The SRm160/300 splicing coactivator subunits

**BENJAMIN J. BLENCOWE,^{1,2} GÖRAN BAURÉN,^{1,4} ADAM G. ELDRIDGE,^{1,2,5}
ROBBYN ISSNER,^{1,6} JEFFREY A. NICKERSON,³ EMANUEL ROSONINA,²
and PHILLIP A. SHARP¹**

¹Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

²Banting and Best Department of Medical Research, C.H. Best Institute, University of Toronto,
Toronto, Ontario M5G 1L6, Canada

³Department of Cell Biology, University of Massachusetts Medical School, Worcester,
Massachusetts 01655, USA

ABSTRACT

The SRm160/300 splicing coactivator, which consists of the serine/arginine (SR)-related nuclear matrix protein of 160 kDa and a 300-kDa nuclear matrix antigen, functions in splicing by promoting critical interactions between splicing factors bound to pre-mRNA, including snRNPs and SR family proteins. In this article we report the isolation of a cDNA encoding the 300-kDa antigen and investigate the activity of it and SRm160 in splicing. Like SRm160, the 300-kDa antigen contains domains rich in alternating S and R residues but lacks an RNA recognition motif; the protein is accordingly named “SRm300.” SRm300 also contains a novel and highly conserved N-terminal domain, several unique repeated motifs rich in S, R, and proline residues, and two very long polyserine tracts. Surprisingly, specific depletion of SRm300 does not prevent the splicing of pre-mRNAs shown previously to require SRm160/300. Addition of recombinant SRm160 alone to SRm160/300-depleted reactions specifically activates splicing. The results indicate that SRm160 may be the more critical component of the SRm160/300 coactivator in the splicing of SRm160/300-dependent pre-mRNAs.

Keywords: nuclear matrix; snRNP; spliceosome; SR protein

INTRODUCTION

A fundamental difference between the splicing processes of eukaryotes is the requirement for a multitude of proteins that contain domains rich in alternating arginine and serine residues (RS domains). These proteins are highly conserved in metazoans, but, with few exceptions, are absent in *Saccharomyces cerevisiae*. Proteins that contain RS domains can be broadly divided into two subgroups: those of the “SR family” and SR proteins that are structurally and functionally distinct from the SR family (referred to below as “SR-related” proteins) (Fu, 1995; Manley & Tacke, 1996; Valcarcel & Green, 1996; Blencowe et al., 1999). Both

groups function at multiple stages during spliceosome formation. For example, they function early to promote the stable recognition of splice sites in constitutively spliced pre-mRNAs, and during the regulation of splice site selection in alternatively spliced pre-mRNAs. They also function at subsequent stages of spliceosome assembly, for example, by facilitating the recruitment of the U4/U6.U5 tri-snRNP particle to pre-mRNA (Roscigno & Garcia-Blanco, 1995).

SR family proteins contain one or two RNA recognition motifs (RRMs) and a phosphorylated C-terminal RS domain (Fu, 1995; Manley & Tacke, 1996; Valcarcel & Green, 1996). Members of the SR family include SRp20, ASF/SF2, SC35, 9G8, SRp30c, SRp40, SRp55, SRp75, and a more distantly related protein, p54 (Zhang & Wu, 1996). Individual members of the SR family activate splicing in cytoplasmic S100 extracts, which contain sufficient concentrations of factors required for splicing, except SR proteins (Krainer et al., 1990b; Zahler et al., 1992). They promote the preferential utilization of intron-proximal splice sites in pre-mRNAs containing *cis*-competing splice sites (Ge & Manley, 1990; Krainer et al., 1990a; Fu et al., 1992; Cáceres et al., 1994;

Reprint requests to: Phillip A. Sharp, Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; e-mail: sharppa@mit.edu.

⁴Present address: Department of Molecular Genome Research, Stockholm University, S-106 91 Stockholm, Sweden.

⁵Present address: Departments of Pathology and Microbiology, Stanford University School of Medicine, Stanford, California 94305, USA.

⁶Present address: Ariad Pharmaceuticals, Cambridge, Massachusetts 02139, USA.

Wang & Manley, 1995). They also promote the recognition of adjacent splice sites in pre-mRNAs by binding to specific sequences within exons called enhancers (Lavigne et al., 1993; Sun et al., 1993; Dirksen et al., 1994; Tanaka et al., 1994; Wang & Manley, 1995).

A large number of SR-related proteins have been identified that function in splicing (Fu, 1995; Lopez, 1998; Blencowe et al., 1999). These include proteins that are homologous to yeast splicing factors; the yeast homologs, however, lack the RS domains. Among these are the U1-snRNP-associated 70-kDa protein and several mammalian proteins that are members of the DEXD/H-box family of RNA-dependent ATPases, which are required for conformational rearrangements in the spliceosome during splicing, and for mRNA release from the spliceosome after splicing (de la Cruz et al., 1999). Many SR-related proteins, however, do not have known homologs or orthologs in yeast. Examples are regulators of alternative splicing that were initially defined genetically in *Drosophila*, such as suppressor of white apricot (SWAP), transformer (Tra), and transformer-2 (Tra-2). SWAP is a structurally and functionally highly conserved protein that autoregulates its expression by suppressing the splicing of its own pre-mRNA (Denhez & Lafyatis, 1994; Spikes et al., 1994). Tra and Tra-2 function as part of a cascade of alternative splicing events involved in the regulation of sex-determination in *Drosophila* (Inoue et al., 1992; Tian & Maniatis, 1993, 1994). Tra contains an RS domain but lacks an RRM, whereas Tra-2 contains a single RRM located between two RS domains. These proteins form a multisubunit complex with SR family proteins on an exonic splicing enhancer (ESE) sequence within exon 4 of the *doublesex* (*dsx*) pre-mRNA. Formation of this complex on exon 4 promotes the utilization of an upstream, sub-optimal, 3' splice site resulting in exon 4 inclusion and, ultimately, differentiation into female flies. In the absence of formation of this complex, exon 4 is excluded, leading to the male differentiation pathway. Recently, two mammalian homologs of *Drosophila* Tra-2 (hTra2 α and hTra2 β) have been identified (Dauwalder et al., 1996; Beil et al., 1997). These proteins preferentially bind to mammalian ESEs containing GAA-repeats and promote ESE-dependent splicing from these sequences in mammalian cell extracts (Tacke et al., 1998).

