

Sip1, a Novel RS Domain-Containing Protein Essential for Pre-mRNA Splicing

WAN-JIANG ZHANG AND JANE Y. WU*

*Department of Pediatrics and Department of Molecular Biology and Pharmacology,
Washington University School of Medicine, St. Louis, Missouri 63110*

Received 1 July 1997/Returned for modification 19 August 1997/Accepted 6 November 1997

Previous studies have shown that protein-protein interactions among splicing factors may play an important role in pre-mRNA splicing. We report here identification and functional characterization of a new splicing factor, Sip1 (SC35-interacting protein 1). Sip1 was initially identified by virtue of its interaction with SC35, a splicing factor of the SR family. Sip1 interacts with not only several SR proteins but also with U1-70K and U2AF65, proteins associated with 5' and 3' splice sites, respectively. The predicted Sip1 sequence contains an arginine-serine-rich (RS) domain but does not have any known RNA-binding motifs, indicating that it is not a member of the SR family. Sip1 also contains a region with weak sequence similarity to the *Drosophila* splicing regulator suppressor of white apricot (SWAP). An essential role for Sip1 in pre-mRNA splicing was suggested by the observation that anti-Sip1 antibodies depleted splicing activity from HeLa nuclear extract. Purified recombinant Sip1 protein, but not other RS domain-containing proteins such as SC35, ASF/SF2, and U2AF65, restored the splicing activity of the Sip1-immunodepleted extract. Addition of U2AF65 protein further enhanced the splicing reconstitution by the Sip1 protein. Deficiency in the formation of both A and B splicing complexes in the Sip1-depleted nuclear extract indicates an important role of Sip1 in spliceosome assembly. Together, these results demonstrate that Sip1 is a novel RS domain-containing protein required for pre-mRNA splicing and that the functional role of Sip1 in splicing is distinct from those of known RS domain-containing splicing factors.

Pre-mRNA splicing takes place in spliceosomes, the large RNA-protein complexes containing pre-mRNA, U1, U2, U4/6, and U5 small nuclear ribonucleoprotein particles (snRNPs), and a large number of accessory protein factors (for reviews, see references 21, 22, 37, 44, and 48). It is increasingly clear that the protein factors are important for pre-mRNA splicing and that studies of these factors are essential for further understanding of molecular mechanisms of pre-mRNA splicing.

Most mammalian splicing factors have been identified by biochemical fractionation and purification (3, 15, 19, 31–36, 45, 69–71, 73), by using antibodies recognizing splicing factors (8, 9, 16, 17, 61, 66, 67, 74), and by sequence homology (25, 52, 74).

Splicing factors containing arginine-serine-rich (RS) domains have emerged as important players in pre-mRNA splicing. These include members of the SR family, both subunits of U2 auxiliary factor (U2AF), and the U1 snRNP protein U1-70K (for reviews, see references 18, 41, and 59). *Drosophila* alternative splicing regulators transformer (Tra), transformer 2 (Tra2), and suppressor of white apricot (SWAP) also contain RS domains (20, 40, 42). RS domains in these proteins play important roles in pre-mRNA splicing (7, 71, 75), in nuclear localization of these splicing proteins (23, 40), and in protein-RNA interactions (56, 60, 64). Previous studies by us and others have demonstrated that one mechanism whereby SR proteins function in splicing is to mediate specific protein-protein interactions among spliceosomal components and between general splicing factors and alternative splicing regula-

tors (1, 1a, 6, 10, 27, 63, 74, 77). Such protein-protein interactions may play critical roles in splice site recognition and association (for reviews, see references 4, 18, 37, 41, 47 and 59). Specific interactions among the splicing factors also suggest that it is possible to identify new splicing factors by their interactions with known splicing factors.

Here we report identification of a new splicing factor, Sip1, by its interaction with the essential splicing factor SC35. The predicted Sip1 protein sequence contains an RS domain and a region with sequence similarity to the *Drosophila* splicing regulator, SWAP. We have expressed and purified recombinant Sip1 protein and raised polyclonal antibodies against the recombinant Sip1 protein. The anti-Sip1 antibodies specifically recognize a protein migrating at a molecular mass of approximately 210 kDa in HeLa nuclear extract. The anti-Sip1 antibodies sufficiently deplete Sip1 protein from the nuclear extract, and the Sip1-depleted extract is inactive in pre-mRNA splicing. Addition of recombinant Sip1 protein can partially restore splicing activity to the Sip1-depleted nuclear extract, indicating an essential role of Sip1 in pre-mRNA splicing. Other RS domain-containing proteins, including SC35, ASF/SF2, and U2AF65, cannot substitute for Sip1 in reconstituting splicing activity of the Sip1-depleted nuclear extract. However, addition of U2AF65 further increases splicing activity of Sip1-reconstituted nuclear extract, suggesting that there may be a functional interaction between Sip1 and U2AF65 in nuclear extract.

MATERIALS AND METHODS

Yeast two-hybrid interaction screening and protein-protein interaction assay.

The yeast two-hybrid interaction system including *Saccharomyces cerevisiae* EGY48, the yeast plasmids, and a HeLa cell cDNA library were kindly provided by R. Brent. SC35 was used as a bait to screen the HeLa cDNA library as described previously (63, 72).

To assay for pairwise interactions between Sip1 and other splicing proteins,

* Corresponding author. Mailing address: Department of Pediatrics and Department of Molecular Biology and Pharmacology, Washington University School of Medicine, One Children's Place, St. Louis, MO 63110. Phone: (314) 454-2081. Fax: (314) 454-2075. E-mail: WU_J@A1.kids.wustl.edu.

yeast plasmids expressing individual splicing proteins as LexA fusion proteins were transformed into yeast strain EGY48 expressing Sip1 as a fusion protein containing the B42 activation domain (63, 72). The liquid assay for β -galactosidase activity was carried out with yeast extracts prepared from at least three independent colonies as described previously (63, 74). β -Galactosidase activities were normalized with protein concentrations of the corresponding yeast extracts. Background was defined as the amount of β -galactosidase activity detected in the yeast expressing the Sip1-activation domain fusion protein and the bait plasmid containing only the LexA without other cDNA sequences.

HeLa cell cDNA library screening and database search. A cDNA fragment encoding Sip1 was isolated from the yeast two-hybrid library vector JG4-5 (72) by digestion with *Eco*RI and *Xho*I. This fragment was labeled by using Klenow enzyme in the presence of [³²P]dCTP and used to screen a HeLa cell λ Zap library (Stratagene) according to standard procedures. Sequencing of cDNA clones was carried out by using a model 373A DNA sequencer with a PRISM Ready Reaction DyeDeoxy Terminator cycle sequencing kit (ABI). Database searches were carried out through the National Institutes of Health mail server, using BLASTN and BLASTP programs. Sequence comparison reveals that the carboxyl-terminal 4.3-kb sequence of Sip1 cDNA is almost identical to that of the human SR129 gene (accession no. Y11251), which was deposited as an unpublished sequence, and that the amino-terminal 1 kb of Sip1 cDNA is different from that of the SR129 gene. It is not clear yet whether Sip1 and SR129 represent different isoforms of the same gene or the Y11251 clone is a hybrid of two different genes. Sequence alignment was performed by using the ClustalW multiple sequence alignment program. Databank searches also revealed the presence of three murine EST cDNA clones which may represent the murine homolog of the *SIP1* gene. No convincing yeast homolog has been identified.

Generation of Sip1 antisera. cDNAs encoding different fragments of Sip1 protein were subcloned into the *Escherichia coli* expression vector pET28A (Novagen) to express His-tagged Sip1 fragments. One of the fragments (amino acid residues 227 to 782) was expressed at a relatively high level. This fragment contained mainly the central part of the Sip1 protein and only a small portion of the RS domain. This His-tagged Sip1 fragment was purified to near homogeneity by using nickel-agarose affinity chromatography in the presence of 6 M urea, and the Sip1 protein was further purified by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The purified Sip1 protein isolated by SDS-PAGE was used to immunize two rabbits to generate polyclonal anti-Sip1 antisera.

Expression and purification of recombinant Sip1 protein by using the baculovirus system. Baculovirus expressing the full-length Sip1 protein as a His-tagged protein was prepared by using the pFastbac system (Gibco). The recombinant Sip1 protein was purified by ammonium sulfate precipitation (45 to 90% saturation) followed by nickel-agarose affinity chromatography according to the manufacturer's instructions. The purified Sip1 protein was dialyzed against BC100 (20 mM HEPES [pH 7.6], 100 mM KCl) containing 10% glycerol and stored at -80°C in aliquots.

Western analyses. Proteins were separated by SDS-PAGE (10% gel) and transferred to nitrocellulose filters. After blocking with 5% nonfat milk in Tris-buffered saline, primary antibodies (anti-Sip1 at 1:2,000 dilution and anti-U2AF65 at 1:1,500 dilution) were applied to the filters, and incubation was carried out at 4°C overnight. After four washes with Tris-buffered saline containing 0.05% Tween 20, horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G [F(ab')₂ fragment specific; Jackson ImmunoResearch Laboratories] was added and incubated for 1 h at room temperature. After four washes, SuperSignal substrate solution (Pierce) was used to produce signals, and the filters were exposed to X-ray films.

