Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 2642–2646, March 1995
Biochemistry

Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5′ splice sites

ALAN M. ZAHLER* AND MARK B. ROTH

Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104

Communicated by Harry F. Noller, University of California, Santa Cruz, CA, December 27, 1994

ABSTRACT **Alternative splicing of precursor messenger RNAs (pre-mRNAs) is an important mechanism for the regulation of gene expression. The members of the SR protein family of pre-mRNA splicing factors have distinct functions in promoting alternative splice site usage. Here we show that SR proteins are required for the first step of spliceosome assembly, interaction of the U1 small nuclear ribonucleoprotein complex (U1 snRNP) with the 5′ splice site of the pre-mRNA. Further, we find that individual SR proteins have distinct abilities to promote interaction of U1 snRNP with alternative 5′ splice junctions. These results suggest that SR proteins direct 5′ splice site selection by regulation of U1 snRNP assembly onto the pre-mRNA.**

Most metazoan precursor messenger RNAs (pre-mRNAs) contain intervening sequences called introns, which are removed prior to export of the RNA from the nucleus (reviewed in ref. 1). Splicing of an intron from pre-mRNA involves two sequential transesterification reactions that are carried out by a set of macromolecules including several small nuclear RNA protein complexes (snRNPs) that are known to assemble on the pre-mRNA in an ordered pathway to form a spliceosome (2). Interactions between the U snRNAs (U1, U2, U5, and U4/U6) and the pre-mRNA in the spliceosome are thought to form the active site for the splicing reactions (3–5).

Spliceosome assembly onto the nascent pre-mRNA transcript is mediated by RNA·RNA interactions between the U snRNAs and the pre-mRNA (reviewed in ref. 2). Metazoan introns can be many kilobases in length and there can be dozens of introns in a precursor message. The consensus sequences for the RNA·RNA interactions between the pre-mRNA and the U snRNAs in metazoans are insufficient to explain the fidelity with which pre-mRNA splicing occurs, indicating that additional mechanisms are required for the fidelity of spliceosome assembly. Several non-snRNP splicing factors have been identified. These include U2 auxiliary factor (U2AF), which binds to the intron near the 3′ splice junction and is required for U2 snRNP binding to the intron branch point (6, 7). Another non-snRNP splicing factor is a family of at least six evolutionarily conserved proteins called the SR protein family (8–13). Any one of the six SR protein family members can be added to complement an otherwise splicing-deficient extract for splicing of simple pre-mRNAs containing single 3′ and 5′ splice sites (8). Individual SR proteins have distinguishable functions when assayed with pre-mRNAs that can be alternatively spliced; specific SR proteins allow for the preferential use of different 5′ splice sites within the same precursor (14). For example, in the simian virus 40 (SV40) early message there are two alternative 5′ splice sites. When a pre-mRNA substrate containing this region of SV40 is spliced in an extract containing only one SR protein, SRp30b promotes splicing at the proximal small tumor antigen (t) 5′ splice site, while SRp40 and SRp55 promote splicing at the distal large tumor antigen (T) splice site (14).

To begin to understand how SR proteins specify 5′ splice site usage, we examined the role of SR proteins in the earliest events of spliceosome assembly. We used a method in which a single residue of 4-thiouridine (4-thioU) positioned in the pre-mRNA substrate is used to detect interactions between the pre-mRNA and various U snRNAs (U1, U2, U5, and U6) during spliceosome assembly (15, 16). When activated by long-wave ultraviolet (UV) light, 4-thioU can form crosslinks to RNA and protein (reviewed in ref. 17). When a 4-thioU residue is placed 2 nt upstream of a 5′ splice site, interactions between this nucleotide and U1 snRNA that occur prior to the first step of splicing have been detected (15). In this report, we use this 4-thioU assay to detect interactions between the pre-mRNA and U1 snRNP at alternative 5′ splice sites in the presence of SR proteins that specify usage of the different splice sites.

MATERIALS AND METHODS

Construction of Pre-mRNA Substrates. Substitution of uridine for adenosine 2 nt 5′ of both 5′ splice sites of the SV40 pre-mRNA substrate was done to make an SV40 alternative splicing pre-mRNA substrate into which 4-thioU could be incorporated. This was accomplished by moving the entire 1235-bp *Hin*dIII fragment of pSVi66 (18) from the pSP64 vector into the *Hin*dIII site of pBluescript KS+. Because of the change of plasmid vectors, there are an additional 50 nt of vector sequence at the 5′ end of the RNA relative to the transcript made from pSVi66. The plasmid template is linearized with *Sty* I prior to T3 transcription. A uridine has been substituted for an adenosine at a position 2 nt 5′ of each of the 5′ splice sites. This was accomplished by two stepwise site-directed mutageneses of the plasmid (19).

