assay, this mutant displayed the same immunoprecipitation pattern (lane 6) as the wild-type substrate (lane 3). None of the 3′ splice site mutations altered the binding of the 5′ splice site fragment (band A) by U1 snRNP's (compare lanes 4, 7, and 10 with lane 1).

The 3′ splice site binding component is highly resistant to micrococcal nuclease.

To analyze further the snRNP responsible for 3′ splice site binding, we first treated splicing extracts with increasing amounts of micrococcal nuclease. After inhibiting the nuclease with EGTA, splicing cofactors and the wild-type splicing substrate (H$\beta\Delta6$) were added and T1 RNase protection assays performed (Fig. 5). Although low levels of nuclease abolished binding to the 5′ splice site (band A) by U1 snRNP's (Fig. 5, lanes 1, 3, 5, 7, and 9), anti-Sm continued to immunoprecipitate the 3′ splice site fragment (band B). Moreover, binding persisted even after extensive nuclease degradation (Fig. 5, lanes 2, 4, 6, 8, and 10). Thus, it seemed that either the RNA moiety of the snRNP involved in 3′ splice site binding is highly resistant to micrococcal nuclease or that the snRNP becomes only partially degraded and can still recognize the 3′ splice site region. Whatever the case, we deduce that the 5′ capped end cannot have been released by micrococcal nuclease digestion from

the snRNA in the snRNP responsible for 3′ splice site binding because band B remained immunoprecipitable by anti-TMG (lanes 11 to 15).

To determine whether one of the abundant Sm snRNP's has a 5′ end highly resistant to micrococcal nuclease, we performed the following experiment. A sonicated extract prepared from $^{32}PO_4$-labeled HeLa cells (23) was treated with various amounts of micrococcal nuclease, and snRNP's were immunoprecipitated with anti-Sm. The labeled snRNA's were then extracted, treated with T1 RNase, and immunoprecipitated with anti-TMG. Because T1 RNase digestion generates a different 2,2,7-trimethyl G-capped oligomer from each of the snRNA's U1, U2, U4, and U5 (Fig. 6A, lanes 1 to 4), this procedure assesses the relative micrococcal nuclease sensitivity of the 5′ ends of the RNA components in the various snRNP's. [Since U6 snRNA has no m_3G-cap (8), it could only be involved in 3′ splice site binding by
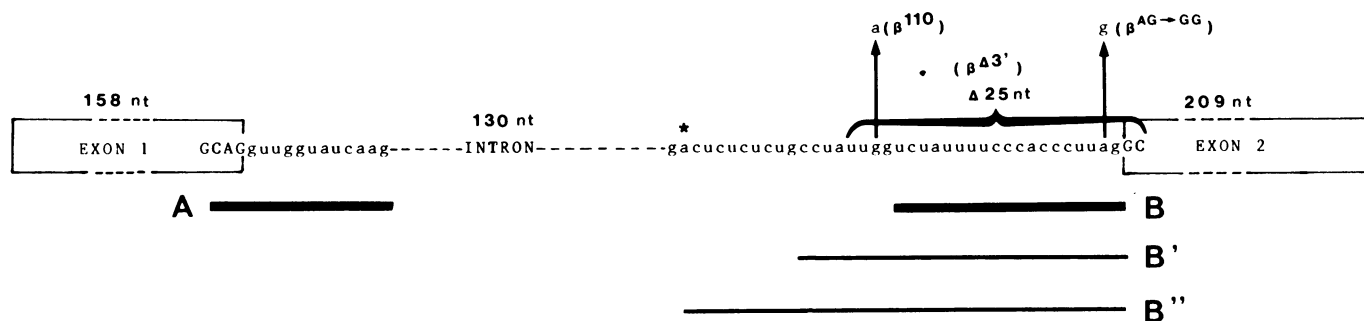

Fig. 2. Partial sequence of the wild-type human beta-globin transcript (H$\beta\Delta6$) and protected regions. In the sequence (10, 34) upper- and lower-case letters indicate exon and intron sequences, respectively. The asterisk over the adenosine residue indicates the site of branch formation (2). Protected regions are underlined and correspond to bands seen in Fig. 1, panel A. The mutations (5) used in Fig. 4 are also shown.

