The EMBO Journal vol.14 no.14 pp.3540–3551, 1995

The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities

Roland Tacke[1] and James L.Manley

Department of Biological Sciences, Columbia University, New York, NY 10027, USA

[1]Corresponding author

ASF/SF2 and SC35 belong to a highly conserved family of nuclear proteins that are both essential for splicing of pre-mRNA *in vitro* and are able to influence selection of alternative splice sites. An important question is whether these proteins display distinct RNA binding specificities and, if so, whether this influences their functional interactions with pre-mRNA. To address these issues, we first performed selection/amplification from pools of random RNA sequences (SELEX) with portions of the two proteins comprising the RNA binding domains (RBDs). Although both molecules selected mainly purine-rich sequences, comparison of individual sequences indicated that the motifs recognized are different. Binding assays performed with the full-length proteins confirmed that ASF/SF2 and SC35 indeed have distinct specificities, and at the same time provided evidence that the highly charged arginine–serine region of each protein is not a major determinant of specificity. In the case of ASF/SF2, evidence is presented that binding specificity involves cooperation between the protein's two RBDs. Finally, we demonstrate that an element containing three copies of a high-affinity ASF/SF2 binding site constitutes a powerful splicing enhancer. In contrast, a similar element consisting of three SC35 sites was inactive. The ASF/SF2 enhancer can be activated specifically in splicing-deficient S100 extracts by recombinant ASF/SF2 in conjunction with one or more additional protein factors. These and other results suggest a central role for ASF/SF2 in the function of purine-rich splicing enhancers.

Key words: ASF/SC35/SF2/splicing enhancer/SR proteins

Introduction

Pre-mRNA splicing, the precise excision of introns from mRNA precursors, is a critical step in gene expression that requires accurate selection and pairing of corresponding 5′ and 3′ splice sites. The multi-component protein–RNA complex that catalyzes the splicing reaction, the spliceosome, consists of several small nuclear ribonucleoprotein particles (snRNPs) and a still unknown number of accessory proteins (for review, see Green, 1991; Moore *et al.*, 1993). Selection of splice sites occurs during assembly of these components on the pre-mRNA. Spliceosome constituents in higher eukaryotes that recognize the pre-mRNA during the early steps of spliceosome assembly include the U1 and U2 snRNPs, the RNA components of which can undergo base pairing interactions with the 5′ splice site and branch point sequence, respectively (Zhuang and Weiner, 1986, 1989; Wu and Manley, 1989), and the U2 auxiliary factor (U2AF), which binds specifically to the polypyrimidine tract and facilitates interaction of U2 snRNP with the branch site (Zamore and Green, 1989). Although U1 snRNP and U2AF are present in the earliest detectable prespliceosome complexes (Michaud and Reed, 1991, 1993; Jamison *et al.*, 1992), it is clear that interaction of these factors with the pre-mRNA is not sufficient to account for accurate intron specification and splice site selection.

Several observations suggest that an important role in these processes is played by members of a structurally and functionally highly related family of evolutionary conserved proteins, termed SR proteins (Zahler *et al.*, 1992), that includes the splicing factors ASF/SF2 (Ge *et al.*, 1991; Krainer *et al.*, 1991) and SC35 (Fu and Maniatis, 1992a). First, SR proteins are essential splicing factors, capable of complementing splicing-deficient S100 fractions for splicing activity *in vitro*. In particular, ASF/SF2 and SC35 are required for an early step in spliceosome assembly, the formation of the first specific ATP-dependent complex detected in the assembly pathway (Krainer and Maniatis, 1985; Krainer *et al.*, 1990a; Fu and Maniatis, 1992b). Second, SR proteins can function as alternative splicing factors, initially shown to influence the selection of alternative 5′ splice sites in a concentration-dependent manner *in vitro* (Ge and Manley, 1990; Krainer *et al.*, 1990b; Zahler *et al.*, 1993). A common feature of ASF/SF2 and SC35 is that they both promote the use of proximal 5′ and 3′ splice sites at high concentrations (Fu *et al.*, 1992). Other SR proteins behave similarly, although in some cases qualitatively and quantitatively different effects on 5′ splice site selection *in vitro* have been observed (Kim *et al.*, 1992; Zahler *et al.*, 1993). These similarities in the behavior of SR proteins raised the possibility that they may be functionally redundant. Finally, the observation that ASF/SF2 and SC35 are each sufficient to commit specific pre-mRNAs to splicing (Fu, 1993), together with recent findings indicating that high concentrations of SR proteins can circumvent the need for U1 snRNP in *in vitro* splicing (Crispino *et al.*, 1994; Tarn and Steitz, 1994), suggests that SR proteins interact with the pre-mRNA very early during spliceosome assembly.