To determine if sequences around the 5′ splice junctions are important for 5′ splice site specificity by SR proteins, a pre-mRNA substrate was constructed in which the sequences of the two alternative 5′ splice sites of the pSVi66 pre-mRNA substrate have been substituted for each other. The regions that were switched consisted of 50 nt that were centered at each of the 5′ splice sites (arbitrarily named A and B). In the original pSVi66 construct the order of the two 50-nt 5′ splice site regions went 5′ → 3′, A then B. In the switched construct, the order of these two 50-nt regions is B then A. As a first step in constructing the template for the switched RNA, the pSVi66 insert was separated into two halves by cutting with the restriction endonuclease *Nde* I, and these halves, each containing one of the alternative 5′ splice sites, were then subcloned into pBluescript KS+. Site-directed mutagenesis was then done to change the sequences of each of the 5′ splice sites.

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: snRNP, small nuclear ribonucleoprotein; pre-mRNA, precursor mRNA; 4-thioU, 4-thiouridine; SV40, simian virus 40; T, large tumor; t, small tumor.

*To whom reprint requests should be sent at the present address: Department of Biology, Sinsheimer Laboratories, University of California, Santa Cruz, CA 95064.

Both sequences to be changed were 50 nt in length, 25 bases on either side of the 5′ splice site. These modifications were performed using 90-base oligodeoxynucleotides containing the 50-nt sequence to be substituted flanked by 20 nt of sequence on each side that are already present in the clone. After the two 50-nt regions were switched, the two subclones were cut out of the plasmids with *Hin*dIII and *Nde* I, ligated together, and cloned as a *Hin*dIII fragment into pBluescript KS+, such that transcription would be under the control of the T3 promoter.

Purification of SR Proteins. SR proteins were purified from calf thymus as described and each is at a concentration of 400 ng/μl (14). This preparation involves separating the SR proteins on preparative SDS/polyacrylamide gels followed by recovery and renaturation of the individual proteins. We have previously shown that this purification procedure does not affect the specific activity of the proteins or the specificity of these SR proteins for alternative 5′ splice sites (14). SR protein concentration was determined by comparing the intensity of Coomassie staining of an SDS/PAGE gel to a bovine serum albumin standard.

Synthesis of 4-ThioU-Containing Pre-mRNA Substrates. The basic protocols for synthesis of RNAs with site-specific incorporation of 4-thioU nucleotides have been described (15, 20). Four DNA templates were required, two for each half of the two 4-thioU-containing substrates. To make the 5′ templates, the T3 primer 5′-ATTAACCCTCACTAAG-3′ with either 5′-CAGTTGCATCCCAGAAG-3′ or 5′-TAGAGCTTTAAATCTC-3′ for the large T 4-thioU and small t 4-thioU 5′ RNAs, respectively, were used in a PCR with the modified pSVi66 plasmid as a substrate. To make the 3′ templates, PCRs were done using two T7 promoter-containing oligodeoxynucleotides, either 5′-TAATACGACTCACTATAGGTATTTGCTTCTTCCTTAA-3′ for the large T 4-thioU 3′ template or 5′-TAATACGACTCACTATAGGTAAATATAAAATTTTTAAG-3′ for the small t 4-thioU 3′ template, with the oligodeoxynucleotide 5′-CTTGGGGTCTTCTACC-3′. The PCRs were done on a large scale, 1.6-ml reaction mixtures for each, to generate sufficient quantities of each of the four DNAs to serve as templates for transcription. All PCR products were purified from agarose gels. RNAs were synthesized from these templates in 100-μl transcription reaction mixtures containing 6 μg of DNA template, 40 mM Tris·HCl (pH 8.0), 8 mM $MgCl_2$, 4 mM spermidine, 10 mM dithiothreitol (DTT), 0.8 unit of RNase inhibitor per μl (Boehringer Mannheim), and 2 mM (each) ATP, CTP, GTP, and UTP. Additionally, 5′ RNA transcription reaction mixtures contained 3 mM 7-methyl-GpppG (Pharmacia) and 4 units of T3 RNA polymerase per μl (Boehringer Mannheim), while 3′ RNA transcription reaction mixtures also contained 4 mM 4-thiouridylyl-3′,5′-guanosine (Sigma) and 5 units of T7 RNA polymerase per μl (Boehringer Mannheim). Reaction mixtures were incubated for 90 min at 37°C. At this time, 5 μl of DNase I (10 units/μl; Boehringer Mannheim) was added to each reaction and the reaction mixtures were incubated an additional 10 min at 37°C. Transcription reactions were then phenol and chloroform extracted and nucleic acids were ethanol precipitated. The RNA was then separated on a 5% polyacrylamide (0.25% bis) denaturing gel to purify full-length product. The 4-thioU-containing 3′ RNAs were treated with polynucleotide kinase in a 10-μl reaction containing 2 μM RNA, 50 mM Tris·HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 7 μM [γ-^{32}P]ATP (7000 Ci/mmol; 1 Ci = 37 GBq), and 10 units of T4 polynucleotide kinase (Boehringer Mannheim). Reactions were incubated at 37°C for 30 min at which time 10 volumes of formamide were added to stop the reaction. The "kinased" RNAs were then purified on a denaturing polyacrylamide gel. The 5′ and 3′ half RNAs were then ligated together using the protocol of Moore and Sharp (20) in a 10-μl reaction. The mixture contained the 3′ RNA at 1 μM, the bridge oligonucleotide (5′-GGAAGAAGCAAATACCACAGTTGCATCCCAGAAG-3′ for large T 4-thioU bridge or 5′-AAATTTTATATTTACCATAGAGCTTTAAATCTCT-3′ for the small t 4-thioU bridge) at 2 μM, the 5′ RNA at 3 μM, 50 mM Tris·HCl (pH 7.6), 10 mM $MgCl_2$, and 10 mM DTT. This was heated to 75°C for 2 min and cooled at room temperature for 3 min to promote annealing. Then ATP was added to 1 mM and 10 units of T4 DNA ligase (United States Biochemical) was added. Ligations were carried out at 37°C for 30 min. Reactions were stopped by addition of 10 volumes of formamide, and ligated RNAs were electrophoretically purified. Approximate ligation efficiencies relative to the amount of 3′ RNA in the reactions were 5% for the small t 4-thioU RNA substrate and 20% for the large T 4-thioU RNA substrate. Approximately 1×10^5 cpm of pre-mRNA substrate was added to each 25-μl *in vitro* splicing reaction.