In vitro splicing, splicing complex formation, immunodepletion, and complementation assays. Splicing reactions were carried out with β -globin pre-mRNA prepared by in vitro transcription in the presence of [³²P]UTP, using construct H β T7 as described previously (74). The labeled pre-mRNA was incubated with HeLa cell nuclear extract, HeLa nuclear extract depleted of Sip1, or the Sip1-depleted HeLa nuclear extract supplemented with recombinant proteins. The splicing products were analyzed on a 6% polyacrylamide gel containing 6 M urea and detected by autoradiography. For native gel analysis of splicing complexes, splicing reactions were performed with ³²P-labeled pPIP85.A pre-mRNA (43) and separated on a non-denaturing 4% (80:1) polyacrylamide gel in 50 mM Tris-glycine. The gels were dried and exposed to X-ray films for autoradiography.

Anti-Sip1 antiserum or the preimmune serum was used for depletion experiments, performed by a protocol modified from that of Zuo and Maniatis (77). In this protocol, 0.5 ml (1:1 slurry) of protein A-Trisacryl (Pierce) was washed three times and blocked at 4°C for 1 h in immunodepletion (ID) wash buffer (20 mM HEPES [pH 7.8], 150 mM NaCl, 0.5% Nonidet P-40, 2 mM EDTA, 2 mg of bovine serum albumin per ml). The protein A-Trisacryl resin was then packed into a minicolumn and washed with 5 ml of ID wash buffer. Preimmune serum or anti-Sip1 antiserum was diluted with 2.5 ml of ID wash buffer and loaded onto the protein A-Trisacryl column three times. The columns were washed with 5 ml of ID wash buffer and 5 ml of BC100 (20 mM HEPES [pH 7.8], 100 mM KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol). HeLa cell nuclear extract (0.5 ml at approximately 12 mg of protein per ml) was thawed, prewarmed to room temperature for 20 min, and passed five times at room temperature through an anti-Sip1 immunoaffinity column or a mock affinity column prepared with the preimmune serum. The extracts were then used or quick frozen and then stored as immu-

nodepleted nuclear extracts. The individual batches of depleted extracts were examined by Western blotting using the anti-Sip1 and anti-U2AF65 antisera.

For the splicing complementation assay, recombinant SC35, ASF/SF2, and U2AF65 were purified. SC35 and ASF/SF2 were purified as described previously (63). Recombinant U2AF65 protein was purified from *E. coli* BL21(DE3)pLysS transformed with a plasmid expressing U2AF65 from a T7 vector (63). *E. coli* expressing the U2AF65 protein after isopropylthiogalactopyranoside induction was harvested and sonicated in a solution containing 20 mM HEPES (pH 8.0), 0.1M NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, and 5% glycerol. MgCl₂ was added to a final concentration of 20 mM to precipitate the proteins. The recombinant U2AF65 protein was purified to near homogeneity on a MonoQ column with a 0.1 to 1 M NaCl gradient. The purified protein was dialyzed against BC100 containing 10% glycerol and stored in aliquots at -80°C .

Nucleotide sequence accession number. The GenBank accession number for Sip1 cDNA is AF030234.

RESULTS

Identification of Sip1 by its interaction with the essential splicing factor SC35. SC35 is an essential splicing factor of the SR family (16, 17; for a review, see reference 18). It interacts with several other known splicing factors (63, 74). Using SC35 as a bait in the yeast two-hybrid interaction cloning system (72), we have searched for proteins interacting with SC35. After screening approximately 2×10^6 independent colonies, we isolated several groups of cDNAs, including those encoding U1-70K, U2AF35, and two novel SC35-interacting proteins, named Sip1 and Sip2. In the previous study, we characterized specific interactions of SC35 with U1-70K and U2AF35, leading to the proposal of a model for SR protein function in spliceosome assembly (63). It was not clear whether the new proteins identified from the yeast two-hybrid screening were functionally important for splicing. We have now further characterized one of these new SC35-interacting proteins, Sip1.

Using the Sip1 cDNA fragment obtained from the yeast two-hybrid system, we screened a HeLa cell cDNA library and isolated 16 cDNA clones. After analyzing the sequences of these overlapping cDNAs, a probable full-length cDNA of 5,298 bp was obtained by ligation of two overlapping cDNAs. Based on a number of criteria, we conclude that this cDNA represents the full-length cDNA encoding Sip1 protein. First, the cDNA contains an open reading frame encoding 1,403 amino acids (Fig. 1), having a presumptive ATG initiation codon located 416 bp downstream from the 5' end with surrounding sequence conforming to the consensus for mammalian translational initiation sequence (30). Second, there are stop codons present in all three reading frames upstream of the presumptive start codon. Third, the cDNA contains a polyadenylation signal in the 3' untranslated region (62) with a poly(A) tail at the 3' end. Fourth, a major band of approximately 5 kb, consistent with the cDNA size (5,298 bp), was detected in human tissues and cell lines by Northern blotting (Fig. 2). Finally, the protein products of the Sip1 cDNA obtained either from in vitro translation or from the baculovirus expression system migrate at the same size as the Sip1 protein present in HeLa cells as detected by Western blotting using anti-Sip1 antiserum (see Fig. 6; also data not shown). Taken together, these results indicate that the Sip1 protein sequence shown in Fig. 1 is very likely to be of full length.

Predicted sequence of Sip1 protein and its limited sequence similarities to known splicing factors. The predicted Sip1 protein contains 1,403 amino acid residues (Fig. 1) with an estimated molecular mass of 157.9 kDa. It has a high percentage of serine (13.8%) and charged residues (8.3% lysine, 7.1% arginine, and 9.1% glutamine). Data bank searches reveal that Sip1 protein has certain sequence features of proteins involved in pre-mRNA splicing. First, the central portion of the Sip1 protein contains a domain rich in arginine-serine sequences.

1 MCFMTCILKWAETLASCPIIDRKPFQAVFKFSALEGYVKVQVKQLRETKDKKNENSFEKQVSCHEMSKSCIRRKAIIVREDLLSAKVCDL 90

91 KWIHRNSLYSETGGKKNAAIKINKPQRSNWSTNQCFRNFFSNMSSVSHSGESSFTYRAYCTEFIEASEISALIRQKRHELELSWFPDTL 180

181 PGIGRIGFIPWNVETEVLPLLISSVLPRTIFPTSTISFEHFGTSCKGYALHAHQEGEEKQTSGETSNTRGSRRKPAAMTTPTRSTRNTRAE 270

271 TASQSORSPISDNSGCDAPGNSNPSSLVSPSSAESEKQTRQAPKRKSVRRGRKPPLLKLLRSSVAAPKSSSNDSDVEETAESDTSPLVE 360

361 KEHQPDVDSNICITVQTHVENQSANCLKSCNEQIEESEKHTANYDTEERVGSSSESQAQDLPLVLVGEEGEVKLENTGIEANVLCLESE 450

450 ISENILEKGGDPLEKQDQISGLSQSEVKTQVTVHLPNDPFTCLTSESKVYQPVSCPLSDLSENVESVVNEEKITESSLVEITEHKDFTL 540

541 KTEELIESPKLESSEGEIIQTVDRQSVKSPEVQLLGHVETEDVEIIATCTDFGNEDFNNIQDSENLLKNNLLNTKLEKSLEEKNESLTE 630

631 HPRSTELPKTHIEQIQKHFSEDNNEMIPMECDSEFCSDQNESEVEPSVNADLKQMNENSVTHCSENMPSSDLADEKVETVSPSESPKDT 720

721 IDKTKKPRTRRSRFHSPSTTWSPNKDTPQEKKRQSPSPRRETGKESRSQSPPKNE SARGRKKGRSQSPKDIARERROSRSRSPKRD 810

811 TTRESRRSESLSPRRETSSRENKRSQPRVKDSSPGEKRSQSREERESDRDQRRERERRTRKWSRSRSHSRSPSCRTRKSKSSSFGRIDRD 900

901 SYSPRWKGRWANDGWRCPRGNDRYRKNDEPKQENRTRKEKNDIHLDADDPSADKHRNDCPNWITEKINSQDPDPRTRNPEKLKESHWEEEN 990

991 RNENSGNSWNKNFGSGWVSNRGRGRGRGRGTYSRFAYKDQENRWNQNRKPLSGNSSSGSESFKFVEQQSYKRSKSEQEFSFDTPADRS 1080

1081 GWTASASSWAVRKTLPADVQNYYSRRGRNSGGPQSGWMQEEETSGQDSSLKQDINQQVDGSQLPINMMQPMNVMMQQMNAQHQPMMNIFP 1170

1171 YPVGVHAPLMNIQRNPFNIHPQLPLHLHTGVPLMQVATPTSVSQGLPPPPPPPPSQVNYIASQPDGKQLQGIPSSSHVSNMSTPVL 1260

1261 APTAAPGNTGMVQGPSSGNTSSSSHSKASNAAVKLAESKVSVAVEASDSSKTDKLLQIQEKAAQEVKLAIKPFYQNKDITKEEYKEIVR 1350

1351 KAVDKVCCHSKSGEVNSTKVANLVKAYVDKIKYSRKGSQKKTLEEPVSTEKNIG*

FIG. 1. Amino acid sequence of Sip1 predicted from the full-length cDNA sequence. SR or RS dipeptides are underlined. Eight imperfect repeats of RRSRSXSX are in boldface and underlined. The CTD-binding motif is doubly underlined.