Recently, we identified a splicing coactivator, SRm160/300 (complex of the SR-related nuclear matrix protein of 160 kDa and a 300-kDa nuclear matrix antigen), that is important for the processing of a subset of constitutively spliced pre-mRNAs (Blencowe et al., 1998). SRm160/300 is also required for a mammalian ESE consisting of six GAA repeats to promote splicing of a pre-mRNA derived from exons 3 and 4 of the *Drosophila doublesex* gene (*dsx* pre-mRNA) (Eldridge et al., 1999). SRm160/300 likely promotes splicing of these substrates by forming multiple interactions with factors bound directly to a pre-mRNA; its association with pre-

mRNA requires U1 snRNP, SR proteins bound to an ESE, and is stabilized by U2 snRNP.

The 300-kDa subunit of the SRm160/300 coactivator was initially detected by the monoclonal antibody B4A11 as a nuclear matrix antigen concentrated in interchromatin granule clusters or "speckle domains" in vivo that is also preferentially associated with splicing complexes containing exon sequences in vitro (Blencowe et al., 1994). In the present study, we report the isolation of a cDNA encoding the 300-kDa subunit and demonstrate that it is a novel protein that, like SRm160, contains RS domains but lacks an RRM. Efficient depletion of the protein, named below as "SRm300," does not prevent splicing of pre-mRNAs previously shown to be dependent on SRm160/300. Consistent with this finding, we demonstrate that SRm160 alone activates splicing in reactions depleted of SRm160/300.

RESULTS

The 300-kDa nuclear matrix antigen recognized by mAb-B4A11 is a novel SR protein, SRm300, that lacks an RRM

The B4A11 nuclear matrix antigen was purified as a 300-kDa protein from HeLa nuclear extract as described in a previous study (see Fig. 4 in Blencowe et al., 1995) and was subjected to partial proteolysis with lysC. Microsequences obtained from two of the released peptides were used to search sequence databases, resulting in the identification of a human expressed sequence tag cDNA (EST 186680) with a predicted ORF containing an exact sequence match to one of the peptides. The insert of EST 186680 was used as a probe to screen a phage λ cDNA library prepared from the human U937 cell line (kind gift of J. Borrow), from which several overlapping clones were isolated. A fragment from the 5' end of the longest cDNA clone (B4A11.7) was used as a probe to isolate additional cDNA clones extending in the 5' direction. Multiple independent cDNA clones from both rounds of library screening were sequenced on both strands to determine the entire ORF sequence shown in Figure 1A. The corresponding cDNA sequence has been deposited in the GenBank database (accession number: AF201422). This cDNA sequence is probably complete because clones extending furthest in the 5' direction contain an in-frame ORF methionine codon preceded by stop codons in all three reading frames. The 3' end of the ORF terminates in a stop codon, followed by a predicted 3' UTR of 646 bases containing a consensus polyadenylation signal and a poly A tail. The predicted ORF contained both peptide sequences obtained from the purified protein (underlined in Fig. 1A). These and other data described below indicate that the predicted ORF in Figure 1A corresponds to the 300-kDa B4A11 nuclear matrix antigen, which

The predicted SRm300 ORF has a remarkably high content of serine (S), arginine (R), and proline (P) residues (23%, 17%, and 12%, respectively). In particular, it contains numerous clusters of SR dipeptides and two polyserine domains of unprecedented length (25 and

FIGURE 2. Association of SRm300 with splicing complexes. **A:** Total HeLa nuclear extract was separated on a 10% SDS-polyacrylamide gel and immunoblotted with an antigen affinity-purified polyclonal antiserum raised to a GST-SRm300 fusion protein containing amino acids 4–138 of SRm300 (rAb-SRm300, lane 1), and the corresponding preimmune serum (lane 2). **B:** Immunoprecipitation of splicing complexes with rAb-SRm300 from reactions incubated for 40 min containing PIP85A pre-mRNA. RNA recovered following immunoprecipitation (lanes 2–5) and RNA recovered directly from a parallel splicing reaction (lane 1) were separated on a 15% denaturing polyacrylamide gel. RNA loaded in lane 1 represents 25% of the total amount recovered, whereas RNA loaded in lanes 2–5 from each immunoprecipitation represents 50% of the total amounts recovered. Immunoprecipitations were performed with a nonspecific control Ab (rabbit anti-mouse, lane 2), mAb-B1C8 (lane 3), rAb pre-immune serum (lane 4), and rAb-SRm300 (lane 5). Bands corresponding to the intermediates and products of the splicing reaction are indicated.