4-ThioU Crosslinking Assays. Standard *in vitro* splicing reaction mixtures containing splicing-deficient extract complemented with SR proteins and a 4-thioU-containing pre-mRNA substrate were assembled. After a 10-min incubation at 30°C the reaction mixtures were pipetted in 12.5-μl droplets onto a Parafilm-covered piece of aluminum chilled on ice. A long-wave UV source with no filter (model UVL-56; Ultraviolet Products, San Gabriel, CA) was then set on top of this so that the bulb was 2 cm from the samples. Samples were irradiated in this manner for 6 min and the reaction mixtures were then recovered and digested with proteinase K as described (8). The nucleic acids were purified and separated on a 4.5% polyacrylamide (0.225% bis) denaturing gel in TBE (90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3) for 9 hr at 1700 V. Bands were visualized with a Molecular Dynamics PhosphorImager (21).

RNase H Digestions. RNase H digests of purified crosslinked RNA were done in 10-μl volumes as described by Wyatt *et al.* (15). The reaction mixtures contained 40 mM Tris·HCl (pH 7.6), 50 mM KCl, 5 mM $MgCl_2$, 0.16 μg of oligodeoxynucleotide per μl, and 0.3 unit of RNase H per μl (Boehringer Mannheim). Reaction mixtures were incubated at 30°C for 60 min and the nucleic acids were purified and electrophoretically separated. An oligodeoxynucleotide complementary to nt 64–75 of human U1 snRNA (5′-CGGAGTGCAATG-3′) was used to confirm that crosslink bands contain U1 snRNA.

***In Vitro* Splicing Reactions.** *In vitro* splicing reactions were done in a volume of 25 μl. Preparation of HeLa S100 splicing-deficient extracts, preparation of ^{32}P-labeled substrate pre-mRNA, *in vitro* splicing conditions, and urea/PAGE analysis were previously described (8).

RESULTS

SR Proteins Are Required for U1 snRNP Interaction with the pre-mRNA. We used 4-thioU crosslinking to examine early steps of spliceosome assembly on a pre-mRNA substrate containing two alternative 5′ splice sites. The substrate used for these experiments contains a region of the SV40 early pre-mRNA with two alternative 5′ splice sites and a single 3′ splice site (18). Use of the 5′ splice site proximal or distal to the 3′ splice site leads to an mRNA encoding small t or large T antigen, respectively. We made two different pre-mRNA substrates: one with a 4-thioU residue 2 nt upstream of the large T 5′ splice site and another with a corresponding substitution at the small t 5′ splice site. To do this we first had to mutate the DNA template for this substrate (18) such that uridine is substituted for adenosine 2 nt upstream of both of the 5′ splice sites (Fig. 1*A*). These mutations had no effect relative to the wild-type pSVi66 sequence on the ability of SR proteins to influence 5′ splice site selection *in vitro* (Fig. 1*B*). As with wild-type pre-mRNA, SRp30b specifies use of the