Fig. 3 (left). Binding and protection analyses of the globin substrate after prior treatment of the splicing extract with oligonucleotides complementary to U1 or U2 RNA and RNase H. The reaction mixtures contained 40 μl of nuclear extract, 20 μl of polyvinylalcohol (13 percent), 4 μl each of ATP (12.5 mM), MgCl$_2$ (80 mM) and creatine phosphate (0.5M), 2 μl each of RNasin (30 unit/μl), RNase H (1.9 unit/μl) and KCl (1M), and 20 μl of either 5′C (0.1 mg/ml) (lanes 5 to 8) or E15 (0.063 mg/ml) (lanes 9 to 12), which are oligonucleotides complementary to the 5′ end of U1 and U2 RNA, respectively (10). One reaction mixture (mock) contained 20 μl of H$_2$O instead of oligonucleotide (lanes 1 to 4). After incubation at 30°C for 1 hour, the H$\beta\Delta6$ transcript (4 μl, 2 × 10^6 cpm) was added to the three reaction mixtures on ice, and each was split into four 25-μl portions. T1 RNase and anti-(U1)RNP (lanes 1, 5, and 9), anti-(U2)RNP (lanes 2, 6, and 10), anti-Sm (lanes 3, 7, and 11) or anti-TMG (lanes 4, 8, and 12) was added and the immunoprecipitated fragments were isolated as described (10).   Fig. 4 (right). Binding and protection analyses of mutant beta-globin substrates. Wild-type (H$\beta\Delta6$) and three mutant transcripts (see Fig. 2) labeled with [α-^{32}P]GTP were assayed for snRNP binding at time zero, as in Fig. 1A. Protected fragments were immunoprecipitated with anti-(U1)RNP (lanes 1, 4, 7, and 10), anti-(U2)RNP (lanes 2, 5, and 8), or anti-Sm (lanes 3, 6, 9, and 11). The