41 residues, respectively). Both of these sequence features are similar to features of SRm160, which is S/R/P-rich and contains multiple SR repeats and polyserine stretches (Blencowe et al., 1998). Also similar to SRm160, SRm300 lacks an RNA recognition motif (RRM) typical of proteins belonging to the SR family of splicing factors. This is consistent with the observation that SRm160/300 does not associate with pre-mRNA in the absence of other splicing factors, including U1 snRNP and SR family proteins (Blencowe et al., 1998; Eldridge et al., 1999).

In addition to SR repeats and polyserine domains, SRm300 also contains three distinct types of consensus repeat sequences (Table 1). These consensus sequences are rich in S, R, P residues, and two of them also contain glycine (G) or glutamic acid (E) residues. They are distinct from the S, R, P-rich repeats present in SRm160 and also do not resemble repeats in other SR-related proteins that have recently been identified (Yuryev et al., 1996; Zhang & Wu, 1998). However, like the repeats in SRm160, the SRm300 repeat consen-

TABLE 1. Repetitive sequences in the SRm300 ORF.^a

Repeat consensus	Location ^b
(PA) RRGRSR (SRTP)	490, 528, 557, 568, 578, 609, 640, 652, 663, 675, 686, 698, 709, 718, 1800
(SX) SSPE (PK)	322, 743, 780, 819, 838, 913, 1218, 1421, 1460, 1482, 1499, 1579, 1598, 1618, 1672, 1691, 1729
SRS GSS(S) P	778, 797, 890, 908, 829, 968, 988, 1008, 1439, 1458, 1577, 1596, 1707

^aThe SRm300 ORF sequence was analyzed for statistically significant repetitive sequences using the SAPS program (Brendel et al., 1992). Many repetitive sequences scored as significant, consisting of separated, tandem, or overlapping repeats. Only consensus sequences of repeats occurring 10 or more times are listed; the corresponding sequences are boxed in Figure 1. Consensus sequences highlighted in bold type are present in all of the repeats listed, whereas the sequences in parentheses are present in only a subset of these repeats. X denotes any amino acid residue.

^bRepresented as the first amino acid residue in the repeat consensus (as boxed in Fig. 1A).

sus sequences contain potential phosphorylation sites for multiple kinases; in particular, the presence of multiple SR repeats is a strong indicator that SRm300 is a substrate for kinases belonging to the SRPK and Clk/Sty families, which phosphorylate serine residues in SR repeats (Gui et al., 1994a,b; Colwill et al., 1996). Consistent with this prediction, SRm300 is detected by the monoclonal antibody mAb104, which binds to phosphoserine residues within SR repeats, and is a substrate for SRPK1 in vitro (Blencowe et al., 1995; B.J. Blencowe & P.A. Sharp, unpubl. observations).

The ORF in Figure 1A, with the repetitive sequences masked, was used to search databases for related proteins. Although no previously characterized proteins were identified, the entire ORF matches predicted exon segments within a complete human genomic clone derived from chromosome 16 (accession number: AC004493; data not shown). Moreover, the ORF amino acids 970–1530 are identical to an ORF within a previously characterized cDNA from brain (Nagase et al., 1997). The N-terminal 159 amino acids of the ORF showed similarity to the N-terminal regions of hypothetical proteins from yeast to human (Fig. 1B). The percent identities of these sequences to the human protein are: *Drosophila melanogaster*, 45.5%, *Caenorhabditis elegans*, 38.7%, and *S. cerevisiae*, 22.2%. The ORF C-terminal to the homologous region in the *C. elegans* sequence diverged significantly from the human sequence (data not shown), whereas the *D. melanogaster* sequence C-terminal to this region has not been characterized. It should be noted that the predicted ORF identified in the *S. cerevisiae* genomic database is only six amino acid residues longer than the sequence shown in Figure 1B. These data indicate that the N-terminal region of SRm300 is highly conserved

in metazoans and is less related to a considerably smaller (hypothetical) protein in yeast. This feature of SRm300 is in contrast to SRm160, which contains a distinct N-terminal domain that is highly conserved in metazoans but absent in *S. cerevisiae* (Blencowe et al., 1998; Blencowe & Ouzounis, 1999).

Association of SRm300 with splicing complexes

SRm300 was previously shown, as the B4A11 nuclear matrix antigen, to: (1) associate with SRm160 and a subset of SR family proteins, (2) associate preferentially with splicing complexes containing exon sequences, and (3) concentrate in nuclear speckle structures (Blencowe et al., 1994, 1998). To verify that the predicted ORF in Figure 1A encodes a protein with these properties, a polyclonal antiserum (rAb-SRm300) was raised to a GST-fusion protein containing SRm300 ORF amino acids 4–138. The serum was affinity purified over the GST-SRm300 fusion protein and then used in immunoblotting, immunoprecipitation, and immunofluorescence microscopy experiments. The affinity-purified serum specifically recognized a single antigen of ~300 kDa in total HeLa cell nuclear extract (Fig. 2A, lane 1). Moreover, it specifically immunoprecipitated splicing complexes assembled on a constitutively spliced pre-mRNA, PIP85A, in an identical manner to mAb-

B1C8 (which is specific for SRm160), preferentially enriching for exon-containing complexes and not the lariat product-containing complex (Fig. 2B, lane 5; compare lanes 1, 3, and 5). The corresponding preimmune serum did not detect any nuclear protein (Fig. 2A, lane 2), nor did it immunoprecipitate splicing complexes (Fig. 2B, lane 4). rAb-SRm300, but not the preimmune serum (data not shown), also specifically immunolabeled speckles in interphase nuclei in a pattern very similar to that of mAb-B1C8 (Fig. 3A–C). However, in contrast to antibodies to SRm160, which specifically immunolabel the spindle poles and spindle apparatus in mitotic cells (Blencowe et al., 1998), rAb-SRm300 immunolabels in a diffuse pattern that is excluded by the spindle apparatus and condensed chromatin in mitotic cells (Fig. 3D–F). These properties of rAb-SRm300 confirm that the ORF in Figure 1A corresponds to the 300-kDa nuclear matrix antigen recognized by mAb-B4A11. The results also indicate that SRm160 and SRm300, although coincident in speckles in interphase cells, have distinct distributions during mitosis.