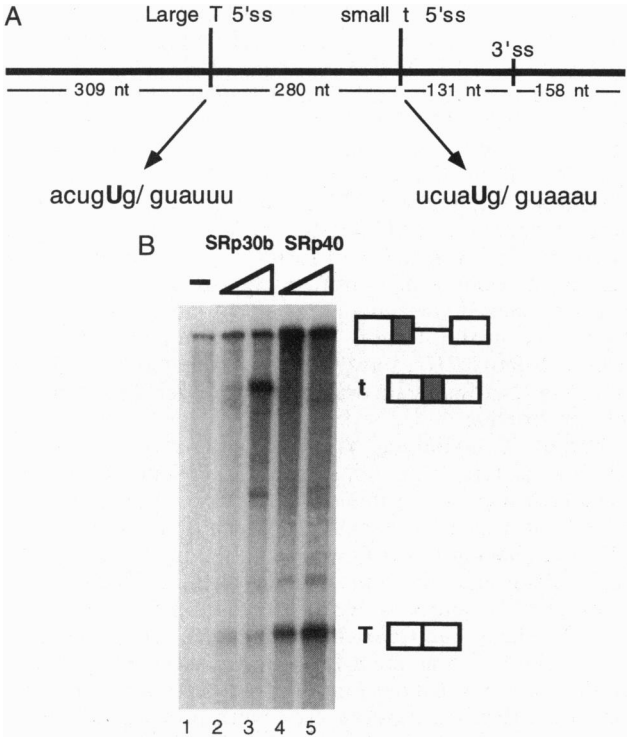

FIG. 1. Modified SV40 (pSVi66) alternative splicing substrate. (*A*) A modified SV40 substrate pre-mRNA was constructed to test if substitution of uridine for adenosine at a position 2 nt 5′ of both 5′ splice sites would alter the specificity of SR proteins in splice site selection for this substrate *in vitro*. The sequences of the 5′ splice sites and the location of the 4-thioU substitutions (uppercase) used in later experiments are indicated. The number of nucleotides separating the splice sites from the ends of the pre-mRNA and from each other are indicated. (*B*) To determine if different SR proteins have distinct specificities in promoting alternative 5′ splice site usage on the modified pSVi66 pre-mRNA substrate, the substrate RNA was incubated under pre-mRNA splicing conditions in a HeLa cell S100 splicing-deficient extract to which different SR proteins were added. Lane 1, control *in vitro* splicing reaction mixture containing HeLa cell S100 extract and ^{32}P-labeled pre-mRNA splicing substrate. Products from similar reactions that also contained 1 or 2 μl of SRp30b (lanes 2 and 3) or 1 or 2 μl of SRp40 (lanes 4 and 5) are shown. The reaction products were separated on a 5% polyacrylamide (0.25% bis) denaturing gel and visualized by autoradiography. Spliced products run at the exact same positions as the products of wild-type pSVi66 splicing on these gels, indicating that the substitution of uridine for adenosine at the -2 position of these 5′ splice sites does not affect their usage (data not shown). Spliced products are indicated with exons shown as boxes, alternatively spliced exons as hatched boxes, and introns as lines.

small t splice site (lanes 2 and 3), while SRp40 specifies use of the large T splice site (lanes 4 and 5).

Using this substrate RNA, we were able to show that interactions between U1 snRNA and the pre-mRNA are SR protein dependent. SV40 pre-mRNAs with 4-thioU incorporated 2 nt before the small t or large T 5′ splice site were incubated in the splicing-deficient extract alone (Fig. 2, lanes 1 and 5) or in the splicing-deficient extract with a mixture of SR proteins that specify splicing at both 5′ splice sites (Fig. 2, lanes 2 and 6). After 10 min the splicing reaction mixtures were crosslinked and the extracted RNAs were electrophoretically separated. In the presence of SR proteins, two anomalously migrating bands for each substrate are detectable (Fig. 2, lanes 2 and 6). During electrophoresis, these molecules migrated only 3–4 cm into the gel, while the pre-mRNA substrate migrated 20 cm into the gel. Formation of these products is dependent on the presence of 4-thioU in the pre-mRNA and on treatment of the reaction mixture with UV light (data not shown). In the presence of antisense U1 oligonucleotide and RNase H these