Eight imperfect repeats of RRSRSXSX are found in this RS domain, and this type of sequence motif is present in several SR proteins (8, 9, 13, 17, 19, 34, 52, 67), spliceosomal protein U1-70K (46, 54), and *Drosophila* splicing regulators Tra and SWAP (40, 42). In addition to this repeat motif, the arginine-serine-rich region of Sip1 shows primary sequence similarity to the RS domain of several SR proteins, with the highest per-

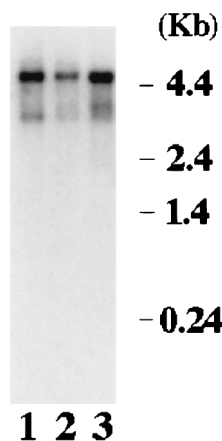


FIG. 2. Northern analysis of Sip1 mRNA. Approximately 2 μ g of poly(A)-selected RNA prepared from two human cell lines (HPB-ALL [lane 1] and Raji [lane 2]) or human thymus tissue (lane 3) was loaded in each lane. The RNA was transferred to a nitrocellulose filter, and the filter was probed with a 32 P-labeled probe prepared from Sip1 cDNA. The autoradiograph of the filter is shown.

centage of amino acid identity to that of SRp75 (Fig. 3A). The sequence similarity in this region is not limited to SR dipeptides or the RRSRSXSX motif; there are 27% identity and 52% similarity over a stretch of 112 amino acid residues (Fig. 3A). Second, Sip1 shows 17% identity and 31% similarity to SWAP in a region of 158 residues from amino acids 273 to 431 (Fig. 3B). This region partially overlaps the first SURP module (12, 36, 53) of the *Drosophila* SWAP. The SURP motif has been identified in a number of splicing proteins, including SWAP, PRP21, SF3a120, and their homologs in different species (12, 36, 53). However, the highly conserved there is of the SURP motif are not present in Sip1, suggesting that Sip1 is not a member of this SURP family. Another feature of the Sip1 protein is that at the very carboxyl terminus there is an 80-amino-acid motif (Fig. 1) which has been proposed to be the domain on two rat proteins which interacts with the carboxyl-terminal domain (CTD) of RNA polymerase II (65). The presence of this putative CTD-binding domain in Sip1 protein raises the possibility that Sip1 is involved in communication between the splicing machinery and the transcription machinery.

Interaction of Sip1 with U2AF65, U1-70K, and SR proteins. Since splicing factors containing RS domains have been shown to mediate protein-protein interactions important for splicing, we examined protein-protein interactions of Sip1 with other splicing factors, especially the proteins containing RS domains. In the yeast two-hybrid interaction assays, Sip1 interacts not only with SR proteins SC35, ASF/SF2, SRp75, and SRp20 but also with U1-70K and U2AF65 (Fig. 4). The *Drosophila* splicing regulators Tra and Tra2 as well as the *Drosophila* SR

A.	Sip1 : 722	<u>D</u> <u>K</u> <u>T</u> <u>K</u> <u>K</u> <u>P</u> <u>R</u> <u>T</u> <u>R</u> <u>S</u> <u>R</u> <u>F</u> <u>H</u> <u>S</u> <u>P</u> <u>S</u> <u>T</u> <u>T</u> <u>W</u> <u>S</u> <u>P</u> <u>N</u> <u>K</u> <u>D</u> <u>T</u> <u>P</u> <u>O</u> <u>E</u> <u>K</u> <u>K</u> <u>R</u> <u>P</u> <u>S</u> <u>P</u> <u>S</u> <u>P</u> <u>R</u> <u>E</u> <u>T</u> <u>G</u> <u>K</u> <u>E</u> <u>S</u> <u>R</u> <u>K</u> <u>S</u> <u>O</u> <u>S</u> <u>P</u> <u>S</u> <u>P</u> <u>K</u> <u>N</u> <u>E</u> <u>S</u> <u>A</u> <u>R</u>	
	SRp75: 322	<u>E</u> <u>O</u> <u>E</u> <u>K</u> <u>L</u> <u>R</u> <u>Q</u> <u>S</u> <u>L</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>K</u> <u>A</u> <u>G</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>K</u> <u>S</u> <u>K</u> <u>D</u> <u>K</u> <u>R</u> <u>K</u> <u>S</u> <u>R</u> <u>K</u> <u>R</u> <u>S</u> <u>E</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>R</u> <u>S</u> <u>K</u> <u>S</u> <u>E</u> <u>S</u> <u>R</u> <u>K</u> <u>R</u> <u>G</u>	
	Sip1 : 782	<u>G</u> <u>R</u> <u>K</u> <u>K</u> <u>S</u> <u>R</u> <u>S</u> <u>S</u> <u>O</u> <u>S</u> <u>P</u> <u>K</u> <u>K</u> <u>D</u> <u>I</u> <u>A</u> <u>R</u> <u>E</u> <u>R</u> <u>R</u> <u>O</u> <u>S</u> <u>O</u> <u>S</u> <u>R</u> <u>S</u> <u>P</u> <u>K</u> <u>R</u> <u>D</u> <u>T</u> <u>I</u> <u>R</u> <u>E</u> <u>S</u> <u>R</u> <u>R</u> <u>S</u> <u>E</u> <u>S</u> <u>L</u> <u>S</u> <u>P</u> <u>R</u> <u>E</u> <u>T</u> <u>S</u> <u>R</u> <u>E</u> <u>N</u> <u>K</u> <u>R</u> <u>S</u>	834
	SRp75: 382	<u>S</u> <u>K</u> <u>R</u> <u>D</u> <u>S</u> <u>K</u> <u>A</u> <u>G</u> <u>S</u> <u>S</u> <u>K</u> <u>K</u> <u>K</u> <u>K</u> <u>E</u> <u>D</u> <u>T</u> <u>D</u> <u>R</u> <u>S</u> <u>O</u> <u>S</u> <u>R</u> <u>S</u> <u>P</u> <u>S</u> <u>R</u> <u>S</u> <u>V</u> <u>S</u> <u>K</u> <u>E</u> <u>R</u> <u>E</u> <u>H</u> <u>A</u> <u>K</u> <u>S</u> <u>E</u> <u>S</u> <u>S</u> <u>O</u> <u>R</u> <u>E</u> <u>G</u> <u>R</u> <u>G</u> <u>E</u> <u>S</u> <u>E</u> <u>N</u> <u>A</u>	434
B.	Sip1: 273	<u>S</u> <u>O</u> <u>S</u> <u>O</u> <u>R</u> <u>S</u> <u>P</u> <u>I</u> <u>S</u> <u>D</u> <u>N</u> <u>S</u> <u>G</u> <u>C</u> <u>D</u> <u>A</u> <u>P</u> <u>G</u> <u>N</u> <u>S</u> <u>N</u> <u>P</u> <u>S</u> <u>L</u> <u>S</u> <u>V</u> <u>P</u> <u>S</u> <u>S</u> <u>A</u> <u>E</u> <u>S</u> <u>E</u> <u>K</u> <u>Q</u> <u>T</u> <u>R</u> <u>Q</u> <u>A</u> <u>P</u> <u>K</u> <u>R</u> <u>K</u> <u>S</u> <u>V</u> <u>R</u> <u>R</u> <u>G</u> <u>R</u> <u>K</u> <u>P</u> <u>L</u> <u>L</u> <u>K</u> <u>K</u> <u>L</u> <u>R</u> <u>S</u>	
	SWAP: 307	<u>T</u> <u>M</u> <u>K</u> <u>Q</u> <u>H</u> <u>A</u> <u>I</u> <u>E</u> <u>K</u> <u>T</u> <u>A</u> <u>R</u> <u>F</u> <u>I</u> <u>A</u> <u>T</u> <u>O</u> <u>G</u> <u>A</u> <u>Q</u> <u>M</u> <u>E</u> <u>L</u> <u>L</u> <u>I</u> <u>K</u> <u>A</u> <u>K</u> <u>Q</u> <u>A</u> <u>N</u> <u>N</u> <u>T</u> <u>Q</u> <u>F</u> <u>D</u> <u>F</u> <u>L</u> <u>T</u> <u>Q</u> <u>G</u> <u>H</u> <u>L</u> <u>Q</u> <u>P</u> <u>Y</u> <u>R</u> <u>H</u> <u>L</u> <u>A</u> <u>A</u> <u>I</u> <u>K</u> <u>A</u> <u>K</u>	
	Sip1: 333	<u>S</u> <u>V</u> <u>A</u> <u>A</u> <u>P</u> <u>E</u> <u>K</u> <u>S</u> <u>S</u> <u>N</u> <u>D</u> <u>S</u> <u>V</u> <u>D</u> <u>E</u> <u>E</u> <u>T</u> <u>A</u> <u>E</u> <u>S</u> <u>D</u> <u>T</u> <u>S</u> <u>P</u> <u>V</u> <u>L</u> <u>E</u> <u>K</u> <u>E</u> <u>H</u> <u>Q</u> <u>P</u> <u>D</u> <u>V</u> <u>D</u> <u>S</u> <u>S</u> <u>N</u> <u>I</u> <u>C</u> <u>T</u> <u>V</u> <u>O</u> <u>T</u> <u>H</u> <u>V</u> <u>E</u> <u>N</u> <u>O</u> <u>S</u> <u>A</u> <u>N</u> <u>C</u> <u>L</u> <u>K</u> <u>S</u> <u>C</u> <u>N</u>	
	SWAP: 367	<u>F</u> <u>P</u> <u>P</u> <u>A</u> <u>P</u> <u>Q</u> <u>T</u> <u>P</u> <u>L</u> <u>D</u> <u>Q</u> <u>N</u> <u>T</u> <u>D</u> <u>K</u> <u>E</u> <u>A</u> <u>P</u> <u>S</u> <u>A</u> <u>D</u> <u>D</u> <u>H</u> <u>S</u> <u>E</u> <u>E</u> <u>V</u> <u>A</u> <u>G</u> <u>G</u> <u>R</u> <u>R</u> <u>N</u> <u>P</u> <u>N</u> <u>Q</u> <u>V</u> <u>V</u> <u>I</u> <u>T</u> <u>V</u> <u>P</u> <u>T</u> <u>I</u> <u>K</u> <u>Y</u> <u>K</u> <u>P</u> <u>S</u> <u>A</u> <u>N</u> <u>C</u> <u>A</u> <u>T</u> <u>L</u> <u>Q</u>	
	Sip1: 392	<u>E</u> <u>Q</u> <u>I</u> <u>E</u> <u>S</u> <u>E</u> <u>K</u> <u>H</u> <u>T</u> <u>A</u> <u>N</u> <u>Y</u> <u>D</u> <u>T</u> <u>E</u> <u>E</u> <u>R</u> <u>V</u> <u>G</u> <u>S</u> <u>S</u> <u>S</u> <u>E</u> <u>S</u> <u>C</u> <u>A</u> <u>Q</u> <u>D</u> <u>L</u> <u>P</u> <u>V</u> <u>L</u> <u>V</u> <u>G</u> <u>E</u> <u>E</u> <u>G</u> <u>E</u>	431
	SWAP: 425	<u>I</u> <u>S</u> <u>K</u> <u>I</u> <u>K</u> <u>G</u> <u>V</u> <u>P</u> <u>L</u> <u>Q</u> <u>A</u> <u>V</u> <u>L</u> <u>Q</u> <u>E</u> <u>D</u> <u>E</u> <u>S</u> <u>S</u> <u>N</u> <u>P</u> <u>G</u> <u>N</u> <u>S</u> <u>C</u> <u>H</u> <u>S</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>S</u> <u>P</u> <u>A</u> <u>L</u> <u>S</u> <u>C</u> <u>R</u> <u>S</u> <u>E</u> <u>G</u> <u>H</u>	466