SRm300 is not required for splicing of constitutive and exonic enhancer-dependent pre-mRNAs

Previously, we demonstrated that SRm160/300 is required for the splicing of a subset of constitutively spliced

FIGURE 3. Immunolocalization of SRm300 during interphase and metaphase. **A–C:** Interphase human CaSki cells were double-immunolabeled with rAb-SRm300 (green) and mAb-B1C8 (red). **D–F:** Metaphase human CaSki cells were double-immunolabeled with rAb-SRm300 (green) and an antitubulin antibody (red). The immunolabeled cells were visualized by confocal microscopy. Each image corresponds to a single section proximal to the midline of the nucleus. The images were superimposed to reveal sites of overlap between the immunostaining patterns (yellow). Bar = 5 μ m.

pre-mRNAs, and also for activation of splicing by an ESE sequence consisting of six GAA repeats (Eldridge et al., 1999). This ESE promotes the splicing of a pre-mRNA derived from exons 3 and 4 of the *doublesex* gene of *Drosophila* ($dsx(GAA)_6$ pre-mRNA). However, these studies did not address whether one or both of the SRm160/300 subunits are important for splicing of these substrates. To investigate this, nuclear extract specifically immunodepleted of SRm300 with rAb-SRm300 was tested for splicing activity (Fig. 4). Unlike rAb-SRm160, which efficiently coimmunodepletes SRm300 with SRm160 (Blencowe et al., 1998, Fig. 3), immunodepletion with rAb-SRm300 efficiently depletes SRm300 without resulting in a significant reduction in the level of SRm160 (Fig. 4A, compare lanes 2 and 3). Immunoblot analysis of the pellet fractions from the rAb-SRm300 immunodepletion revealed that only a minor level of SRm160 was coimmunoprecipitated with rAb-SRm300 (data not shown). This indicates that, unlike rAb-SRm160, rAb-SRm300 may disrupt the SRm160/300 complex during immunodepletion. However, it is also possible that a significant fraction of SRm160 is in

excess of SRm300 in nuclear extract. In contrast to depletion of SRm160/300 with rAb-SRm160 (Fig. 4B,C, lanes 3), efficient depletion of SRm300 alone with rAb-SRm300 did not inactivate splicing of the PIP85A pre-mRNA (Fig. 4B, lane 5), or the $dsx(GAA)_6$ pre-mRNA (Fig. 4C, lane 5). These results suggest that SRm160, but not SRm300, may be the more important component of the SRm160/300 splicing coactivator for the splicing of these substrates (see Discussion).

SRm160 activates splicing in SRm160/300-depleted reactions

The above observations suggested that SRm160 alone might complement an extract depleted of both SRm160 and SRm300. To test if this is the case, a recombinant SRm160 (rSRm160) protein containing six N-terminal histidine residues was expressed in insect cells from a baculovirus vector and purified to homogeneity (refer to Materials and Methods). The final fraction from the purification contained a single polypeptide migrating at ~180 kDa (Fig. 5A, lane 1). This corresponds to the

FIGURE 4. Immunodepletion of SRm300 does not block splicing. **A:** Immunodepletion of SRm300. Nuclear extract immunodepleted with rAb-SRm300 (lane 3) or mock-depleted with preimmune serum (lane 2) was separated alongside a sample of untreated nuclear extract (lane 1) on a 7% SDS polyacrylamide gel and immunoblotted with mAb-B4A11 (specific for SRm300) and rAb-SRm160. **B:** Immunodepletion of SRm300 does not block splicing of PIP85A pre-mRNA. Splicing reactions incubated for 1 h with PIP85A pre-mRNA contained a regular (untreated) nuclear extract (lane 1), nuclear extracts depleted of SRm160/300 (lane 3) or SRm300 (lane 5), and nuclear extracts mock-depleted with preimmune sera (lanes 2 and 4) from rabbits immunized with GST-SRm160 and GST-SRm300 fusion proteins, respectively. RNA recovered from the splicing reactions was separated on a 15% denaturing polyacrylamide gel. The splicing reaction intermediates and products are indicated. **C:** Immunodepletion of SRm300 does not prevent splicing of an ESE-dependent pre-mRNA. Splicing reactions were incubated for 1 h with a pre-mRNA derived from exons 3 and 4 of the *Drosophila doublesex* gene, which contains a mammalian ESE consisting of six GAA repeats ($dsx(GAA)_6$ pre-mRNA). The reactions contained the same nuclear extracts as described in Figure 3B. RNA recovered from the splicing reactions was separated on a 7% denaturing polyacrylamide gel.