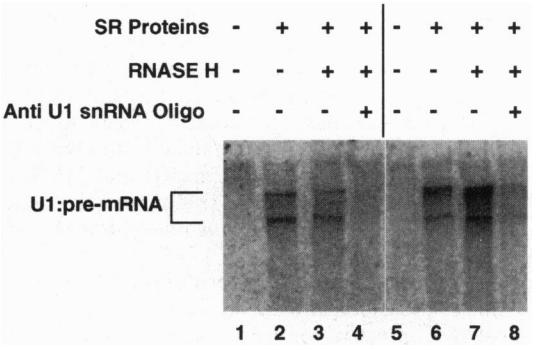

FIG. 2. SR proteins are required for interactions between the 5′ splice sites and U1 snRNA. Splicing reactions were performed using two different 4-thioU-containing substrate pre-mRNAs. Lanes 1–4 are derived from splicing reactions with a pre-mRNA substrate containing a [^{32}P]phosphate and a 4-thioU 2 nt 5′ of the small t 5′ splice site. Lanes 5–8 are derived from splicing reactions with a pre-mRNA substrate containing a [^{32}P]phosphate and a 4-thioU 2 nt 5′ of the large T 5′ splice site. Lanes 1, 2, 5, and 6, products from standard splicing reactions in the S100 splicing-deficient extract containing the indicated 4-thioU pre-mRNA substrate and either no SR proteins (lanes 1 and 5) or a mixture of SR proteins purified from calf thymus (lanes 2 and 6). After a 10-min incubation at 30°C the reaction mixtures were irradiated with long-wave UV light for 6 min. The nucleic acids were purified and separated on a 4.5% polyacrylamide (0.225% bis) denaturing gel in TBE for 9 hr at 1700 V. Bands were visualized with a Molecular Dynamics PhosphorImager; only the first 7 cm of the gel from the origin are shown. Lanes 3, 4, 7, and 8, products of RNase H digestions of RNA purified from crosslinking reactions identical to those shown in lanes 2 and 6. This RNA has been treated with RNase H in the presence of no oligodeoxynucleotide (lanes 3 and 7) or an oligodeoxynucleotide complementary to nt 64–75 of human U1 snRNA.

products are specifically degraded, indicating that they are formed by crosslinking of the labeled pre-mRNA to U1 snRNA (Fig. 2, lanes 3 and 4 and 7 and 8). In addition, these products are sensitive to treatment with RNase H in the presence of an oligodeoxynucleotide complementary to the pre-mRNA but are not sensitive to treatment with RNase H in the presence of an oligodeoxynucleotide complementary to U5 snRNA (data not shown). As seen for another substrate by Wyatt *et al.* (15), two U1 snRNA–pre-mRNA crosslinked species were detected. The two U1 snRNA–pre-mRNA crosslinked bands for each of the SV40 pre-mRNA substrates had very similar mobilities. The fact that the two 5′ splice sites are each ≈300 nt from an end of the substrate pre-mRNA could explain the similar mobilities of the crosslinks for both 4-thioU-containing substrates. These experiments indicate that SR proteins are required for binding of U1 snRNP to the 5′ splice site of the pre-mRNA and are in agreement with the observation that there is a cooperative effect of SR proteins on U1-snRNP binding to a 5′ splice site (22).

SR Proteins Have Distinct Specificities in Recruitment of U1 snRNP to Alternative 5′ Splice Sites. Having established an assay for interaction of U1 snRNP with the two different 5′ splice sites, we determined the pattern of U1 snRNP binding to both 5′ splice sites in the presence of different SR proteins. The two different 4-thioU-containing pre-mRNA substrates were incubated in the splicing-deficient extract with no SR proteins, increasing amounts of SRp30b (small t-specific), or increasing amounts of SRp40 (large T-specific). After 10 min, an aliquot was removed and irradiated (Fig. 3*A*). The remainder was incubated at 30°C for 2 hr and the corresponding products of splicing are shown in Fig. 3*B*. The results of several independent trials show that SRp40 allows significantly greater U1 snRNP binding to the site that is used (Fig. 3, lanes 12–14) than it does with the site that is not used (Fig. 3, lanes 5–7). In contrast, SRp30b promotes binding of U1 snRNP to