identity of the 18-nt fragment (B*) generated from $\beta^{AG\rightarrow GG}$ (lane 11) was confirmed by comparing its migration after T1 treatment with a T1 digest of the original transcript. In this experiment, 3′ splice site binding appears low relative to that in Fig. 1A since the transcripts were labeled with [α-^{32}P]GTP rather than [α-^{32}P]CTP.
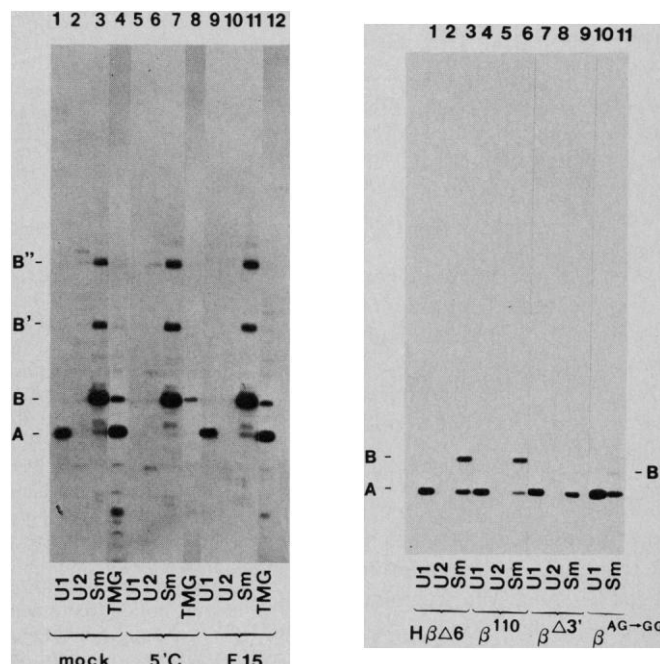
virtue of its association with U4 in snRNP particles (24).] The results of such an assay are shown in Fig. 6A, lanes 5 to 11. We see that both U2 and U4 RNA's have highly sensitive 5' termini since the recovery of their capped 5' ends dropped to less than 10 percent at the lowest concentration of micrococcal nuclease used (500 unit/ml). The capped oligomer derived from U1 RNA is also exposed; it disappeared after digestion with higher concentrations of micrococcal nuclease. Only U5 RNA has a very resistant 5' end: its capped oligonucleotide was retained and protected in the U5 snRNP even after the extract was treated with the highest concentration of nuclease. This resistance of the 5' end of U5 RNA parallels the remarkable nuclease insensitivity of the 3' splice site binding activity in the splicing extract (Fig. 5). However, because a cell extract prepared by sonication differs from a splicing extract, we repeated the binding and T1 protection experiments using a sonicated cell extract and the labeled beta-
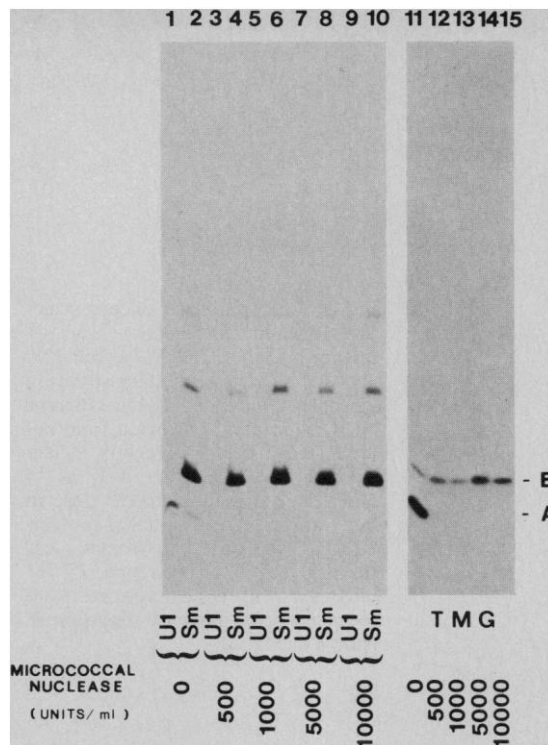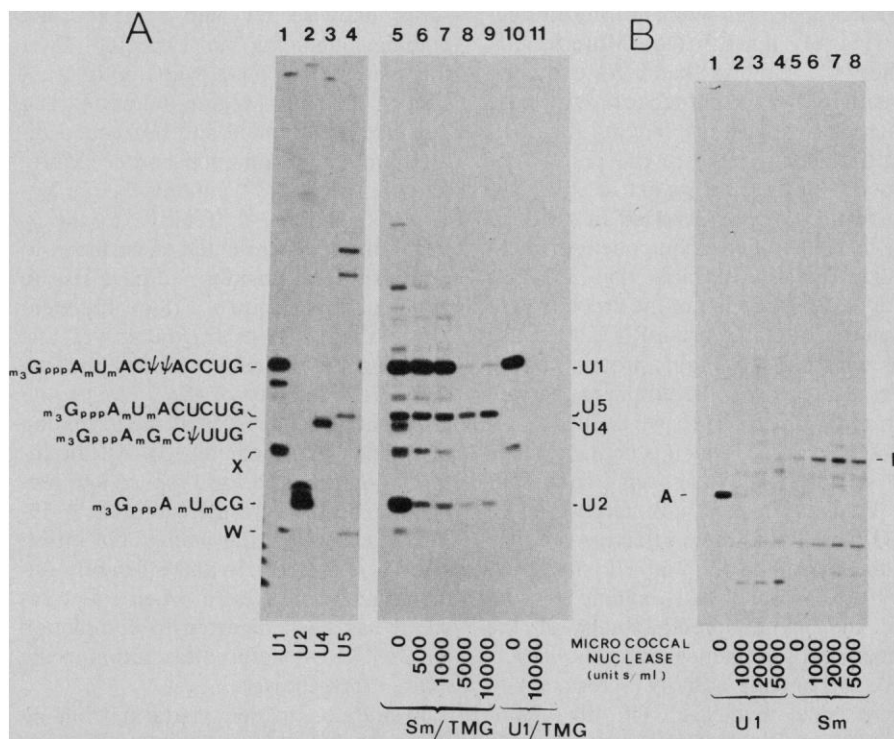