FIGURE 5. Reconstitution of SRm160/300-depleted splicing reactions with recombinant SRm160. **A:** Recombinant SRm160 containing six N-terminal histidine residues (rSRm160) was expressed in insect cells using a baculovirus vector, purified to homogeneity (see Materials and Methods), and was separated on a 12% SDS polyacrylamide gel and stained with Coomassie (lane 1). Lane 2 shows a sample of molecular weight standards (sizes are indicated in kilodaltons). **B:** rSRm160 promotes splicing of PIP85A pre-mRNA in SRm160/300-depleted reactions. Splicing reactions incubated for 1 h contained nuclear extract mock-depleted with preimmune serum (lane 1) or SRm160/300-depleted nuclear extract supplemented with 0, 1.5, or 3.0 μL of rSRm160 (~100 μg/μL) (lanes 2–4, respectively). 3.0 and 1.5 μL of protein buffer were added to the reactions in lanes 2 and 3, respectively, to control for buffer effects. RNA recovered from the splicing reactions was separated on a 15% denaturing polyacrylamide gel.

mobility of a fraction of SRm160 in HeLa nuclear extract that is highly phosphorylated, indicating that rSRm160 is also highly phosphorylated. Immunoblotting with antibodies to SRm160 confirmed the identity of the recombinant protein as SRm160 (data not shown).

Addition of increasing amounts of rSRm160 to splicing reactions containing nuclear extract immunodepleted of SRm160/300 activated splicing of PIP85A pre-mRNA (Fig. 5B, compare lane 1 with lanes 2–4). The highest level of splicing activity obtained upon addition of rSRm160 corresponded to approximately 30% of the level of splicing activity in a control reaction containing a mock-depleted nuclear extract (Fig. 5B, compare lane 4 with lane 1). Addition of higher amounts of rSRm160 resulted in a decrease in splicing efficiency (data not shown). As was observed previously for biochemically purified SRm160/300, rSRm160 did not activate splicing when added to a cytoplasmic S100 extract, confirming that it is distinct in activity from SR

family proteins, which activate splicing in these extracts (data not shown).

Addition of levels of rSRm160 to SRm160/300-depleted reactions that activated splicing of the PIP85A pre-mRNA did not activate splicing of the dsx(GAA)₆ pre-mRNA (data not shown). Although it cannot be excluded that an additional component besides SRm160 is missing from the depleted extract that is specifically required for ESE function, this seemed unlikely, as the immunodepletion procedure did not alter the level of any SR family or hTra 2 protein detected by the anti-SR protein antibody mAb104 (Blencowe et al., 1998, Fig. 3; data not shown). Moreover, it was found that levels of rSRm160 that activated PIP85A splicing in the SRm160/300 depleted reaction were inhibitory to splicing of the dsx(GAA)₆ pre-mRNA in a mock-depleted nuclear extract. These findings suggest that splicing of the dsx(GAA)₆ pre-mRNA may be difficult to reconstitute in SRm160/300-depleted reactions because this substrate is particularly sensitive to altered ratios of SRm160 to other splicing factors. Nevertheless, the results demonstrate that specific depletion of SRm160, but not SRm300, prevents both constitutive and ESE-dependent splicing, and that rSRm160 alone can activate splicing in the absence of SRm300.

DISCUSSION

The 300-kDa subunit of the SRm160/300 splicing coactivator is identified as a novel protein with a remarkably high content of serine (S), arginine (R), and proline (P) residues and is named "SRm300." Similar to SRm160, SRm300 contains numerous clusters of SR dipeptide repeats but lacks an RRM. SRm300 contains several unique sequence features, including a highly conserved N-terminal domain that lacks homology with previously characterized proteins, many copies of three different consensus S/R/P-rich motifs (Table 1), and two long polyserine tracts. Surprisingly, in contrast to depletion of SRm160/300, specific depletion of SRm300 did not prevent splicing of the PIP85A pre-mRNA, nor splicing of an ESE-dependent pre-mRNA derived from the *Drosophila doublesex* gene (dsx(GAA)₆; see Introduction). When added alone to splicing reactions, recombinant SRm160 specifically activated splicing. The results provide evidence that SRm160 is the more important component of the SRm160/300 complex in splicing reactions in vitro.

SRm300 is remarkable for its high content of SR dipeptide repeats and other repetitive sequence features. The presence of extensive SR repeat-rich domains in SRm300 strongly suggests that it is capable of interacting simultaneously with many factors, including SRm160 and other SR proteins. Such a function could be important for the organization of SR proteins and other splicing factors within the cell nucleus (see below). The two very long polyserine tracts in SRm300

are reminiscent of serine-rich domains that have been identified in other nuclear proteins, including Nopp140, SCAF10, and RNPS1. Nopp140 is a nucleolar phosphoprotein that appears to shuttle along "tracks" between the nucleolus and the cytoplasm (Meier & Blobel, 1992). SCAF10 is an SR-related cyclophilin concentrated in speckles that is associated with the carboxyl terminal domain of RNA polymerase II large subunit (Bourquin et al., 1997). RNPS1 is a recently identified splicing factor that contains an RRM but lacks an RS domain (Mayeda et al., 1999). Like SRm160/300, RNPS1 stimulates splicing in reactions containing limiting concentrations of SR family proteins. However, unlike SRm160/300, which is required for the splicing of specific pre-mRNA substrates, RNPS1 appears to function as a more general activator of splicing.

The results in the present study extend our earlier observations indicating that SRm160/300 functions as a coactivator in the splicing of specific constitutively spliced pre-mRNAs and also an ESE-dependent pre-mRNA. In the previous studies, it was demonstrated that the association of SRm160/300 with a pre-mRNA requires U1 snRNP, SR proteins bound to an ESE, and is stabilized by U2 snRNP (Blencowe et al., 1998; El-dridge et al., 1999). Moreover, in the absence of pre-mRNA, it was found that SRm160/300 interacts weakly but specifically with U2 snRNP and with a subset of SR proteins, including the ESE-binding protein hTra2 β (El-dridge et al., 1999). The results of the present study suggest that, in the absence of SRm300, SRm160 may mediate critical interactions between these components bound to pre-mRNA.