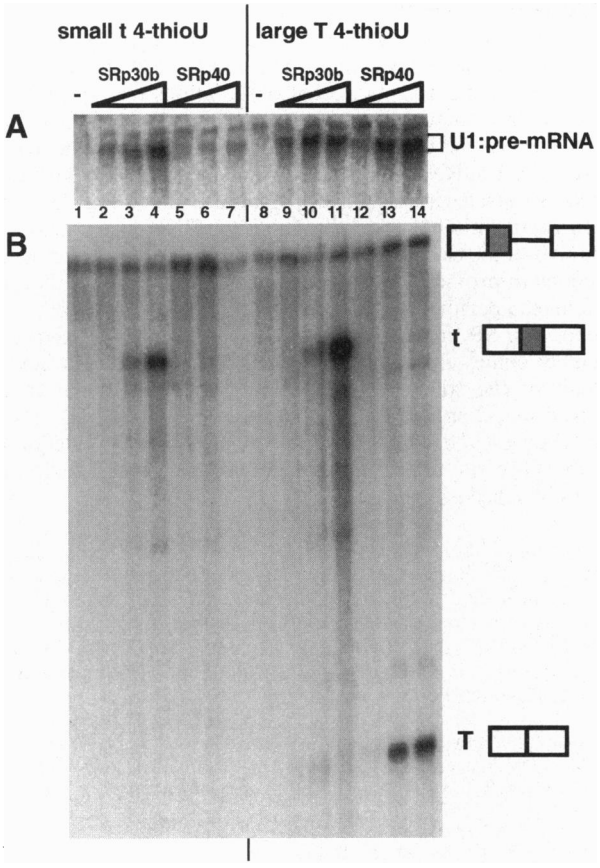

FIG. 3. SR proteins have different abilities to promote U1 snRNA recruitment to alternative 5′ splice sites. Pre-mRNA substrates contain the 4-thioU nucleotide 2 nt 5′ of the small t splice site (lanes 1–7) or 2 nt 5′ of the large T splice site (lanes 8–14). Standard 25-μl splicing reaction mixtures were made containing splicing-deficient extract alone (lanes 1 and 8), splicing-deficient extract with either 0.5, 1, or 2 μl of SRp30b (lanes 2, 3, and 4, respectively, and lanes 9, 10, and 11, respectively), or splicing-deficient extract with 0.5, 1, or 2 μl of SRp40 (lanes 5, 6, and 7, respectively, and lanes 12, 13, and 14, respectively). Reaction mixtures were incubated at 30°C. After 10 min of incubation, 12 μl of each reaction was removed and irradiated with 365-nm light for 6 min. The nucleic acids were then recovered from these samples and prepared for electrophoresis. The remaining reaction mixtures were incubated at 30°C for a total of 120 min and the nucleic acids were recovered and prepared for electrophoresis. All samples were run on a 4.5% denaturing polyacrylamide gel (0.225% bis) in TBE at 1700 V for 6.75 hr. (*A*) Molecular Dynamics PhosphorImager image of the first 5 cm of the gel from the origin for the UV light-irradiated samples. (*B*) Autoradiograph of the 120-min splicing reactions, showing a region of the gel containing the pre-mRNAs and the splicing products. The U1–pre-mRNA crosslink as well as the pre-mRNA and the two splice products are indicated. It is important to note that the large T splice product band cannot be detected in lanes 1–7 because the only [^{32}P]phosphate in the small t 4-thioU substrate pre-mRNA would be in the intron when the large T splice junction is used. Reaction intermediates containing these introns were detected in the SRp40 lanes as expected (data not shown).

both 5′ splice sites (Fig. 3, lanes 2–4 and 9–11), even though only the small t splice junction is used for splicing. SRp30b promoted U1 snRNP assembly on the small t 5′ splice site at least 5-fold better than SRp40, while both SR proteins promoted equal U1 snRNP assembly at the large T splice site. These results, summarized in Fig. 4, indicate that SRp40 and SRp30b have distinct abilities to recruit U1 snRNP to the two different 5′ splice sites. SRp40 may specify the splice site that is used at or before U1 recruitment, while splice site selection appears to occur at another step for SRp30b.

The 5′ Splice Site Sequences and SR Protein Specificity of Alternative Splice Site Usage. SR proteins have distinct abilities to promote U1 snRNP interactions with 5′ splice sites. These interactions occur by the base-pairing of complementary sequences in the 5′ end of the U1 snRNA to the 5′ splice site. We wondered whether the relative strengths of this RNA·RNA complementarity could explain SR protein function in U1 snRNP recruitment to the pre-mRNA. It has been proposed that the strength of complementarity of a 5′ splice site to the first 9 nt of U1 snRNA at the 5′ end (5′-CAG/GUAAGU-3′) could explain alternative 5′ splice site preference (23, 24). Because the small t site of SV40 has a greater complementarity to the 5′ end of U1 snRNA (six of nine possible matches versus five of nine at the large T splice junction), we were able to test the significance of the sequences at the 5′ splice sites by creating a pre-mRNA substrate in which 50-nt blocks centered around each of the two 5′ splice sites were exactly switched (Fig. 5*A*). This altered pre-mRNA shows the same pattern of splicing as the original pre-mRNA. SRp55, which has the same specificity of promoting usage of the large T splice site as SRp40 (14), was used in these experiments. SRp30b and SRp55, which promote usage of the proximal (Fig. 5*B*, lanes 2 and 3) and distal (Fig. 5*B*, lanes 4 and 5) 5′ splice sites, respectively, in the wild-type pre-mRNA, still show the same proximal (Fig. 5*B*, lanes 7 and 8) and distal (Fig. 5*B*, lanes 9 and 10) 5′ splice site preferences in the "switched" substrate. Thus, sequences surrounding the 5′ splice site, including the region of U1 complementarity, cannot account for the SR protein-dependent alternative splicing observed.

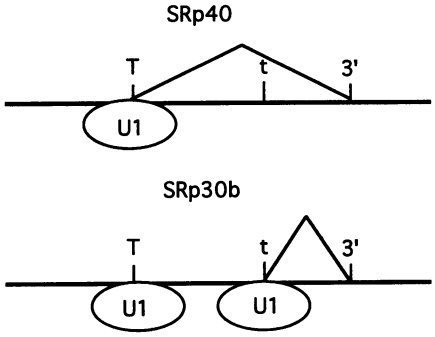

FIG. 4. Model for U1 snRNP recruitment to 5′ splice sites of SV40 substrate pre-mRNA by SR proteins. In the absence of SR proteins, no U1 snRNP is detected at the 5′ splice sites and no splicing occurs. In the presence of SRp40, U1 snRNP is recruited primarily to the large T splice site and splicing occurs at that site. In the presence of SRp30b, U1 snRNP is recruited to both 5′ splice sites, and splicing occurs at the small t splice site.