Fig. 5. Binding and protection analyses of the human globin pre-mRNA following micrococcal nuclease pretreatment of splicing extracts. Five 50 μl portions of nuclear extract (see legend to Fig. 1) containing 1 μl of CaCl$_2$ (50 mM) were treated for 15 minutes at 30°C with micrococcal nuclease at either 0 (lanes 1, 2, and 11), 500 (lanes 3, 4, and 12), 1000 (lanes 5, 6, and 13), 5000 (lanes 7, 8, and 14), or 10,000 (lanes 9, 10, and 15) unit/ml. After inactivation of the nuclease by adding 1 μl of EGTA (0.1M) to each portion, the splicing substrate ([α-^{32}P]CTP-labeled HβΔ6) and cofactors were added as in Fig. 1A (final volume 83 μl). Each reaction was split into three 25-μl aliquots that were then immunoprecipitated with either anti-(U1)RNP (lanes 1, 3, 5, 7, and 9), anti-Sm (lanes 2, 4, 6, 8, and 10), or anti-TMG (lanes 11 to 15) in the presence of T1 RNase, as described in Fig. 1A.

Fig. 6. Analysis of the capped 5' ends of snRNP's after treatment with micrococcal nuclease. (A) ^{32}P-labeled whole-cell extracts [prepared as in (23)] were treated with varying amounts of micrococcal nuclease (lanes 5 to 11). SnRNP's were immunoprecipitated with anti-Sm (lanes 5 to 9) or anti-(U1)RNP (lanes 10 and 11). After PCA (phenol:chloroform:isoamylalcohol, 50:49:1) extraction and ethanol precipitation with 50 μg of glycogen as carrier, the digested snRNA's were treated with T1 RNase (2.9 unit/μl) for 10 minutes at room temperature. Then, 50 μl of a slurry containing anti-TMG bound to Protein A Sepharose (2.5 mg) were added in NET-2 buffer [50 mM tris-HCl, pH 7.5, 150 mM NaCl, 0.05 percent (by volume) NP-40] and incubated for 30 minutes at 0°C. After three washings with NET-2 buffer, the pellet was extracted with PCA, precipitated with ethanol in the presence of carrier RNA, and loaded on a 25 percent polyacrylamide–8M urea gel. To obtain intact 5' end fragments as markers, a portion of the labeled extract was extracted with PCA and small RNA's were fractionated on a 10 percent polyacrylamide–7M urea gel. The uniformly labeled U1, U2, U4, and U5 snRNA's were cut out, eluted, and treated with T1 RNase before immunoprecipitation with anti-TMG (lanes 1 to 4). [T1 RNase will not cut 3' to the 2' O-methyl-G near the 5' end

of U4 RNA; and the T1 digestion of U5 snRNA (lane 4) was only partial. In addition, digestion of U1 and U5 RNA's produced a small amount of a band migrating as a 4-nt fragment (band W, lanes 1 and 4). We do not know why these products, as well as the U1-derived 7-nt fragment (band X), appeared.] Densitometer tracing of lanes 5 and 9 was used to quantitate the disappearance of the various 5' end oligomers: 80- and 60-fold decreases for the U1 and U2 RNA 5' end oligomers, respectively, and at least a 25-fold decrease for the 5' end T1 oligomer of U4 RNA were measured. The recovery of the capped oligomer from U5 snRNA suffered only a 2.5-fold diminution over a range of micrococcal nuclease concentrations from 0 to 10,000 unit/ml. (B) Binding and protection assays of the globin transcript with antibody-bound snRNP's. Fifty-μl portions (50-μl) of whole-cell extract (10^6 cells) (23) were incubated for 15 minutes at 30°C with micrococcal nuclease at 0 (lanes 1 and 5), 1000 (lanes 2 and 6), 2000 (lanes 3 and 7), and 5000 (lanes 4 and 8) unit/ml. After addition of EGTA, the treated extracts were incubated for 1 hour at 4°C with anti-(U1)RNP (lanes 1 to 4) or anti-Sm (lanes 5 to 8) already bound to Protein A Sepharose. After three 1-ml washings with BB1X buffer (0.1M tris-HCl, pH 8.0, 0.1M NaCl, 3 mM MgCl$_2$, 0.1 mM EDTA, 0.5 mM DTT, 0.05 percent NP-40), the resuspended pellet was incubated for another hour in 50 μl of BB1X buffer containing transcript (HβΔ6) DNA (500,000 cpm) and carrier DNA (sonicated salmon sperm) (0.5 μg/ml). After another washing 50 μl of BB1X buffer and 5 μl of T1 RNase (29 unit/μl) were added and the mixture was incubated for 15 minutes on ice. The pellet was washed three times, and the labeled RNA fragments were extracted with PCA, precipitated with ethanol and fractionated on a 15 percent polyacrylamide–8M urea gel in TBE buffer. Note several experimental differences from our previous unsuccessful attempts to detect 3' splice site binding by Sm snRNP's: a mouse beta-globin transcript was G-labeled and biochemically purified snRNP's were used (18).