Although we cannot exclude the possibility that a minor level of SRm300 remaining in the SRm300- or SRm160/300-depleted reactions is important for splicing, the results indicate that SRm160 is the more rate-limiting component for the splicing of SRm160/300-dependent substrates. However, the observation that SRm300 is stably associated with SRm160 and with splicing complexes through both steps of the splicing reaction on different pre-mRNA substrates suggests that it participates in splicing. It is possible that other substrates are more dependent on SRm300 and/or that SRm300 has functions in splicing not detected *in vitro*. Several kinases have been identified recently, including members of the SRPK and Clk/Sty families, which phosphorylate RS domains and are known to influence protein-protein interactions between SR proteins and splicing activity (Misteli, 1999). These kinases could also influence critical interactions in splicing involving SRm160/300, which are both phosphoproteins and contain phosphorylated RS domains (Blencowe et al., 1995). Such a mechanism could operate at a more indirect level in the case of SRm300. For example, changes in the level of phosphorylation of SRm300, which is tightly bound to the nuclear matrix in speckle structures, could influence the nuclear distribution of

SRm160 and/or other associated SR proteins, and thereby modulate splicing activity by regulating the availability of these proteins at sites of nearby transcription. Similarly, the upregulation of kinases, such as members of the SRPK family during mitosis (Gui et al., 1994a), may account for the separate distributions of the SRm160/300 subunits observed in metaphase cells (see Fig. 3). This apparent change in the association of the SRm160 and SRm300 proteins could be important for the regulation of splicing activity during the cell cycle. These and other possibilities will require more detailed investigations of the functions of the SRm160/300 splicing coactivator subunits.

MATERIALS AND METHODS

SRm300 purification

SRm160/300 proteins were purified as described by Blencowe et al. (1995) and separated on a 10% SDS polyacrylamide gel. Following Coomassie staining, SRm300 was excised from the gel and digested *in situ* with lysC. The released peptides were microsequenced by the Massachusetts Institute of Technology (MIT) biopolymers facility.

cDNA isolation and sequencing

Sequence analysis of EST clone 186680 and phage λ cDNA clones (B4A11.2, B4A11.7, and B4A11-AE.1-6) isolated from a library prepared from U937 cell mRNA (gift of J. Borrow, MIT) allowed the assembly of a 7.6-kb contig containing the entire SRm300 ORF (see Fig. 1). All of the cDNA clones were sequenced on both strands using *ampliTaq* DNA polymerase reactions and an ABI PRISM automated sequencing machine (Perkin Elmer) as per the manufacturer's instructions. Sequencing was performed by the MIT Center for Cancer Research Biopolymers Facility. Assembly of contigs and the analysis of cDNA sequences were performed using software from DNASTAR (Madison, Wisconsin).

Antibodies

A cDNA fragment encoding SRm300 ORF amino acids 4–138 was generated by PCR amplification from B4A11-AE.1. and inserted into the pGEX-KG expression vector (Pharmacia) to generate pGST-SRm300. The GST-SRm300 fusion protein was expressed and purified from cell lysates on glutathione-Sepharose (Pharmacia), as per the manufacturer's instructions. Female New Zealand white rabbits were immunized with GST-SRm300 according to standard immunization protocols by Covance Co. (Denver, Pennsylvania). Serum antibodies were affinity purified on immobilized GST-SRm300 fusion protein and eluted with high and low pH buffers, as described by Harlow and Lane (1988). Total serum immunoglobulin was purified on protein A Sepharose beads (Pharmacia) and eluted using Tris-glycine buffer (pH 3.0), also as described by Harlow and Lane (1988). mAb-B1C8 has been previously described by Blencowe et al. (1994) and Wan et al. (1994).

Nuclear extracts

HeLa nuclear extracts were prepared essentially as described by Dignam et al. (1983). Nuclear extracts depleted of SRm160/300 rAb-SRm160 were identical to those described in Blencowe et al. (1998). Depletion of SRm300 was performed with rAb-SRm300 using the same procedure as described for SRm160/300 in Blencowe et al. (1998).

Cloning and expression of recombinant SRm160 in Baculovirus

SRm160 cDNA was cloned into a pFASTBAC HT (His-tagged) vector, using three overlapping sequenced EST clones (EST 78873, EST 121604, and EST 138500). Recombinant virus was produced using the BAC-to-BAC baculovirus system as outlined by the manufacturer (Gibco BRL). One hundred microliters of virus (10^9 pfu/mL) were used to infect 1×10^7 Hi5 cells in 150 mm tissue culture plates. Cells were harvested 48 h after infection and lysed in 50 mM Tris (pH 8.5), 400 mM KCl, 1% NP40, 5 mM 2-mercaptoethanol, and 0.2 mM Pefabloc (Boehringer) by brief sonication and incubation on ice for 30 min. Cell debris was removed by centrifugation $20,000 \times g$ for 20 min. The sample was added to a Ni-NTA resin column (Gibco-BRL), and washed with 15 vol 20 mM Tris (pH 8.5), 1 M KCl, 10 mM imidazole, 5 mM 2-mercaptoethanol, and 10% (v/v) glycerol. Proteins were eluted with 5–10 vol 20 mM Tris (pH 8.5), 100 mM KCl, 100 mM imidazole, 5 mM 2-mercaptoethanol, and 10% (v/v) glycerol. One-half milliliter fractions were collected and analyzed. The SRm160-containing fractions were dialyzed for 3 h against 3×2 -L changes of SR dialysis buffer [SRDB: 65 mM KCl, 15 mM NaCl, 10 mM HEPES (pH 7.6), 1 mM EDTA, 2 mM DTT, 5 mM KF, 5 mM β -glycerophosphate, 0.2 mM Pefabloc (Boehringer)]. SRm160 was precipitated by adding MgCl_2 to a final concentration of 20 mM, incubated on ice for 2 h, and centrifuged at $13,000 \times g$ for 30 min. The resulting pellet was washed with SRDB containing 20 mM MgCl_2 and resuspended in Dignam buffer D (Dignam et al., 1983) containing 5% (v/v) glycerol. All purification steps were carried out at 4°C.