FIG. 3. Sequence similarities between Sip1 and SRp75 (A) and between Sip1 and *Drosophila* SWAP (B). Identical amino acid residues are in boldface and underlined. Conserved (I-L-V, D-E, K-R, S-T) amino acids are underlined.

protein dSRp55 also interact with Sip1 in this assay. These interactions are not due to nonspecific association between the arginine-serine-rich sequences because no significant interactions were detected between Sip1 and U2AF35 or p54, although both of these proteins contain RS domains. Therefore, the profile of protein-protein interactions of Sip1 is different from that observed with SR proteins such as SC35 and ASF/SF2, consistent with the structural differences between Sip1 and SR proteins.

To confirm the results from the yeast two-hybrid system, we examined protein-protein interactions in vitro. Sip1, U2AF65, and U2AF35 proteins were labeled by in vitro translation in the presence of [³⁵S]methionine. Coimmunoprecipitation experiments were carried out to examine possible protein-protein interactions between Sip1 and U2AF65 or U2AF35 (Fig. 5). While the anti-Sip1 antibody (see below and Materials and Methods) immunoprecipitated the in vitro-translated Sip1 protein (lane 6), it did not precipitate either U2AF35 (lane 2) or U2AF65 (lane 4). When Sip1 was coincubated with U2AF65, the anti-Sip1 antibody precipitated both U2AF65 and Sip1 (lane 5). However, when Sip1 was coincubated with U2AF35, the anti-Sip1 antibody precipitated Sip1 but not U2AF35 (lane 3). These results demonstrate that Sip1 interacts with U2AF65 but not U2AF35. The coprecipitation of Sip1 and U2AF65 was

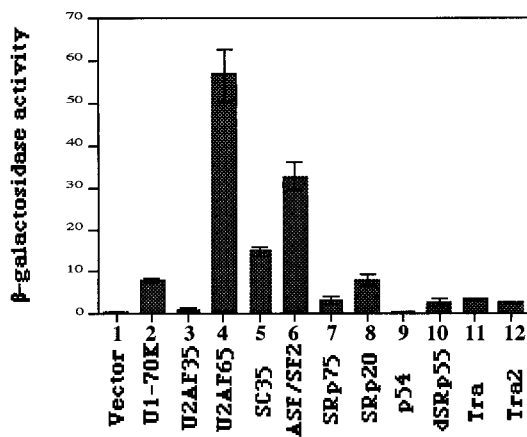


FIG. 4. Interactions between Sip1 and different RS domain-containing splicing factors (as indicated under the x axis) detected by the yeast two-hybrid interaction assay. Quantitative liquid β-galactosidase assays were performed with yeast extracts from at least three independent yeast isolates for each combination. Relative β-galactosidase activities shown represent fold activation above background (see Materials and Methods).

not dependent on the presence of RNA, because RNase treatment did not affect the coprecipitation (data not shown). Similar immunoprecipitation experiments indicate that Sip1 also interacts with U1-70K and SC35 in vitro. The results from the yeast two-hybrid system and coimmunoprecipitation experiments were thus consistent in revealing specific protein-protein interactions between Sip1 and these known splicing factors. The observation that Sip1 interacts with a number of essential splicing factors strongly suggests that Sip1 may be involved in splicing.

Expression of recombinant Sip1 protein and generation of anti-Sip1 antiserum. To investigate the biochemical function of Sip1, we decided to express recombinant Sip1 protein and generate specific antibodies recognizing Sip1. After several attempts, we were unable to express the full-length Sip1 in *E. coli*. However, a fragment of Sip1 containing amino acid residues 227 to 782 was expressed as a His-tagged protein in *E. coli*. The His-tagged Sip1 protein fragment was purified to apparent homogeneity and was used to raise polyclonal antibodies in rabbits.

The anti-Sip1 antibodies specifically recognize a protein in HeLa cell nuclear extract of approximately 210 kDa (Fig. 6A), larger than the molecular mass predicted from the full-length

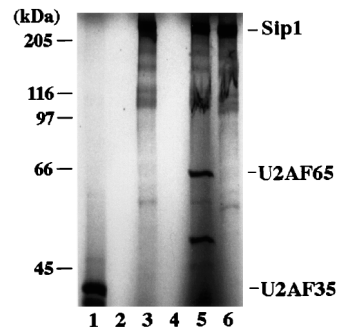


FIG. 5. Interaction of Sip1 with U2AF65 but not U2AF35 as assayed by coimmunoprecipitation. Sip1, U2AF65, and U2AF35 were labeled with [³⁵S]methionine by in vitro translation reactions. Individual proteins alone or proteins after coincubation were immunoprecipitated with anti-Sip1 antiserum (lanes 2 to 6). Lane 1 contains U2AF35 immunoprecipitated by anti-FLAG antibody, which recognizes the epitope tag on the in vitro-translated U2AF35. Lanes 2 and 4 contain immunoprecipitates from U2AF35 (lane 2) and U2AF65 (lane 4) in vitro translation products precipitated by anti-FLAG antibody, showing that this antibody does not cross-react with either U2AF35 or U2AF65. Lanes 3 and 5 contain immunoprecipitates formed after coincubation of Sip1 with U2AF35 (lane 3) or with U2AF65 (lane 5), using anti-Sip1 antibody. Lane 6 contains Sip1 in vitro translation products precipitated by anti-Sip1.