DISCUSSION

We have shown that SR proteins are required for interaction of U1 snRNP with the pre-mRNA and that individual SR proteins have distinct functions in directing U1 snRNP to bind to a specific region of the pre-mRNA. Different SR proteins may be able to bind the pre-mRNA at different sites based on pre-mRNA sequence or structure. Indeed, SR protein binding sites within exon sequences downstream of the splice sites have been detected in a number of genes (25–27). SR proteins can also interact with a protein component of the U1 snRNP, the U1 70k protein (22, 28). When an SR protein binds to the pre-mRNA, it may then bind or already be bound to U1 snRNP through the U1 70k protein. This interaction would bring U1 snRNP in proximity to a specific region of the pre-mRNA, thus increasing the local concentration of U1 snRNP and promoting binding to that region of the pre-mRNA. This correlates with the experiment shown in Fig. 5 and with what is known about mutations in the 5′ splice site of human β-globin that lead to β-thalessemia (23, 29). Mutations in the original 5′ splice site sequence lead to use of one of three adjacent

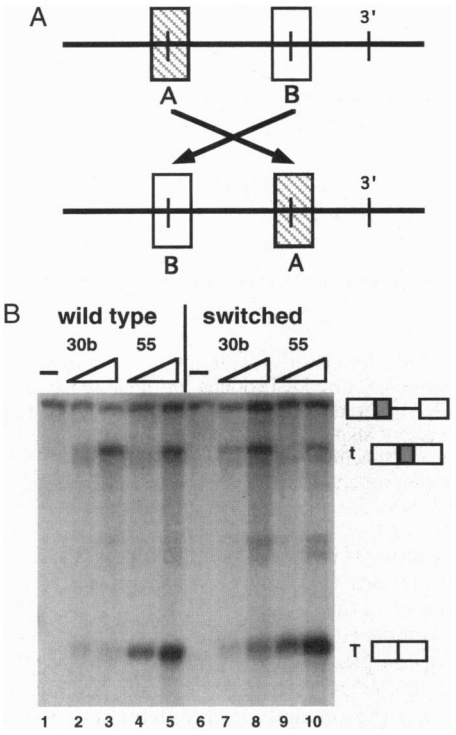

FIG. 5. The 5′ splice site sequences and SR protein specificity. (*A*) To determine if sequences around the 5′ splice sites are important for 5′ splice site specificity by SR proteins, a pre-mRNA substrate was constructed in which the sequences of the two alternative 5′ splice sites of the pSVi66 pre-mRNA substrate have been substituted for each other. The regions that were switched consisted of 50 nt that were centered at each of the 5′ splice sites (arbitrarily named A and B). In the original pSVi66 construct the order of the two 50-nt 5′ splice site regions went 5′ → 3′, A then B. In the switched construct, the order of these two 50-nt regions is B then A. (*B*) *In vitro* splicing of the "wild-type" and "switched" pSVi66 pre-mRNA substrates in the presence of different SR proteins. Splicing reactions contained either the wild-type pre-mRNA substrate pSVi66 (lanes 1–5) or the substrate "pSVi66 switched" in which 50 nt at each 5′ splice site have been substituted for one another (lanes 6–10). Reaction mixtures contained either splicing-deficient extract alone (lanes 1 and 6), splicing-deficient extract with 1 or 2 μl of SRp30b (lanes 2 and 3, respectively, and lanes 7 and 8, respectively), or splicing-deficient extract with 1 or 2 μl of SRp55 (lanes 4 and 5, respectively, and lanes 9 and 10, respectively). Reaction mixtures were incubated at 30°C for 2 hr and nucleic acids were separated on a 5% polyacrylamide (0.25% bis) denaturing gel in TBE for 4.5 hr at 1700 V. The positions of the pre-mRNA and the two alternative splice products are indicated.

"cryptic" 5′ splice junctions, which are located between 14 and 38 nt from the original 5′ splice site. These cryptic sites have weaker U1 snRNA consensus sequences than the original 5′ splice site. When the wild-type 5′ splice site is mutated, a nearby cryptic splice junction is chosen. This may be due to the strong influence of SR proteins to promote U1 snRNP binding and spliceosome assembly in that region of the pre-mRNA. Our results indicate that SR proteins bind to a site on the pre-mRNA outside of the 5′ splice site region. However, there is evidence that SR proteins bind to the 5′ splice site region (30, 31). Perhaps SR proteins interact with the pre-mRNA at the 5′ splice site and at another region, with the non-5′ splice site region being responsible for alternative splicing specificity.

Comparison of the relative levels of U1 recruitment to each of the two 5′ splice sites when each of the two different SR proteins are used suggests the possibility that different mechanisms may be involved in this alternative splice site selection. SRp40 shows a bias toward more U1 recruitment to the 5′ splice site that is used, suggesting that splice site selection may have occurred at or before U1 recruitment. In contrast, SRp30b may specify splice site selection after U1 recruitment because this SR protein shows equivalent U1 recruitment to both sites. This possibility is supported by recent observations that suggest SR proteins carry out later functions in addition to the earliest step of U1 snRNP recruitment. Splicing extract depleted or debilitated for U1 snRNP can still function in splicing certain pre-mRNA substrates if excess SR proteins are added (32, 33). This is consistent with the model that the U1 snRNP does not appear to be part of the catalytic core of the spliceosome and most likely functions to promote spliceosome assembly. It appears that SR proteins can perform this role of U1 snRNP in its absence. Other evidence for SR protein function at several steps in spliceosome assembly comes from interaction studies. SR proteins bind a subunit of the splicing factor U2AF (28) and appear to be involved in U2 snRNP binding to the pre-mRNA (34). These results suggest the possibility that SR proteins may function to specify splice site selection at one or more of these additional step(s) of spliceosome assembly.