Splicing and immunoprecipitation assays

Immunoprecipitation of splicing complexes was performed exactly as described in Blencowe et al. (1994). Splicing assays using the PIP85A pre-mRNA were carried out as described by Blencowe et al. (1998). Splicing assays using the dsx(GAA)₆ pre-mRNA were performed as described by Eldridge et al. (1999). The dsx template is as described previously by Yeakley et al. (1996). The splicing reactions shown in Figure 2B were analyzed on a 15% denaturing polyacrylamide gel; dsx splicing reactions were analyzed on 7% denaturing polyacrylamide gels.

SDS-PAGE and immunoblot analysis

SDS-PAGE and immunoblots were performed essentially as described by Harlow and Lane (1988). The immunoblot in Figure 2A was developed using a secondary antibody con-

jugated to horseradish peroxidase and chemiluminescence detection (NEN), as per the manufacturer's instructions.

Immunocytochemistry

Immunolabeling of CaSki cells (CRL 1550) was performed exactly as described by Blencowe et al. (1998). The samples were examined by epifluorescence microscopy and confocal microscopy using a Leica TCS system.

ACKNOWLEDGMENTS

We are grateful to C. Burge and J. Bowman for providing help with sequence analysis and to E. Reifenberg, Yu Li, and Guo-qin Zhong for expert technical assistance. J. Augliera and R. Cook in the Biopolymers Laboratory at MIT are thanked for performing sequencing of SRm300 cDNAs and peptides. J. Yeakley and X.-D. Fu kindly provided the dsx pre-mRNA template. We also thank S. Penman for mAb-B1C8. S. McCracken and J. Bowman kindly made helpful comments on the manuscript. G. Baurén gratefully acknowledges fellowship support from the Wenner-Gren Center Foundation and The Swedish Foundation for International Cooperation in Research and Higher Education. This work was supported in part by a United States Public Health Service MERIT award R37-GM34277 and RO1-AI32486 from the National Institutes of Health to P.A.S. and partially by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute. This work was also supported in part by grants from the U.S. Department of Defense Breast Cancer Research Program and the Medical Research Council of Canada to B.J.B.

Received September 9, 1999; returned for revision October 18, 1999; revised manuscript received October 29, 1999

REFERENCES

- Beil B, Screaton G, Stamm S. 1997. Molecular cloning of htra2-beta-1 and htra2-beta-2, two human homologs of tra-2 generated by alternative splicing. *DNA Cell Biol* 16:679–691.
- Blencowe BJ, Bowman JAL, McCracken S, Rosonina E. 1999. SR-related proteins and the processing of messenger RNA precursors. *Biochem Cell Biol* 77:277–291.
- Blencowe BJ, Issner R, Kim J, McCaw P, Sharp PA. 1995. New proteins related to the Ser-Arg family of splicing factors. *RNA* 1:852–865.
- Blencowe BJ, Issner R, Nickerson JA, Sharp PA. 1998. A coactivator of pre-mRNA splicing. *Genes & Dev* 12:996–1009.
- Blencowe BJ, Nickerson JA, Issner R, Penman S, Sharp PA. 1994. Association of nuclear matrix antigens with exon-containing splicing complexes. *J Cell Biol* 127:593–607.
- Blencowe BJ, Ouzounis CA. 1999. The PWI motif: A new protein domain in splicing factors. *Trends Biochem Sci* 24:179–180.
- Bourquin JP, Stagljar I, Meier P, Moosmann P, Silke J, Baechli T, Georgiev O, Schaffner W. 1997. A serine/arginine-rich nuclear matrix cyclophilin interacts with the C-terminal domain of RNA polymerase II. *Nucleic Acids Res* 11:2055–2061.
- Brendel V, Bucher P, Nourbakhsh I, Blaisdell BE, Karlin S. 1992. Methods and algorithms for statistical analysis of protein sequences. *Proc Natl Acad Sci USA* 89:2002–2006.
- Cáceres JF, Stamm S, Helfman DM, Krainer AR. 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* 265:1706–1709.
- Colwill K, Pawson T, Andrews B, Prasad J, Manley J, Bell J, Duncan P. 1996. The Clk/Sty protein kinase phosphorylates SR splicing