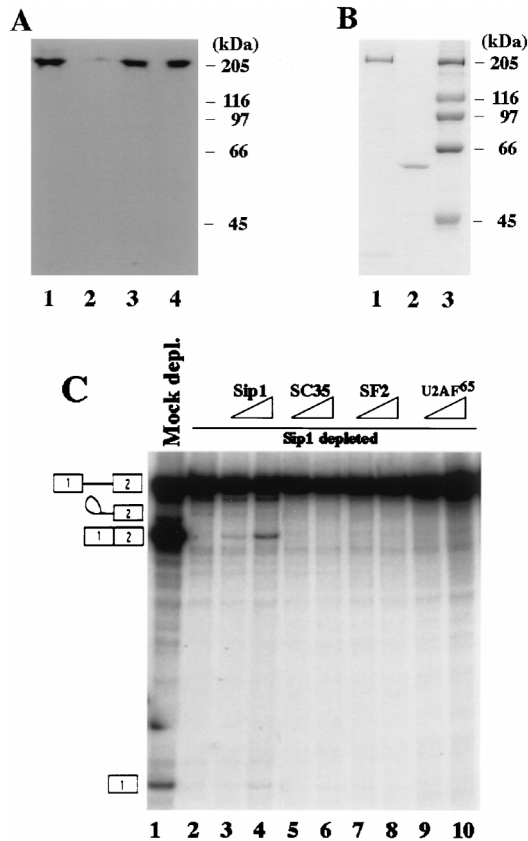


FIG. 6. The essential role of Sip1 in pre-mRNA splicing is not redundant with that of SC35, ASF/SF2, or U2AF65. (A) Sip1 protein in HeLa cell nuclear extract could be depleted by anti-Sip1 but not preimmune serum. Western blotting analysis was performed with the anti-Sip1 antibody, using HeLa nuclear extract (lane 1), the extract after immunodepletion with anti-Sip1 or preimmune serum (lane 3), or the purified recombinant Sip1 protein from the baculovirus expression system (lane 4). (B) Coomassie blue staining of the purified recombinant Sip1 protein (lane 1), U2AF65 protein (lane 2), and protein molecular mass markers (lane 3). (C) The recombinant Sip1, but not SC35, ASF/SF2, or U2AF65, could partially restore splicing activity of the Sip1-depleted nuclear extract. In vitro splicing of human β -globin pre-mRNA was carried out for 90 min with HeLa nuclear extract after immunodepletion by the preimmune serum (Mock depl., lane 1) or anti-Sip1 (lanes 2 to 4). The splicing extract was supplemented with 100 and 300 ng of purified Sip1 (lanes 3 and 4), SC35 (lanes 5 and 6), ASF/SF2 (lanes 7 and 8), or U2AF65 (lanes 9 and 10). The splicing reaction products were analyzed on a 6% polyacrylamide gel.

cDNA (158 kDa). The 210-kDa protein was confirmed to be Sip1 by the observation that both the in vitro-translated product prepared from the full-length Sip1 cDNA and recombinant Sip1 protein purified from the baculovirus system migrated at a molecular mass of approximately 210 kDa (Fig. 5, 6A, and 6B). The specificity of the anti-Sip1 antibodies was also shown by the findings that the antibodies precipitated the 210-kDa protein prepared by in vitro translation of the mRNA made from the full-length Sip1 cDNA but did not precipitate other RS domain-containing proteins or unrelated proteins and that Western blotting using the antibodies detected only one protein band of 210 kDa in the nuclear extracts prepared from HeLa cells and did not cross-react with other proteins (Fig. 6A and data not shown). These results suggest that the difference between the predicted molecular mass and the apparent size determined by SDS-PAGE may be due to posttranslational modification such as phosphorylation, which is known to cause aberrant migration of several RS domain-containing proteins in SDS-PAGE (for a review, see reference 18).

For functional studies, we expressed a full-length recombinant Sip1 protein in the baculovirus system. The recombinant Sip1 protein from the baculovirus system was also recognized by the anti-Sip1 antiserum (Fig. 6A, lane 4) and migrated similarly to the Sip1 protein in HeLa nuclear extract (Fig. 6A, lanes 1 and 3), suggesting that the baculovirus-expressed recombinant Sip1 protein underwent posttranslational modification(s) similar to that in HeLa cells.

Sip1 protein as an essential component in HeLa nuclear extract for its pre-mRNA splicing activity. The observation that Sip1 not only has sequence features of proteins involved in pre-mRNA splicing but also interacts with several proteins important for splicing, including SR proteins and U2AF65, suggests that Sip1 plays a role in splicing. To test this, we first investigated whether depletion of Sip1 from HeLa cell nuclear extract could affect the splicing activity. Under appropriate conditions, 80 to 90% of Sip1 protein from the nuclear extract could be removed by immunodepletion using the anti-Sip1 antiserum (Fig. 6A; compare lane 2 with lane 1), whereas the level of Sip1 was not reduced in the mock-depleted extract with the preimmune serum (Fig. 6A; compare lane 3 with lane 1). We examined the effect of immunodepletion on the splicing activity of HeLa nuclear extract. Splicing reactions were carried out with β -globin pre-mRNA prepared by in vitro transcription in the presence of [32 P]UTP with the H β T7 construct as described previously (74). When the labeled pre-mRNA was incubated with HeLa nuclear extract, the β -globin pre-mRNA was efficiently spliced (Fig. 7B, lanes 8 and 16). The splicing activity was not affected in the nuclear extract treated with the preimmune serum (Fig. 6C, lane 1; Fig. 7B, lanes 1 and 9). By contrast, β -globin pre-mRNA was not spliced by the nuclear extract immunodepleted with the anti-Sip1 antiserum (Fig. 6C, lane 2; Fig. 7B, lanes 2 and 10).

To test whether Sip1 was essential for splicing, the recombinant Sip1 protein purified from the baculovirus system was added back to the immunodepleted nuclear extract. Addition of other RS domain-containing proteins, including SC35, ASF/SF2, or U2AF65, had no effect on the splicing activity of the Sip1-depleted extract (Fig. 6C, lanes 5 to 10). The recombinant Sip1 protein which had been purified to apparent homogeneity (Fig. 6B, lane 1), on the other hand, could restore the splicing activity of the Sip1-depleted extract (Fig. 6C, lanes 3 and 4; Fig. 7B). In the range of 100 to 300 ng of Sip1 protein, the restoration of splicing activity was dependent on the amount of Sip1 protein added (Fig. 6C, lanes 3 and 4; Fig. 7B, lanes 3 to 5 and 11 to 13). Further increases in the amount of Sip1 protein did not increase the splicing activity further (data not shown). The restoration of the splicing activity is due to the activity of the purified recombinant Sip1 protein but not due to contaminating proteins from Sf9 cells because mock protein preparations and other RS domain-containing proteins purified from Sf9 cells did not show any activity, and pretreatment of the purified recombinant Sip1 protein with the anti-Sip1 antibody completely removed the splicing reconstitution activity (data not shown). The observation that addition of purified recombinant Sip1 protein but not SC35, ASF/SF2, or U2AF65 protein to the Sip1-depleted extract could restore the splicing activity indicates that the lack of splicing activity in the Sip1-depleted extract was due to Sip1 deficiency in the extract. These results demonstrate that Sip1 protein is essential for splicing and that the functional role of Sip1 in splicing is not redundant with that of known RS domain-containing proteins such as SC35, ASF/SF2 or U2AF65.

U2AF65 further enhances splicing activity in the Sip1-reconstituted nuclear extract. We noticed that the addition of Sip1 could only partially restore the splicing activity to the

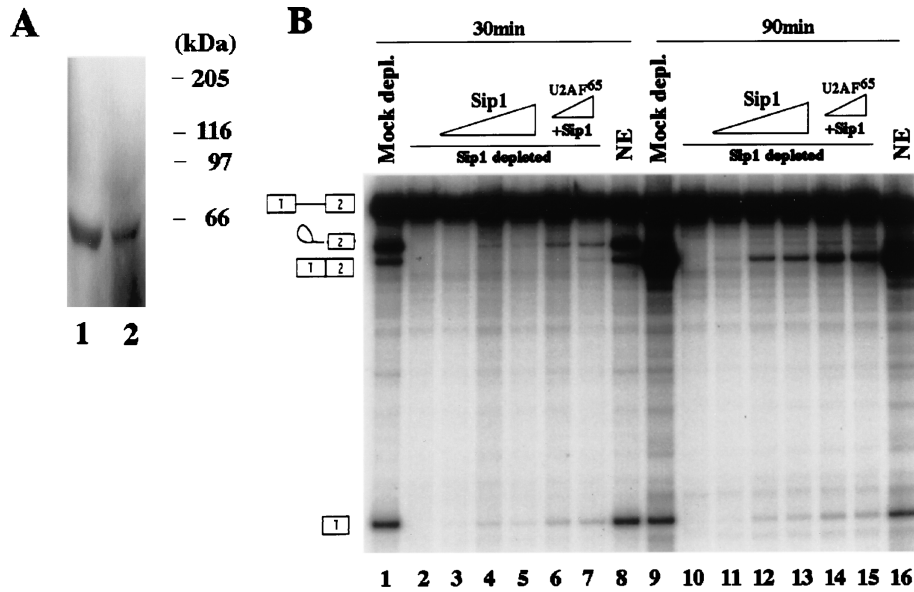


FIG. 7. Sip1 interacts with U2AF65 in HeLa cell nuclear extract, and U2AF65 enhances splicing activity of the Sip1-reconstituted nuclear extract. (A) The level of U2AF65 in the Sip1-immunodepleted nuclear extracts was less than in the mock-depleted nuclear extracts as detected by Western blotting. Approximately 150 μ g of nuclear extracts which had been immunodepleted by using either preimmune (lane 1) or anti-Sip1 (lane 2) serum was fractionated by SDS-PAGE (10% gel), transferred to a nitrocellulose membrane, and probed with anti-U2AF65 antibody (77). (B) U2AF65 further enhanced the splicing reconstitution by Sip1 in the Sip1-depleted nuclear extract. Lanes 1 to 9 contain the products of in vitro splicing of human β -globin pre-mRNA after a 30-min incubation; lanes 9 to 16 are the corresponding reaction products after a 90-min incubation. The splicing reaction products obtained with untreated HeLa nuclear extract (NE) are shown in lanes 8 and 16. The splicing products obtained with HeLa nuclear extract after immunodepletion by the preimmune serum (Mock depl.) are shown in lanes 1 and 9. Lanes 2 to 5 and 10 to 13 contain the splicing reaction products obtained with Sip1-immunodepleted HeLa nuclear extract supplemented with 0, 100, 200, and 300 ng of purified recombinant Sip1 protein after 30- or 90-min incubation. Lanes 6 and 7 and lanes 14 and 15 show the splicing products obtained with the Sip1-depleted extract supplemented with 200 ng of purified Sip1 and 100 or 200 ng of purified U2AF65 protein, respectively.