Fig. 7 (RNA secondary structure):

```
5'   2,7,7G pppA U  A - U C U U A A C C C A  A U U U U U U G A - U    A OH  3'
      m3    m m                                           |         G - C
                      C - G  |                            90        C - G
                      U - A  80                                     C - G
                   C     G U   U                         100 -      U - A
                 U  G        C   C                                  U - A -110
                   G - C   A A                                    G - C
              10 - U - A       \70                             C       G
                   U - A                                     C   U U  G
                   U · G
                   C - G
                   U - A
                   C - G
                   U - A
              20   U · G - 60
                 A   C - G
               U   G A - U
             C  A     G   C
             G       A - U  C
             C  A U  A - Ψ
                30     A - U
                 U - A
                   C - G - 50
                   U - A
                   U - A
                   U - A
                 C        Ψ
               Gm       Cm
              C         A
               C      Ψ
              40  U  U  U
                       m
```
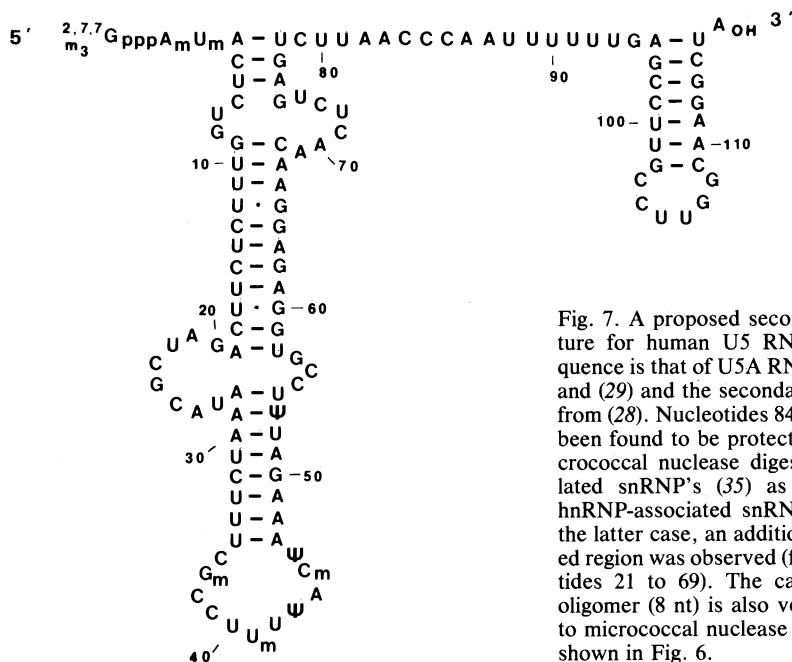
Fig. 7. A proposed secondary structure for human U5 RNA. The sequence is that of U5A RNA from (28) and (29) and the secondary structure from (28). Nucleotides 84 to 116 have been found to be protected from micrococcal nuclease digestion in isolated snRNP's (35) as well as in hnRNP-associated snRNP's (26). In the latter case, an additional protected region was observed (from nucleotides 21 to 69). The capped 5' T1 oligomer (8 nt) is also very resistant to micrococcal nuclease digestion as shown in Fig. 6.

globin pre-mRNA substrate. Even in such extracts both the 5' and 3' splice site fragments appeared in immunoprecipitates generated with anti-Sm antibodies (Fig. 6B, lanes 5 to 8). Moreover, 5' splice site binding (band A) exhibited sensitivity to micrococcal nuclease whereas 3' splice site binding (band B) did not, comparable to the results obtained with the splicing extract (Fig. 5).

Are U5 snRNP's involved in splicing? We have identified a component in mammalian splicing extracts that interacts with the 3' splice site of the first intron of human beta-globin pre-mRNA. Because this component contains protein or proteins that have Sm determinants as well as a m_3G-cap structure, we deduce that it is an Sm snRNP. That it is neither the U1 nor the U2 snRNP has been demonstrated by experiments with antibodies to m_3G (anti-TMG) after selective removal of the 5' end of U1 and U2 snRNA's; both the 3' splice site binding activity and the immunoprecipitability of the component are unchanged. Most important, the binding activity is resistant to micrococcal nuclease. Of the major snRNP's, only the snRNP containing U5 RNA has a comparable resistance to this nuclease, suggesting a role for the U5 snRNP in 3' splice site recognition.

In our binding assays, regions immunoprecipitated from the splicing reaction at time zero overlapped almost completely with the consensus sequence necessary for functioning of a 3' splice site. In the human beta-globin pre-mRNA, the major fragment was a single T1 RNase oligomer corresponding to the last 19 nucleotides of the intron. Compa-

rable experiments with anti-Sm and a splicing substrate containing the first intron of the adenovirus major late transcript likewise revealed 3' splice site binding, including an expected 18-nt fragment (3). Thus, with both transcripts, the bound region did not extend beyond the 3' splice site but ended directly at the intron-exon border. Moreover, in the $\beta^{AG \rightarrow GG}$ mutant, the nucleotide at position −1 (relative to the 3' junction) is not protected since the A to G transition at position −2 gave rise to an immunoprecipitated 18-nt fragment rather than a 19-nt fragment; yet this mutation dramatically reduced the immunoprecipitability of the 3' splice site fragment (Fig. 4) as well as the formation of splicing intermediates (5). All the information required for 3' splice site recognition seems to reside within the 3' 19-nt fragment from the beta-globin intron since it is selectively and efficiently immunoprecipitated even when the transcript has been digested to completion with T1 RNase before its addition to the extract (not shown).