The structure of SR proteins is characterized by a C-terminal domain rich in arginine–serine dipeptide repeats (RS region) and one or two N-terminally located RNP-type RNA binding domains (RBDs). RBDs, which typically encompass 80–90 amino acids with two highly conserved elements, designated RNP-1 and RNP-2, are found in a number of RNA binding proteins involved in various

aspects of mRNA biogenesis (for reviews, see Kenan *et al.*, 1991; Burd and Dreyfuss, 1994a). Amongst splicing factors, these include, besides SR proteins, U2AF (Zamore *et al.*, 1992) and regulators of alternative splicing in *Drosophila* (Amrein *et al.*, 1988; Goralski *et al.*, 1989; Sosnowski *et al.*, 1989). In several cases, proteins with RBDs have been shown to bind RNA with sequence specificity (for review, see Burd and Dreyfuss, 1994a). ASF/SF2 and SC35 each contain one classical RBD (Ge *et al.*, 1991; Krainer *et al.*, 1991; Fu and Maniatis, 1992a), and ASF/SF2 contains an additional degenerate RBD, located between the first RBD and the RS region. While the RBDs of ASF/SF2 have been shown previously to have RNA binding activities (Caceres and Krainer, 1993; Zuo and Manley, 1993, 1994), comparable information about SC35 and other SR proteins is lacking.

Although the significance of SR protein function in constitutive and alternative splicing is well documented, the molecular events that underly their various activities are just beginning to be understood. In agreement with their proposed role in early spliceosome assembly, ASF/SF2 and SC35 were shown to participate in protein–protein interactions with other essential splicing components. SC35, and perhaps other SR proteins, may help bridge U1 and U2 snRNPs via interaction with U1 70 kDa protein and U2AF (Wu and Maniatis, 1993), while ASF/SF2 can facilitate binding of U1 snRNP to the pre-mRNA 5′ splice site, presumably via direct interaction with the U1 snRNP 70 kDa protein (Kohtz *et al.*, 1994). Furthermore, the RS region of ASF/SF2, which is required for activity in S100 complementation assays (Caceres *et al.*, 1993; Zuo and Manley, 1993), is essential for the interaction with the U1 snRNP 70 kDa protein (Kohtz *et al.*, 1994).

Relatively little information is available regarding the potential target RNA sequences of SR proteins. While ASF/SF2 has been shown to recognize 5′ splice sites specifically in two different pre-mRNAs (Zuo and Manley, 1994), SR proteins including ASF/SF2 have also been implicated in alternative splicing events that rely on exonic sequences, designated splicing enhancer elements (Lavigueur *et al.*, 1993; Sun *et al.*, 1993). Such enhancer elements have been identified in a number of natural pre-mRNAs where they appear to be involved in the control of stage- or tissue-specific splicing events (Hoshijima *et al.*, 1991; Tian and Maniatis, 1992; Watakabe *et al.*, 1993, and references therein; Huh and Hynes, 1994). Some but not all splicing enhancers are characterized by the presence of purine-rich sequences (Watakabe *et al.*, 1993). ASF/SF2 can bind to a purine-rich splicing enhancer in the bovine growth hormone pre-mRNA, although the exact site of binding has not been determined (Sun *et al.*, 1993). SR proteins have also been shown to interact with another purine-rich enhancer (Lavigueur *et al.*, 1993), and to be present in the enhancer complex of the female-specific exon of the *Drosophila* doublesex pre-mRNA (Tian and Maniatis, 1993, 1994). Altogether, these observations seem to suggest that SR proteins can interact with a variety of different sequences, but to understand their mechanism of action it is necessary to know whether the different SR proteins have distinct, sequence-specific RNA binding activities.