nuclear extract immunodepleted by the anti-Sip1 antiserum. To determine whether this was due to the absence or reduction of other splicing factors which interact with Sip1, we examined the protein level of U2AF65 in the mock- and Sip1-immunodepleted nuclear extracts because U2AF65 interacts strongly with Sip1 in both the yeast two-hybrid assay and in vitro biochemical experiments (Fig. 4 and 5). Western blotting experiments using anti-U2AF65 (77) showed that the level of U2AF65 was consistently lower in Sip1-immunodepleted extracts than in the mock-immunodepleted extracts with the preimmune serum (Fig. 7A). Consistent with this, U2AF65 protein was detectable in the beads by Western blotting with anti-U2AF65 after immunoprecipitation of the nuclear extract by anti-Sip1 antibodies (data not shown). Because the anti-Sip1 antibody did not cross-react with U2AF65 either by Western blotting or in the very sensitive immunoprecipitation experiments with 35 S-labeled U2AF65 protein (Fig. 6 and 5), U2AF65 protein itself was not directly precipitated by the anti-Sip1 antibody in the absence of Sip1 (Fig. 5 and data not shown). The most likely explanation is that the level of U2AF65 was reduced in the Sip1-immunodepleted extract because of protein-protein interaction between U2AF65 and Sip1. We then tested whether combination of Sip1 and U2AF65 could further increase the splicing activity of the Sip1-immunodepleted extracts. While the addition of U2AF65 alone to the Sip1-depleted extract did not lead to any detectable splicing activity (Fig. 6C, lanes 9 and 10), addition of 100 to 200 ng of U2AF65 protein together with 200 ng of Sip1 protein further enhanced splicing activity (Fig. 7B; compare lanes 6 and 7 with lane 5 and lanes 14 and 15 with lane 13). Together, these results suggest that U2AF65 may interact with Sip1 in the spliceosome and functionally cooperate with Sip1 in the splicing reaction.

That Sip1 may functionally interact with other splicing fac-

tors was suggested by the fact that the level of splicing activity after both Sip1 and U2AF65 addition was only about 20 to 30% of that detected with the untreated nuclear extract (Fig. 7B, lanes 8 and 16) or with the mock-depleted nuclear extract with the preimmune serum (Fig. 7B, lanes 1 and 9). The possibility of functional interactions between Sip1 and other splicing factors in the splicing reaction is consistent with the protein-protein interaction profile revealed by the yeast two-hybrid assay and immunoprecipitation studies.

Sip1 is involved in spliceosome assembly. To dissect the functional mechanism of Sip1 in splicing, we examined protein-protein interactions between Sip1 and other splicing factors, especially the proteins involved in the early steps of spliceosome assembly such as U1-70K, SR proteins, and U2AF65, as described above (Fig. 4 and 5). Interaction of Sip1 with several proteins essential for early steps of spliceosome assembly strongly suggests that Sip1 may be involved in this process. We therefore investigated the role of Sip1 in spliceosome assembly. Nonspecific H complex, prespliceosomal complex A, and spliceosomal complex B can be separated by native gel electrophoresis (28, 29). Using 32 P-labeled pPIP85.A pre-mRNA (43), we demonstrated that in the Sip1-depleted nuclear extract, the formation of both A and B complexes was deficient whereas the formation of these splicing complexes in the mock-depleted nuclear extract was not affected (Fig. 8), indicating that Sip1 plays an important role in spliceosome assembly.

Our results demonstrate that depletion of Sip1 leads to blockage of splicing prior to the first step without any detectable cleaved first exon or lariat intermediate (Fig. 6 and 7) and that assembly of both A and B splicing complexes is defective in the Sip1-depleted nuclear extracts (Fig. 8). Because Sip1 interacts with several proteins essential for the earliest steps of

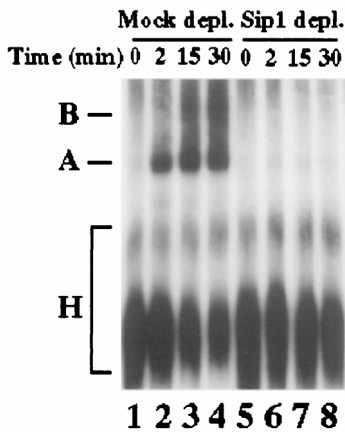


FIG. 8. Involvement of Sip1 in spliceosome assembly. The splicing complexes were formed with ^{32}P -labeled pPIP85.A pre-mRNA at different time points in the presence of ATP, using nuclear extracts which were immunodepleted (depl.) by using the preimmune or anti-Sip1 serum (lanes 1 to 4 and 5 to 8, respectively). The splicing complexes were separated on a native 4% polyacrylamide gel by electrophoresis. The autoradiograph of the gel is shown. Positions of A and B complexes and nonspecific H complex are indicated.

spliceosome assembly such as U2AF65 and SR proteins, it is very likely that Sip1 is involved in these earliest steps of spliceosome assembly.

DISCUSSION

We have demonstrated that it is possible to identify new proteins involved in pre-mRNA splicing through their specific interactions with other splicing factors. A protein identified by this approach, Sip1, has been shown to be a new factor essential for mammalian pre-mRNA splicing.

Structural features of Sip1 protein. Sip1 is a novel protein with sequence similarity to previously identified splicing factors. It contains eight imperfect RRSRSXSX repeats, a motif present in the RS domain of several proteins involved in constitutive splicing and alternative splicing regulation, such as SR proteins and the *Drosophila* splicing regulators Tra and SWAP. RS domains have been implicated in both protein-protein and protein-RNA interactions (1, 23, 27, 40, 56, 60, 63, 77). We are currently investigating whether the RS domain of Sip1 plays a role in mediating protein-RNA and protein-protein interactions. Several proteins such as U2AF35 and Tra do not contain an RNA-binding motif (42, 73), yet they play important roles in splicing and alternative splicing regulation (42, 77). Sip1 has additional sequence similarity outside of the RS domain with the *Drosophila* splicing regulator SWAP. We will further investigate the functional significance of this sequence similarity.

Sip1 protein migrates at a molecular mass of approximately 210 kDa, similar to the apparent molecular masses of the U5 snRNP-specific 200- and 220-kDa proteins. Recent information on the partial peptide sequences of these two proteins shows that the 200-kDa protein is a member of DEXH-box family of putative RNA helicases and the 220-kDa protein is a homolog of yeast PRP8 (38). Thus, Sip1 is distinct from these proteins. In affinity-purified spliceosomes assembled on adenovirus major late promoter pre-mRNA or tropomyosin pre-mRNA and spliceosomes purified by gel filtration, in addition to the doublet corresponding to the U5-specific proteins 200 and 220 kDa, there appear to be a few protein spots in the 200-kDa range as detected by silver staining of two-dimensional gels (2). In addition, a group of SR protein-related

polypeptides have been identified (5, 5a). It remains to be determined whether Sip1 is one of these spliceosome-associated proteins or SR protein-related polypeptides.

The presence of the 80-amino-acid RNA polymerase II CTD-binding motif in the Sip1 protein suggests that Sip1 may be able to interact with the RNA polymerase II CTD, which has been implicated in pre-mRNA processing (65), and play a role in linking the processes of transcription and pre-mRNA splicing. This possibility will be further investigated.

Interactions between Sip1 and spliceosomal components.

We have tested whether Sip1 is a spliceosomal snRNP by examining the presence of snRNAs including U1, U2, U4, U5, and U6 in the immunoprecipitates from HeLa nuclear extracts formed by anti-Sip1 antibodies. We found that no significant amount of spliceosomal snRNPs could be detected in the Sip1 immunoprecipitates (74a), suggesting that Sip1 is not an snRNP protein.