In addition to demonstrating that an Sm snRNP is involved in 3' splice site binding, our data indicate that the interaction occurs very early during the splicing process. Neither added ATP nor incubation at 30°C is required. Many studies (5, 6, 13, 14, 16, 25) have suggested that 3' splice site recognition is essential for recognition of the intron branch point, a process which involves U2 snRNP's (10). The immunoprecipitation of longer 3' splice site fragments B' and B'' (which contain additional upstream sequences of the beta-globin transcript)

might reflect the association of other factors at or near the 3' splice site. Moreover, the observation that anti-(U1)RNP and anti-(U2)RNP immunoprecipitate a small amount of 3' splice site fragment (Fig. 1A, lanes 1 and 2) could indicate a weak association between U1 and U2 snRNP's and the 3' splice site binding component. Later in the splicing reaction, a stronger association with the U2 snRNP surely occurs since anti-(U2)RNP has been observed to immunoprecipitate 3' splice site-containing fragments in addition to the branch point region (10). On the other hand, binding at the 3' splice site region cannot by itself dictate which 3' splice junction will be used: we have observed for mutant β^{110} (which utilizes a new upstream AG) that both the protected 3' splice site at time 0 (Fig. 3) and the region protected by U2 snRNP's at later times (13) are the same as those in the wild-type transcript (10). Thus, it could be that snRNP binding to the 3' splice site acts only to position the U2 snRNP and other components required for the first step of the splicing reaction. Once the lariat intermediate has been formed, the splicing machinery might then search for a 3' splice site, preferring the first AG downstream from the branch site. However, a certain distance must be respected since an AG located too close to the branch point is ignored (7).

Most unexpected was our finding that the substrate binding activity of an Sm snRNP can be almost completely resistant to micrococcal nuclease. Nuclease resistance has often been interpreted to indicate a lack of nucleic acid involvement in a particular cellular process. Here, this unusual property has allowed us to eliminate three of the four abundant Sm snRNP's as candidates for the 3' splice site binding component. The capped 5' termini of snRNP's containing U1, U2, and U4 [which exists together with U6 in a single particle (24)] are cleaved off at relatively low concentrations of micrococcal nuclease and therefore cannot account for the unchanged immunoprecipitability of the 3' splice site fragment with anti-TMG.

By contrast, at least eight nucleotides at the 5' end of U5 RNA in the U5 snRNP are almost completely resistant to degradation (Fig. 5A). Therefore, the 3' splice site binding component must be either the U5 particle or a minor Sm snRNP with comparable nuclease resistance. High nuclease resistance of the U5 snRNP has been noted previously: when a heterogeneous nuclear ribonucleoprotein (hnRNP) preparation containing snRNP's was subjected to micrococcal

nuclease treatment at concentrations similiar to those used here, internal protected regions of U5 snRNA accounted for 80 percent of the entire molecule (26) (legend to Fig. 7). In addition, attempts to destroy U5 RNA in splicing extracts by site-directed cleavage using several different complementary deoxyoligonucleotides and RNase H have so far been unsuccessful (10, 27).

If U5 snRNP's are responsible for recognizing the 3' ends of the introns, the extreme conservation of U5 RNA in all organisms where it has so far been identified (28, 29) is not surprising. That U5 snRNP's are less abundant than U1 or U2 snRNP's (about one-fifth) is not expected, but could be explained if U5 snRNP's have higher affinity for their specific binding regions on pre-mRNA's than U1 or U2 snRNP's. Alternatively, U1 and U2 snRNP's may be required in several copies for the splicing of each intron or may be involved in some other nuclear process as well. Whether U5 snRNP's interact with 3' splice sites through RNA-RNA or RNA-protein contacts (or even via some other component) is unknown. The only region of U5 that exhibits extensive complementarity with the entire 3' splice site consensus sequence (residues 55 to 70) is probably sequestered in secondary structure. In contrast, there are several CU sequences (at position 39–40, 71–72, 79–80, and 104–105) that could pair with the conserved AG at the 3' splice junction. It seems likely that proteins of the U5 snRNP may play an important role in substrate recognition.

Work from several laboratories has demonstrated that pre-mRNA splicing occurs on a large complex (sedimenting at 50 to 60S in mammalian systems) termed the spliceosome (14, 30, 31). For spliceosome formation, both the 3' polypyrimidine stretch and an intact 5' splice site are required in the pre-mRNA; ATP is also essential. We now know that at least three Sm snRNP's are components of the mammalian spliceosome, those containing U1, U2, and probably U5 RNA. This raises the question of whether the one remaining abundant Sm snRNP in mammalian cells, which contains both U4 and U6 RNA's, is also involved. In any case, it is likely that hnRNP proteins (32) are also components of the spliceosome. Although much remains to be learned about spliceosome

content and formation, it appears that multiple components assemble onto the pre-mRNA according to an ordered pathway before the covalent reactions of the splicing process occur. Such a series of steps is reminiscent of initiation complex formation during the start of protein synthesis. In splicing, however, more ribonucleoprotein subunits (here snRNP's) are involved than in ribosomes, perhaps because several nonadjacent sites on the pre-mRNA must be recognized and brought together before transformation of the supercomplex into an entity with enzymatic activity.