We thank E. Sontheimer and J. Steitz for their gracious gift of 4-thioUpG. We thank Manny Ares, Mark Groudine, Andrew Kirsch, Karla Neugebauer, Rabiya Tuma, and Hal Weintraub for critical comments on the manuscript. This research was supported by Grant 42786-02 (M.B.R.) from the National Institute for General Medical Sciences. A.M.Z. is a Burroughs–Wellcome Fund Fellow of the Life Sciences Research Foundation.

1. Sharp, P. A. (1994) *Cell* **77,** 805–815.
2. Moore, M. J., Query, C. C. & Sharp, P. A. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 303–357.
3. Sharp, P. A. (1985) *Cell* **42,** 397–400.
4. Cech, T. R. (1986) *Cell* **44,** 207–210.
5. Steitz, T. A. & Steitz, J. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6498–6502.
6. Zamore, P. D. & Green, M. R. (1991) *EMBO J.* **10,** 207–214.
7. Zamore, P. D., Patton, J. G. & Green, M. R. (1992) *Nature (London)* **355,** 609–614.
8. Zahler, A. M., Lane, W. S., Stolk, J. A. & Roth, M. B. (1992) *Genes Dev.* **6,** 837–847.
9. Ge, H., Zuo, P. & Manley, J. L. (1991) *Cell* **66,** 373–382.
10. Krainer, A. R., Mayeda, A., Kozak, D. & Binns, G. (1991) *Cell* **66,** 383–394.
11. Fu, X. D. & Maniatis, T. (1992) *Science* **256,** 535–538.
12. Zahler, A. M., Neugebauer, K. M., Stolk, J. A. & Roth, M. B. (1993) *Mol. Cell. Biol.* **13,** 4023–4028.
13. Mayeda, A., Zahler, A. M., Krainer, A. R. & Roth, M. B. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1301–1304.
14. Zahler, A. M., Neugebauer, K. M., Lane, W. S. & Roth, M. B. (1993) *Science* **260,** 219–222.
15. Wyatt, J. R., Sontheimer, E. J. & Steitz, J. A. (1992) *Genes Dev.* **6,** 2542–2553.
16. Sontheimer, E. J. & Steitz, J. A. (1993) *Science* **262,** 1989–1996.
17. Favre, A. (1990) in *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*, ed. Morrison, H. (Wiley, New York), pp. 379–425.
18. Ge, H. & Manley, J. L. (1990) *Cell* **62,** 25–34.
19. Ner, S. S., Atkinson, T. C. & Smith, M. (1989) *Nucleic Acids Res.* **17,** 4015–4023.
20. Moore, M. J. & Sharp, P. A. (1992) *Science* **256,** 992–997.
21. Johnston, R. F., Pickett, S. C. & Parker, D. L. (1990) *Electrophoresis* **11,** 355.
22. Kohtz, J. D., Jamison, S. F., Will, C. L., Zuo, P. Lührmann, R., Garcia-Blanco, M. A. & Manley, J. L. (1994) *Nature (London)* **368,** 119–124.
23. Nelson, K. K. & Green, M. R. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6253–6257.
24. Eperon, I. C., Ireland, D. C., R. A., S., Mayeda, A. & Krainer, A. R. (1993) *EMBO J.* **12,** 3607–3617.
25. Lavigueur, A., La Branche, H., Kornblihtt, A. R. & Chabot, B. (1993) *Genes Dev.* **7,** 2405–2417.
26. Sun, Q., Mayeda, A., Hampson, R. K., Krainer, A. R. & Rottman, F. M. (1993) *Genes Dev.* **7,** 2598–2608.
27. Tian, M. & Maniatis, T. (1993) *Cell* **74,** 105–114.
28. Wu, J. Y. & Maniatis, T. (1993) *Cell* **75,** 1061–1070.
29. Treisman, R., Orkin, S. H. & Maniatis, T. (1983) *Nature (London)* **302,** 591–596.
30. Zuo, P. & Manley, J. L. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 3363–3367.
31. Fu, X. D. (1993) *Nature (London)* **365,** 82–85.
32. Crispino, J. D., Blencowe, B. J. & Sharp, P. A. (1994) *Science* **265,** 1866–1869.
33. Tarn, W. Y. & Steitz, J. A. (1994) *Genes Dev.* **8,** 2704–2717.
34. Fu, X. D. & Maniatis, T. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1725–1729.