Systematic examination of protein-protein interactions reveals that Sip1 can directly and specifically interact with several known splicing factors. This protein-protein interaction profile is distinct from that of SR proteins such as SC35 and ASF/SF2. Previous studies show that SC35 and ASF/SF2 interact with U1-70K, U2AF35, and several SR proteins, including p54 (63, 74). Similar to SC35 and ASF/SF2, Sip1 interacts with U1-70K, a component of U1 snRNP associated with the 5' splice site. However, of the proteins associated with the 3' splice site, Sip1 can interact with U2AF65 but not with U2AF35. Among the SR proteins, Sip1 can interact with SC35 and ASF/SF2 but not with p54. Sip1 is also different from U2AF35 in protein-protein interactions. Both Sip1 and U2AF35 interact with U2AF65 and with several SR proteins; however, Sip1 interacts with U1-70K, whereas U2AF35 does not directly interact with U1-70K (63). Consistent with the finding that Sip1 interacts with multiple spliceosomal proteins, we have observed that anti-Sip1 antibody but not the preimmune serum can efficiently precipitate splicing complexes containing pre-mRNA and splicing products (74a), indicating that Sip1 protein is associated with the spliceosome.

Although U2AF65 could not be directly immunoprecipitated by the anti-Sip1 antibody, the level of U2AF65 was reduced in the Sip1-immunodepleted nuclear extracts. Addition of U2AF65 alone did not have any detectable effect on splicing activity of the Sip1-depleted extract; however, addition of both Sip1 and U2AF65 proteins to the Sip1-depleted extract could restore the splicing activity to a level higher than that achieved by addition of Sip1 protein alone. The most likely explanation for this result is that Sip1 and U2AF65 are associated with each other in HeLa nuclear extract. Addition of both proteins, and perhaps other splicing factors, is required to fully restore the splicing activity of the depleted nuclear extract.

An essential role for Sip1 in pre-mRNA splicing. A role for Sip1 in splicing was first suggested by the observation that the anti-Sip1 antiserum could deplete splicing activity of HeLa nuclear extract. Purified recombinant Sip1 protein could restore the splicing activity to the Sip1-immunodepleted nuclear extract, demonstrating that Sip1 is indeed required for splicing. That Sip1 is a general splicing factor is supported by the observation that Sip1 is required for splicing of not only human β -globin pre-mRNA substrate but also other model substrates tested, including pPIP85.A (43, 74a). The exact role of Sip1 is not clear, but it appears to be distinct from that of other RS domain-containing splicing factors. This conclusion was based on our findings that the protein-protein interaction profile of Sip1 is different from those of other splicing factors and that other RS domain-containing proteins, including SC35, ASF/

SF2, and U2AF65, could not restore the splicing activity of the Sip1-immunodepleted extract.

The exact step at which Sip1 functions in the splicing reaction remains to be investigated. In the Sip1-immunodepleted extract, both the first step and the second step of the splicing reaction were blocked. It is possible that blockage of the second step is the consequence of the defect in the first step of the splicing reaction. Alternatively, Sip1 may be directly involved in both steps of the splicing reaction.

The ability of Sip1 to interact with U2AF65, U1-70K, and several SR proteins suggests a possible role of Sip1 in the earliest steps of spliceosome assembly. Indeed, in the Sip1-depleted nuclear extract, the formation of both A and B complexes is deficient. U2AF has been shown to be required for A-complex formation (51). Because Sip1 interacts with U2AF65, it is possible that the role of Sip1 in A-complex formation is due to its interaction with U2AF65. However, in the Sip1-depleted nuclear extract, a significant amount of U2AF65 is detected, although less than in the mock-depleted extract. This finding suggests that Sip1 may have a direct role in A-complex formation in addition to its interaction with U2AF. The finding that Sip1 interacts with both U1-70K and U2AF65 suggests that Sip1 may be involved in establishing interactions between the 5' and 3' splice sites during spliceosome assembly. It is also conceivable that by interacting with both U2AF65 and SR proteins, Sip1 may be able to form a bridge between U2AF65 and SR proteins associated with exon elements, thus mediating interactions across the exon in the exon definition model (4). Previous work has shown that U1 snRNP targets U2AF65 to the 3' splice site by interactions spanning the exon (24). Sip1 can interact with both U1-70K and U2AF65 and therefore may play a role in the network of interactions spanning the exon. Similarly, by interacting with U1-70K, SR proteins, and U2AF65, Sip1 may participate in the multiple protein-protein interactions across the intron (63). Interaction with U1-70K would allow Sip1 to influence recognition of the 5' splice site by the U1 snRNP, while interaction between Sip1 and U2AF65 may influence the recruitment of U2 snRNP to the branch site.

SR proteins are critical for pre-mRNA splicing by acting at multiple steps of spliceosome assembly. Previous studies have demonstrated that SR proteins can enhance interaction of U1 snRNP with the 5' splice site (14, 26, 27, 55, 68) and that at high concentrations, SR proteins can abrogate the requirement for U1 snRNP in pre-mRNA splicing (11, 57, 58). In addition, SR proteins can recruit U2 snRNP to the branch site of a pre-mRNA which lacks a 5' splice site (56) and escort U4/6.U5 tri-snRNP to the spliceosome (49). It is therefore possible that by interacting with the SR proteins, Sip1 can also function at multiple steps of spliceosome assembly.

ACKNOWLEDGMENTS

We thank R. Brent for providing the yeast two-hybrid system, T. Maniatis and P. Zuo for providing anti-U2AF65 antiserum, R.-M. Xu for helpful suggestions about protein purification, S. Noblitt for technical assistance, and D. Black, J. Bruzik, A. Goate, and Y. Rao for critical reading of the manuscript.

This work was supported by a grant from NIH (RO1 GM53945), an institutional grant through Washington University from Howard Hughes Medical Institute (76296-538202), and a special fellowship from the Leukemia Society of America.

REFERENCES

- Amrein, H., M. L. Hedley, and T. Maniatis. 1994. The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by Tra2. *Cell* **76**:735-746.
- Beggs, J. D., and P. E. Hodges. 1994. U₂ fulfills a commitment. *Curr. Biol.* **4**:264-267.
- Bennett, M., S. Michaud, J. Kingston, and R. Reed. 1992. Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes Dev.* **6**:1986-2000.
- Bennett, M., and R. Reed. 1992. Correspondence between a mammalian spliceosome component and an essential yeast splicing factor. *Science* **262**:105-108.
- Berget, S. M. 1995. Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**:2411-2414.
- Blencowe, B. J., J. A. Nickerson, R. Issner, S. Penman, and P. A. Sharp. 1994. Association of nuclear matrix antigens with exon-containing splicing complexes. *J. Cell Biol.* **127**:593-607.
- Blencowe, B. J., R. Issner, J. Kim, P. McCaw, and P. A. Sharp. 1995. New proteins related to the Ser-Arg family of splicing factors. *RNA* **1**:852-865.
- Bruzik, J. P., and T. Maniatis. 1995. Enhancer-dependent interaction between 5' and 3' splice sites in trans. *Proc. Natl. Acad. Sci. USA* **92**:7056-7059.
- Cáceres, J. F., and A. R. Krainer. 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J.* **12**:4715-4726.
- Cavaloc, Y., M. Popielarz, J. P. Fuchs, R. Gattoni, and J. Stevenin. 1994. Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. *EMBO J.* **13**:2639-2649.
- Chaudhary, N., C. McMahon, and G. Blobel. 1991. Primary structure of a human arginine-rich nuclear protein that colocalizes with spliceosome components. *Proc. Natl. Acad. Sci. USA* **88**:8189-8193.
- Chiara, M. D., and R. Reed. 1995. A two-step mechanism for 5' and 3' splice-site pairing. *Nature* **375**:510-513.
- Crispino, J. D., B. J. Blencowe, and P. A. Sharp. 1994. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science* **265**:1866-1869.
- Denhez, F., and R. Lafyatis. 1994. Conservation of regulated alternative splicing and identification of functional domains in vertebrate homologs to *Drosophila* splicing regulator, suppressor of white apricot. *J. Biol. Chem.* **269**:16170-16179.
- Diamond, R. H., K. Du, V. M. Lee, K. L. Mohn, B. A. Haber, D. S. Tewari, and R. Taub. 1993. Novel delayed-early and highly insulin-induced growth response genes: identification of HRS, a potential regulator of alternative pre-mRNA splicing. *J. Biol. Chem.* **268**:15185-15192.
- Eperon, I. C., D. C. Ireland, R. A. Smith, A. Mayeda, and A. R. Krainer. 1993. Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO J.* **12**:3607-3617.
- Fetzer, S., L. Jurgen, C. L. Will, and R. Lührmann. 1997. The U4/U6.U5 tri-snRNP-specific 27K protein is a novel SR protein that can be phosphorylated by the snRNP-associated protein kinase. *RNA* **3**:344-355.
- Fu, X.-D., and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**:437-441.
- Fu, X.-D., and T. Maniatis. 1992. Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science* **256**:535-538.
- Fu, X.-D. 1995. The superfamily of arginine/serine rich splicing factors. *RNA* **1**:663-680.
- Ge, H., P. Zuo, and J. L. Manley. 1991. Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. *Cell* **66**:373-382.
- Goralski, T. J., J.-E. Edström, and B. S. Baker. 1989. The sex determination locus transformer-2 of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell* **56**:1011-1018.
- Green, M. R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* **7**:559-599.
- Guthrie, C. 1991. Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* **253**:157-163.
- Hedley, M.-L., H. Amrein, and T. Maniatis. 1995. An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. *Proc. Natl. Acad. Sci. USA* **92**:11524-11528.
- Hoffman, B. E., and P. J. Grabowski. 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes Dev.* **6**:2554-2568.
- Horowitz, D. S., and A. R. Krainer. 1997. A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. *Genes Dev.* **11**:139-151.
- Jamison, S. F., Z. Pasman, J. Wang, C. Will, R. Lührmann, J. L. Manley, and M. A. Garcia-Blanco. 1995. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. *Nucleic Acids Res.* **23**:3260-3267.
- Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Lührmann, M. A. Garcia-Blanco, and J. L. Manley. 1994. Protein-protein interactions and 5' splice site recognition in mammalian mRNA precursors. *Nature* **368**:119-124.
- Konarska, M. M., and P. A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46**:845-855.
- Konarska, M. M., and P. A. Sharp. 1987. Interactions between snRNPs in formation of spliceosomes. *Cell* **49**:763-774.
- Kozak, M. 1989. The scanning model for translation. *J. Cell Biol.* **108**:229-241.