References and Notes

1. S. M. Mount, *Nucleic Acids Res.* **10**, 459 (1982); J. H. Rogers, *Int. Rev. Cytol.* **93**, 188 (1985).
2. B. Ruskin, A. R. Krainer, T. Maniatis, M. R. Green, *Cell* **38**, 317 (1984).
3. R. A. Padgett, M. M. Konarska, P. J. Grabowski, S. F. Hardy, P. A. Sharp, *Science* **225**, 898 (1984).
4. M. M. Konarska, P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Nature (London)* **313**, 552 (1985); S. Zeitlin and A. Efstratiadis, *Cell* **39**, 589 (1984).
5. R. Reed and T. Maniatis, *Cell* **41**, 95 (1985).
6. B. Ruskin and M. R. Green, *Nature (London)* **317**, 732 (1985).
7. B. Ruskin et al., *Cell* **41**, 183 (1985).
8. J. A. Steitz et al., *Cold Spring Harbor Symp. Quant. Biol.* **47**, 893 (1983); H. Busch. R. Reddy, L. Rothblum, Y. C. Choi, *Annu. Rev. Biochem.* **51**, 617 (1982); R. Reddy, D. Henning, H. Busch, *J. Biol. Chem.* **260**, 10930 (1985).
9. A. Krämer et al., *Cell* **38**, 299 (1984).
10. D. L. Black, B. Chabot, J. A. Steitz, ibid. **42**, 737 (1985).
11. A. R. Krainer and T. Maniatis, ibid., p. 725.
12. V. W. Yang, M. R. Lerner, J. A. Steitz, S. J. Flint, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1371 (1981); R. A. Padgett, S. M. Mount, J. A. Steitz, P. A. Sharp, *Cell* **35**, 101 (1983); I. Bozzoni et al., *J. Mol. Biol.* **180**, 1173 (1984); A. Fradin et al., *Cell* **37**, 927 (1984).
13. B. Chabot, unpublished observations.
14. D. Frendewey and W. Keller, *Cell* **42**, 355 (1985).
15. M. R. Lerner, J. A. Boyle, S. M. Mount, S. L. Wolin, J. A. Steitz, *Nature (London)* **283**, 220 (1980); J. Rogers and R. Wall, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1877 (1980).
16. E. B. Keller and W. A. Noon, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7417 (1984).
17. K. Tatei et al., ibid., p. 6281.
18. S. M. Mount, I. Pettersson, M. Hinterberger, A. Karmas, J. A. Steitz, *Cell* **33**, 509 (1983).
19. A. R. Krainer, T. Maniatis, B. Ruskin, M. R. Green, ibid. **36**, 493 (1984).
20. All splicing substrates were synthesized from SP6 recombinant plasmids described previously (5, 19). The globin recombinants were cut with Bst I (an isoschizomer of Bam HI, provided by C. Joyce). Transcription reactions were performed as described (10) with 250 μCi of [α−^{32}P]GTP or −CTP.
21. The anti-(U1)RNP (Ag), anti-(U2)RNP (Ya), and anti-Sm (Y12) were used previously (10, 23). The preparation of anti-2,2,7-trimethylguanosine and of anti-7-methylguanosine has been described [L. Rainen and B. D. Stollar, *Nucleic Acids Res.* **5**, 4877 (1978); R. D. Meredith and B. F. Erlanger, ibid. **6**, 2179 (1979); R. Lührmann et al., ibid. **10**, 7103 (1982); J. H. Smith and G. L. Eliceiri, *J. Biol. Chem.* **258**, 4636 (1983)]. Our preparation of anti-TMG was similar, except that the synthesis of the nucleoside was as follows. 2-,2-dimethylguanosine was prepared from guanosine by two cycles of thiomethylation and Raney nickel desulfurization [P. K. Bridson and C. B. Reese, *Bioorg. Chem.* **8**, 339 (1979); O. Kemal and C. B. Reese, *Synthesis*