- factors and regulates their intranuclear distribution. *EMBO J* 15:265–275.
- Dauwalder B, Amaya-Manzanares F, Mattox W. 1996. A human homolog of the *Drosophila* sex determination factor transformer-2 has conserved splicing regulatory functions. *Proc Natl Acad Sci USA* 93:9004–9009.
- de la Cruz J, Kressler D, Linder P. 1999. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *Trends Biochem Sci* 24:192–198.
- Denhez F, Lafyatis R. 1994. Conservation of regulated alternative splicing and identification of functional domains in vertebrate homologs to the *Drosophila* regulator, *suppressor-of-white-apricot*. *J Biol Chem* 269:16170–16179.
- Dignam JD, Lebovitz RM, Roeder RD. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* 11:1475–1489.
- Dirksen WP, Hampson RK, Sun Q, Rottman FM. 1994. A purine-rich exon sequence enhances alternative splicing of bovine growth hormone pre-mRNA. *J Biol Chem* 269:6431–6436.
- Eldridge AG, Li Y, Sharp PA, Blencowe BJ. 1999. The SRm160/300 splicing coactivator is required for exon-enhancer function. *Proc Natl Acad Sci USA* 96:6125–6130.
- Fu X-D. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* 1:663–680.
- Fu X-D, Mayeda A, Maniatis T, Krainer AR. 1992. General splicing factors SF2 and SC35 have equivalent activities in vitro, and both affect alternative 5' and 3' splice site selection. *Proc Natl Acad Sci USA* 89:11224–11228.
- Ge H, Manley JL. 1990. A protein factor, ASF, controls cell specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* 62:25–34.
- Gui J-F, Lane WS, Fu X-D. 1994a. A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* 369:678–682.
- Gui J-F, Tronchere H, Chandler SD, Fu X-D. 1994b. Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. *Proc Natl Acad Sci USA* 91:10824–10828.
- Harlow E, Lane D. 1988. *Antibodies: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Inoue K, Hoshijima K, Higuchi I, Sakamoto H, Shimura Y. 1992. Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of doublesex primary transcript for sex-specific RNA processing. *Proc Natl Acad Sci USA* 89:8092–8096.
- Krainer AR, Conway GC, Kozak D. 1990a. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* 62:35–42.
- Krainer AR, Conway GC, Kozak D. 1990b. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes & Dev* 4:1158–1171.
- Lavigne A, Branche HL, Kornblihtt A, Chabot B. 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes & Dev* 7:2405–2417.
- Lopez AJ. 1998. Alternative splicing of pre-mRNA: Developmental consequences and mechanisms of regulation. *Annu Rev Genet* 32:279–305.
- Manley J, Tacke R. 1996. SR proteins and splicing control. *Genes & Dev* 10:1569–1579.
- Mayeda A, Badolato J, Kobayashi R, Zhang MQ, Gardiner EM, Krainer AR. 1999. Purification and characterization of human RNPS1: A general activator of pre-mRNA splicing. *EMBO J* 18:4560–4570.
- Meier U, Blobel G. 1992. Nopp140 shuttles on tracks between nucleus and cytoplasm. *Cell* 70:127–138.
- Misteli T. 1999. RNA splicing: What has phosphorylation got to do with it? *Curr Biol* 9:R198–R200.
- Nagase T, Ishikawa K, Nakajima D, Ohira M, Seki N, Miyajima N, Tanaka A, Kotani H, Nomura N, Ohara O. 1997. Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. *DNA Res* 4:141–150.
- Roscigno R, Garcia-Blanco M. 1995. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* 1:692–706.
- Spikes DA, Kramer J, Bingham PM, Van Doren K. 1994. SWAP pre-mRNA splicing regulators are a novel, ancient protein family sharing a highly conserved sequence motif with the prp21 family of constitutive splicing proteins. *Nucleic Acids Res* 22:4510–4519.
- Sun Q, Mayeda A, Hampson RK, Krainer AR, Rottman FM. 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes & Dev* 7:2598–2608.
- Tacke R, Tohyama M, Ogawa S, Manley JL. 1998. Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. *Cell* 93:139–148.
- Tanaka K, Watakabe A, Shimura Y. 1994. Polypurine sequences within a downstream exon function as a splicing enhancer. *Mol Cell Biol* 14:1347–1354.
- Tian M, Maniatis T. 1993. A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. *Cell* 74:105–114.
- Tian M, Maniatis T. 1994. A splicing enhancer exhibits both constitutive and regulated activities. *Genes & Dev* 8:1703–1712.
- Valcarcel J, Green MR. 1996. The SR protein family: Pleiotropic functions in pre-mRNA splicing. *Trends Biochem Sci* 21:296–301.
- Wan K, Nickerson JA, Krockmalnic G, Penman S. 1994. The B1C8 protein is in the dense assemblies of the nuclear matrix and relocates to the spindle and pericentriolar filaments at mitosis. *Proc Natl Acad Sci USA* 91:594–598.
- Wang J, Manley JL. 1995. Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing in vivo in diverse ways. *RNA* 1:346–355.
- Yeakley JM, Morfin JP, Rosenfeld MG, Fu XD. 1996. A complex of nuclear proteins mediates SR protein binding to a purine-rich splicing enhancer. *Proc Natl Acad Sci USA* 93:7582–7587.
- Yuryev A, Patturajan M, Litingtung Y, Joshi RV, Gentile C, Gebara M, Corden JL. 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.
- Zahler AM, Lane WS, Stolk JA, Roth MB. 1992. SR proteins: A conserved family of pre-mRNA splicing factors. *Genes & Dev* 6:837–847.
- Zhang WJ, Wu JY. 1996. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Mol Cell Biol* 16:5400–5408.
- Zhang WJ, Wu JY. 1998. Sip1, a novel RS domain-containing protein essential for pre-mRNA splicing. *Mol Cell Biol* 18:676–684.