Here we describe the selection of binding sites for ASF/SF2 and SC35. Our results indicate that ASF/SF2 and SC35 are sequence-specific RNA binding proteins with distinct specificities. In addition, ASF/SF2, but not SC35, recognizes sequences very similar to purine-rich elements found in various naturally occurring splicing enhancers. We show that the ASF/SF2 consensus recognition motif, but not the SC35 motif, can function as a splicing enhancer element, and that ASF/SF2 binding is necessary but not sufficient for its function.

Results

In order to investigate the RNA binding properties of ASF/SF2 and SC35, we used, in modified form, an approach designed for the identification of high-affinity binding sites for RNA binding molecules, SELEX (Tuerk and Gold, 1990). In short, histidine-tagged recombinant bacterial derivatives of ASF/SF2 and SC35 were purified, immobilized on a solid support via Ni-chelate binding and exposed to approximately equimolar amounts of a pool of 59 base long RNA molecules with 20 bases of random sequence, in the presence of a large excess of tRNA. Bound RNA was isolated and amplified by consecutive steps of reverse transcription, polymerase chain reaction (PCR) and transcription from a T7 promoter. After 7–9 cycles of selection/amplification the PCR products were cloned and the sequences of the selected regions of individual clones determined. The derivatives of ASF/SF2 and SC35 that were used, ASFΔRS and SC35ΔRS, each lacked its RS region. This was largely because it seemed likely that the high positive charge of this region might lead to non-specific interactions with RNA that would obscure any possible sequence-specific binding. This was suggested by previous experiments with both U2AF (Zamore *et al.*, 1992) and ASF/SF2 (Zuo and Manley, 1994), which showed that in each case the presence of the RS domain did appear to reduce RNA binding specificity. Our approach assumes that the RS domain does not contribute to sequence-specific binding, a view supported by the experiments below.

ASFΔRS recognizes purine-rich sequences

An initial SELEX experiment was performed with a truncated version of ASF/SF2 that contained both RBDs but lacked the RS region (ASFΔRS, Zuo and Manley, 1993, see Figure 1). After the final (7 th) cycle, 49 selected sequences were determined (Figure 2A and B). All were found to be highly enriched in purines and to contain at least one motif resembling either the octamer RGAAG-AAC (Figure 2A) or one of two related decamer sequences (AGGACAGAGC and AGGACGAAGC, Figure 2B). Each of the identified motifs contained three or less mismatches compared with its corresponding consensus sequences. The octamer motif was found in 22 sequences of which five contained the motif twice (Figure 2A). Motifs resembling one of the consensus decamers were identified in 36 sequences (Figure 2B). Of these, nine also contained an octamer whose sequence in most cases overlapped the decamer (e.g. see sequence A30 in Figure 2A and B).

The selected sequences should, by definition, bind ASFΔRS with higher affinities than randomly chosen sequences. To confirm this assumption by an independent