(1980), p. 1025]. The thiomethylation was carried out in aqueous solution, as opposed to ethanol-water, with mercaptoacetic acid substituting for both acetic acid and *p*-thiocresol. After refluxing for 1 hour, the *N*-carboxymethylthiomethyl derivative crystallized upon cooling to room temperature. The desulfurization was carried out in methanol on the diisopropylethylammonium salt. The dimethyl derivative was purified from the monomethyl contaminant by elution from Dowex-1 resin in acetate form with an ammonium acetate solution (pH 0.5). Quaternization of the N-7 position was achieved with methyl iodide in dimethyl sulfoxide and the 2,2,7-trimethylguanosine was chromatographically isolated on phosphocellulose [A. G. Saponara and M. D. Enger, *Nature (London)* **223**, 1365 (1969)]. Periodate oxidation of TMG was carried out on ice in a 50 m*M* phosphate solution at pH 7.0. A phosphate buffered solution of keyhole limpet hemocyanin (5 mg/ml) was added, followed by a solution of sodium cyanoborohydride to yield 0.2 percent (weight to volume) final concentration [R. F. Borch and A. I. Hassid, *J. Org. Chem.* **37**, 1673 (1972)]. After 1 hour, the conjugated protein sample was dialyzed against phosphate buffered saline. The nucleoside protein conjugate was emulsified with an equal volume of Freund's complete adjuvant for the first injection and with incomplete adjuvant thereafter. A New Zealand White rabbit was inoculated subcutaneously with 1 to 5 mg of protein conjugate on days 0, 21, 35, and 42 and approximately every 28 days thereafter. Blood was drawn from the rabbit before the first injection and on days 28, 42, and 49, and 7 days after each subsequent boost. The titer of the serum increased until day 49. The sera were screened with sonicates of ^{32}PO$_4$-labeled HeLa cells (23) and showed a pattern of immunoprecipitation including U1, U2, U3, U4, U5, and U6 snRNA's similar to that described by Bringmann et al. [*J. Biol. Chem.* **258**, 2475 (1983)].
22. N. Hernandez and W. Keller, *Cell*, **35**, 89 (1983); S. F. Hardy, P. S. Grabowski, R. A. Padgett, P. A. Sharp, *Nature (London)* **308**, 375 (1984).
23. T. Mimori, M. Hinterberger, I. Pettersson, J. A. Steitz, *J. Biol. Chem.* **259**, 560 (1984).
24. C. Hashimoto and J. A. Steitz, *Nucleic Acids Res.* **12**, 3283 (1984); P. Bringmann et al., *EMBO J.* **3**, 1357 (1984); J. Rinke, B. Appel, M. Digweed R. Lührmann, *J. Mol. Biol.*, **185**, 721 (1985).
25. B. Ruskin and M. R. Green, *Cell* **43**, 131 (1985).
26. J. Sri-Widada, J.-P. Liautard, C. Brunel, P. Jeanteur, *Nucleic Acids Res.* **11**, 6631 (1983).
27. D. Black, unpublished observation.
28. A. Krol, H. Gallinaro, E. Lazar, M. Jacob, C. Branlant, *Nucleic Acids Res.* **9**, 769 (1981).
29. R. Reddy, ibid. **13** (suppl.), r155 (1985).
30. P. J. Grabowski, S. R. Seiler, P. A. Sharp, *Cell* **42**, 345 (1985).
31. E. Brody and J. Abelson, *Science* **228**, 963 (1985).
32. G. Dreyfuss, Y. D. Choi, S. A. Adam, *Mol. Cell. Biol.* **4**, 1104 (1984).
33. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, *Nucleic. Acids Res.* **11**, 6631 (1983); N. Heintz and R. G. Roeder, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1475 (1984).
34. R. M. Lawn, A. Efstratiadis, C. O'Connell, T. Maniatis, *Cell* **21**, 647 (1980).
35. J.-P. Liautard, J. Sri-Widada, C. Brunel, P. Jeanteur, *J. Mol. Biol.* **162**, 623 (1982).
36. We thank Tsuneyo Mimori for the anti-(U2) RNP serum; Adrian Krainer, Robin Reed, and Tom Maniatis for the HβΔ6, β$^{AG \rightarrow GG}$, β^{A3} and β^{110} clones; David Solnick and Tom Maniatis for critically reading the manuscript; Patricia Dinatale, Donna Villano, and Nancy Batter for help in preparing the manuscript; Michael Lerner and Susan Cotmore for help in preparing the anti-TMG serum; and members of the Steitz laboratory for advice. F. M. Richards generously provided laboratory space and materials for the synthesis of 2,2,7-trimethylguanosine. Supported by National Science Foundation grants PCM-8305203 (D.M.L.) and PCM-8215962 (J.A.S.), National Institutes of Health grant GM-26154 (J.A.S.), and a G.E. Foundation Fellowship (D.L.B.)