31. **Kraimer, A. R., and T. Maniatis.** 1985. Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. *Cell* **42**:725-736.
32. **Kraimer, A. R., G. C. Conway, and D. Kozak.** 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.* **4**:1158-1171.
33. **Kraimer, A. R., G. C. Conway, and D. Kozak.** 1990. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**:35-42.
34. **Kraimer, A. R., A. Mayeda, D. Kozak, and G. Binns.** 1991. Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and *Drosophila* splicing regulators. *Cell* **66**:383-394.
35. **Krämer, A., P. Legrain, F. Mulhauser, K. Gröning, R. Brosi, and G. Bilbe.** 1994. Splicing factor SF3a60 is the mammalian homologue of PRP9 of *S. cerevisiae*: the conserved zinc finger-like motif is functionally exchangeable in vivo. *Nucleic Acids Res.* **22**:5223-5228.
36. **Krämer, A., F. Mulhauser, C. Wersig, K. Gröning, and G. Bilbe.** 1995. Mammalian splicing factor SF3a represents a new member of the surp family of proteins and is homologous to the essential splicing factor PRP21p of *S. cerevisiae*. *RNA* **1**:260-272.
37. **Krämer, A.** 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* **65**:367-409.
38. **Laubler, J., P. Fabrizio, S. Teigelkamp, W. S. Lane, E. Hartmann, and R. Lührmann.** 1996. The HeLa 200 kDa U5 snRNP-specific protein and its homologue in *S. cerevisiae* are members of the DEXH-box protein family of putative RNA helicases. *EMBO J.* **15**:4001-4015.
39. **Lerner, E. A., M. R. Lerner, C. A. Janeway, and J. A. Steitz.** 1981. Monoclonal antibodies to nucleic acid-containing cellular constituents; probes for molecular biology and autoimmune disease. *Proc. Natl. Acad. Sci. USA* **78**:2737-2741.
40. **Li, H., and P. M. Bingham.** 1991. Arginine/serine-rich domains of the su(wa) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell* **67**:335-342.
41. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. *Genes Dev.* **10**:1569-1579.
42. **McKeown, M., J. M. Belote, and B. Baker.** 1987. A molecular analysis of tra, a gene in *D. melanogaster* that controls female sexual differentiation. *Cell* **48**:489-499.
43. **Moore, J. M., and P. A. Sharp.** 1992. Site-specific modification of pre-mRNA: the 2' hydroxyl groups at the splice sites. *Science* **256**:992-997.
44. **Moore, J. M., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursor to messenger RNAs by the spliceosome, p. 303-358. *In* R. F. Gesteland and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
45. **Patton, J. G., S. A. Mayer, P. Tempst, and B. Nadal-Ginard.** 1991. Characterization and molecular cloning of a polypyrimidine-tract binding protein: a component of a complex necessary for pre-mRNA splicing. *Genes Dev.* **5**:1224-1236.
46. **Query, C. C., R. C. Bentley, and J. D. Keene.** 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* **57**:89-101.
47. **Reed, R.** 1996. Initial splice site recognition and base pairing during pre-mRNA splicing. *Curr. Opin. Genet. Dev.* **6**:215-220.
48. **Rio, D. C.** 1993. Splicing of pre-mRNA: mechanism, regulation and role in development. *Curr. Opin. Genet. Dev.* **3**:574-584.
49. **Roscigno, R. F., and M. A. Garcia-Blanco.** 1995. SR proteins escort U4/6.U5 tri-snRNP to the spliceosome. *RNA* **1**:692-706.
50. **Roth, M. B., C. Murphy, and J. G. Gall.** 1990. A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. *J. Cell Biol.* **111**:2217-2223.
51. **Ruskin, B., P. D. Zamore, and M. R. Green.** 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52**:207-219.
52. **Screaton, G. R., J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G. Jackson, J. I. Bell, and A. R. Kraimer.** 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J.* **14**:4336-4349.
53. **Spikes, D. A., J. Kramer, P. Bingham, and K. Van Doren.** 1994. SWAP family pre-mRNA splicing regulators are a novel, ancient protein family sharing a highly conserved sequence motif with prp21 family of constitutive splicing factors. *Nucleic Acids Res.* **22**:4510-4519.
54. **Spritz, R. A., K. Strucnk, C. S. Surowy, S. O. Hoch, D. E. Barton, and U. Francke.** 1987. The human U1-70k snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing, and RNA-binding. *Nucleic Acids Res.* **15**:10373-10391.
55. **Staknis, D., and R. Reed.** 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol. Cell. Biol.* **14**:7670-7682.
56. **Tacke, R., Y. Chen, and J. L. Manley.** 1997. Sequence specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. *Proc. Natl. Acad. Sci. USA* **94**:1148-1153.
57. **Tarn, W. Y., and J. A. Steitz.** 1994. SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes Dev.* **8**:2704-2717.
58. **Tarn, W. Y., and J. A. Steitz.** 1995. Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. *Proc. Natl. Acad. Sci. USA* **92**:2504-2508.
59. **Valcarcel, J., and M. R. Green.** 1996. The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.* **21**:296-301.
60. **Valcarcel, J., R. K. Gaur, R. Singh, and M. R. Green.** 1996. Interaction of U2AF65 RS region with pre-mRNA of branch point and promotion of base pairing with U2 snRNA. *Science* **273**:1706-1709.
61. **Vellard, M., A. Sureau, J. Soret, C. Martinerie, and B. Perbal.** 1992. A potential splicing factor is encoded by the opposite strand of the trans-spliced c-myc exon. *Proc. Natl. Acad. Sci. USA* **89**:2511-2515.
62. **Wahle, E., and W. Keller.** 1992. The biochemistry of 3' end cleavage and polyadenylation of messenger RNA precursors. *Annu. Rev. Biochem.* **61**:419-440.
63. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**:1061-1070.
64. **Xiao, S. H., and J. L. Manley.** 1997. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev.* **11**:334-344.
65. **Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden.** 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. USA* **93**:6975-6980.
66. **Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth.** 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6**:837-847.
67. **Zahler, A. M., K. M. Neugebauer, J. A. Stolk, and M. B. Roth.** 1993. Human SR proteins and isolation of a cDNA encoding SRp75. *Mol. Cell. Biol.* **13**:4023-4028.
68. **Zahler, A. M., and M. B. Roth.** 1995. Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites. *Proc. Natl. Acad. Sci. USA* **92**:2642-2646.
69. **Zamore, P. D., and M. R. Green.** 1989. Identification, purification and characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc. Natl. Acad. Sci. USA* **86**:9243-9247.
70. **Zamore, P. D., and M. R. Green.** 1991. Biochemical characterization of U2 snRNP auxiliary factor: an essential pre-mRNA splicing factor with a novel intracellular distribution. *EMBO J.* **10**:207-214.
71. **Zamore, P. D., J. G. Patton, and M. R. Green.** 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355**:609-614.
72. **Zervos, A. S., J. Gyuris, and R. Brent.** 1993. Mix1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. *Cell* **72**:223-232.
73. **Zhang, M., P. D. Zamore, M. Carmo-Fonseca, A. I. Lamond, and M. R. Green.** 1992. Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc. Natl. Acad. Sci. USA* **89**:8769-8773.
74. **Zhang, W.-J., and J. Y. Wu.** 1996. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol. Cell. Biol.* **16**:5400-5408.
- 74a. **Zhang, W.-J., and J. Y. Wu.** Unpublished result.
75. **Zuo, P., and J. L. Manley.** 1993. Functional domains of the human splicing factor ASF/SF2. *EMBO J.* **12**:4727-4737.
76. **Zuo, P., and J. L. Manley.** 1994. The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5' splice sites. *Proc. Natl. Acad. Sci. USA* **91**:3363-3367.
77. **Zuo, P., and T. Maniatis.** 1996. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes Dev.* **10**:1356-1368.