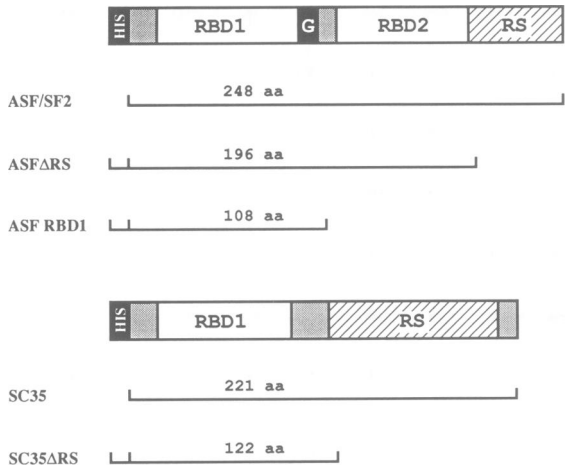

Fig. 1. Schematic representation of the recombinant proteins used in this study. Full-length ASF/SF2 and SC35, purified from baculovirus-infected SF21 cells, contain two and one RNA binding domain(s) (RBD), respectively, and a C-terminal RS region (RS). The two RBDs of ASF/SF2 are separated by a glycine-rich region (G). C-terminally truncated versions of ASF and SC35 were expressed in and purified from bacteria by virtue of an 11–12 amino acid histidine tag (HIS) at the N-terminus. The number of amino acids constituting each protein is indicated.

technique, we performed gel mobility shift experiments with four different ASFΔRS-selected sequences (A7, A14, A18, A39) and two control sequences taken from the random pool (C2, C13, see Figure 2D). A7 contains two overlapping octamer motifs, A14 an octamer and a decamer, A18 a decamer in a pyrimidine-rich context and A39 a variant of the octamer with a relatively poor match to the consensus. C2 is slightly enriched in purines and C13 strongly pyrimidine-rich. The sequences were presented to increasing amounts of ASFΔRS as ^{32}P-labeled RNA probes of ~59 nucleotides, preserving the same sequence context as in the selection procedure. As shown in Figure 3A, stable complexes were formed with all selected sequences but not with the control sequences. Strongest binding was observed with A14, which formed two complexes of different mobilities, suggesting that the A14 sequence may bind two molecules of ASFΔRS simultaneously (Figure 3A, lanes 4–6). In a competition assay containing constant amounts of ASFΔRS and ^{32}P-labeled A14 RNA plus increasing amounts of each of three different unlabeled competitors, self-competition was at least 25 times more efficient than was competition with the control sequence C2 (Figure 3B, compare lanes 6 and 14). A18 displayed an intermediate efficiency as a competitor (Figure 3B, lanes 7–11). These results indicate that ASFΔRS binds specifically to the selected sequences and support the idea that both the consensus octamer and the consensus decamer constitute high-affinity binding sites for ASF/SF2.

Cooperation between two RBDs determines the RNA binding specificity of ASF/SF2

As mentioned above, ASF/SF2 contains two RBDs, an N-terminal one that displays a very high match to the RBD consensus and a second, degenerate RBD that is considerably diverged from the consensus. The significance of each RBD has been confirmed by the effect of point mutations in key residues, and each RBD has been shown to have the ability to bind RNA by itself, albeit with reduced affinity (Caceres and Krainer, 1993; Zuo and Manley, 1993). One possibility was that the canonical RBD was largely responsible for determining specificity, with the second RBD performing an auxiliary role, perhaps to increase affinity. A prediction of this is that high-affinity sequences selected by the N-terminal RBD would be identical or similar to those selected by ASFΔRS. However, after seven cycles of selection/amplification with a truncated version of ASF/SF2 that contained only the first RBD (RBD1, see Figure 1), the selected sequences lacked any resemblance to either the ASF/SF2 octamer or decamer. Instead they contained motifs similar to the consensus heptamer ACGCGCA, sometimes as part of extended stretches of A/GC dinucleotide repeats (see e.g. R2, Figure 2C). These findings suggest that the two RBDs of ASF/SF2 may cooperate in some manner to determine the RNA binding specificity of the intact molecule.

ASF/SF2 and SC35 possess different RNA binding specificities

We next sought to identify optimal RNA binding sites for SC35. To this end, we performed SELEX with SC35ΔRS, essentially as described above for ASFΔRS. Visual inspection of 34 individual sequences after seven cycles of selection indicated that SC35ΔRS, similarly to ASFΔRS, preferentially selected purine-rich sequences, but a sequence motif that was shared by a majority of these sequences was not apparent (data not shown). In contrast, after two additional cycles, all sequences determined (33) contained at least one nonamer motif with high similarity (i.e. three or less mismatches) to one of two related consensus sequences, of which one was strongly purine-rich. Motifs resembling the purine-rich nonamer, AGSA-GAGTA (Figure 4A), were found in 24 selected sequences, while motifs similar to the second nonamer, GTTCG-AGTA, were identified in 11 selected sequences (Figure 4B). Two selected sequences (S19 and S37) contained both motifs. Re-examination of the individual sequences determined after seven cycles revealed that ~50% contained sequences similar to one of the two motifs. It is notable that the purine-rich motif resembles the consensus 5′ splice site (compare C/AAG|GTRAGT and AGSA-GAGTA). Mobility shift experiments performed as above with ASFΔRS confirmed that SC35ΔRS binds specifically to the selected sequences (data not shown).

Given that ASFΔRS and SC35ΔRS both recognize purine-rich sequences, it was important to know whether ASF/SF2 would bind sequences selected by SC35ΔRS and vice versa. In addition, we wished to test our assumption that the RS region does not contribute significantly to RNA binding specificity. To address these questions, we examined the RNA binding properties of full-length recombinant SC35 and ASF/SF2 purified from baculovirus-infected SF21 cells, again using gel mobility shift assays. As shown in Figure 5A, SC35 bound efficiently to the SC35ΔRS-selected sequences S16 and S33 but not to three different non-selected sequences, including an ASFΔRS sequence (A14), an ASF RBD1 sequence (R5) and a random sequence (C2) (see Figure 2). Most importantly, no stable complex was formed with A14 (Figure 5A, lanes 7–9), which contains both an ASF/SF2 octamer

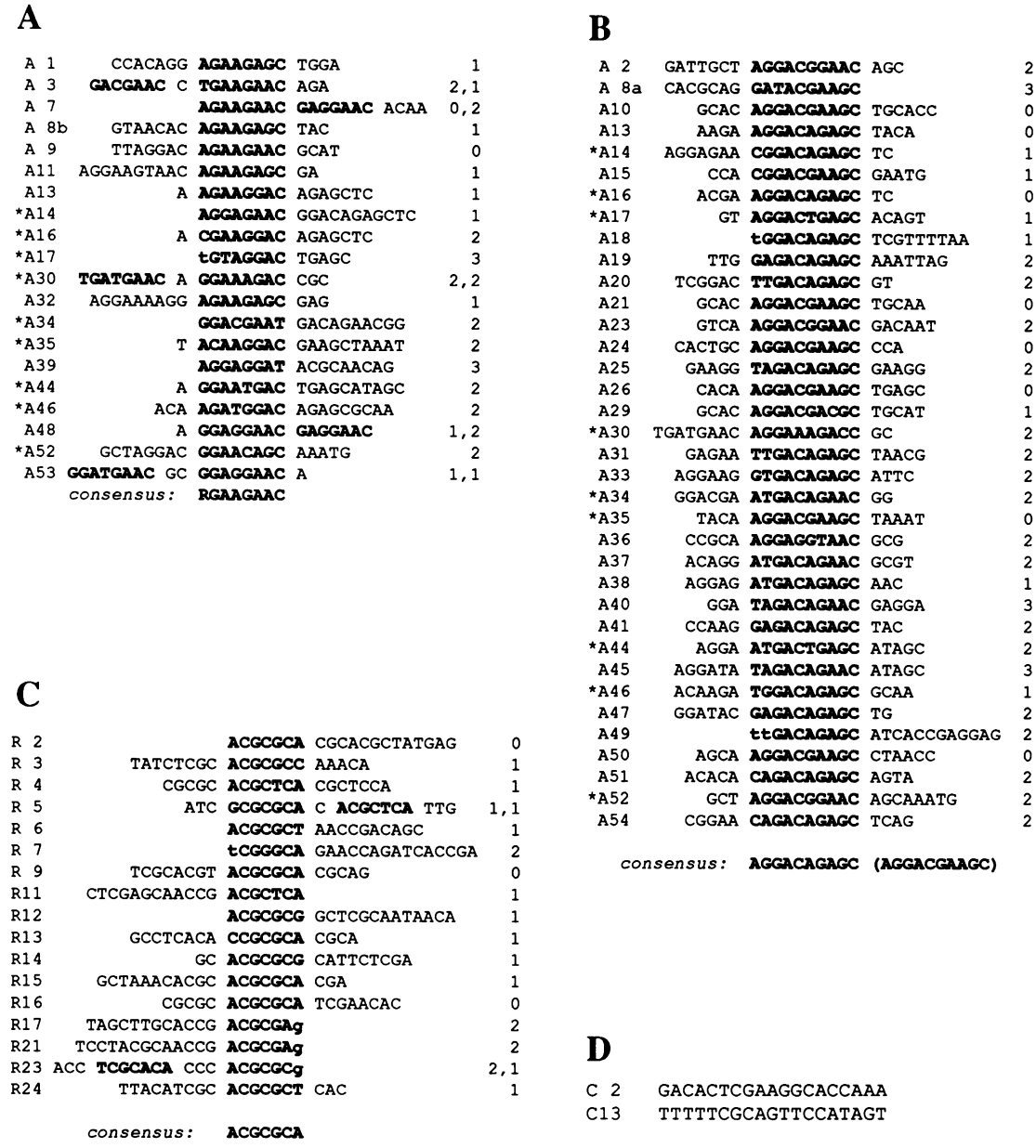

Fig. 2. Selection/amplification with ASFΔRS and ASF RBD1. Shown are the sequences of individual clones selected by ASFΔRS (**A, B**) and ASF RBD1 (**C**), and two sequences of the random pool (**D**). Motifs found in ASFΔRS-selected sequences resembling the octamer RGAACAAC (R = A or G) (A) or the decamer sequences AGGACAGAGC or AGGACGAAGC (B), and motifs in ASF RBD1-selected sequences resembling the heptamer ACGCGCA (C) are represented in bold type characters. Nucleotides that form part of such motifs but belong to the flanking constant region are shown in lower case characters. Numbers shown behind the sequences indicate the number of mismatches between the motifs present in these sequences and the deduced consensus. Numbers in front of the sequences identify the small-scale DNA preparations originally used for sequencing. Plasmids of preparations 5 and 8 contained two inserts (a, b). Sequences that could not be determined due to loss of sample or ambiguous sequencing results account for omitted numbers in A, B and C, and in Figure 4. Asterisks identify sequences listed in both A and B.

and decamer (see Figure 2). When the corresponding experiment was performed with ASF/SF2, strong shifted bands appeared with A7, A14 and an SC35ΔRS-selected sequence, S16 (Figure 5B). Complex formation with other sequences not selected by ASFΔRS was not observed (C2) or required at least five times higher concentrations of ASF/SF2 (R5 and S33) (Figure 5B). Despite the observation that ASF/SF2 bound well to S16, the results strongly support the conclusion that ASF/SF2 and SC35 have distinct RNA binding specificities.

Three copies of the ASF/SF2 octamer AGAAGAAC (A3) can function as a splicing enhancer

SR proteins, and in particular ASF/SF2, have recently been suggested to stimulate splicing by binding to splicing enhancer elements in 3′ exons (Lavigueur *et al.*, 1993; Sun *et al.*, 1993). In order to assess the functional significance of the ASF/SF2 and SC35 binding sites identified by our data, we asked whether these sites could function as splicing enhancers, using a pre-mRNA with inherently weak splicing signals. We chose as splicing substrate (GN in Figure 6A)

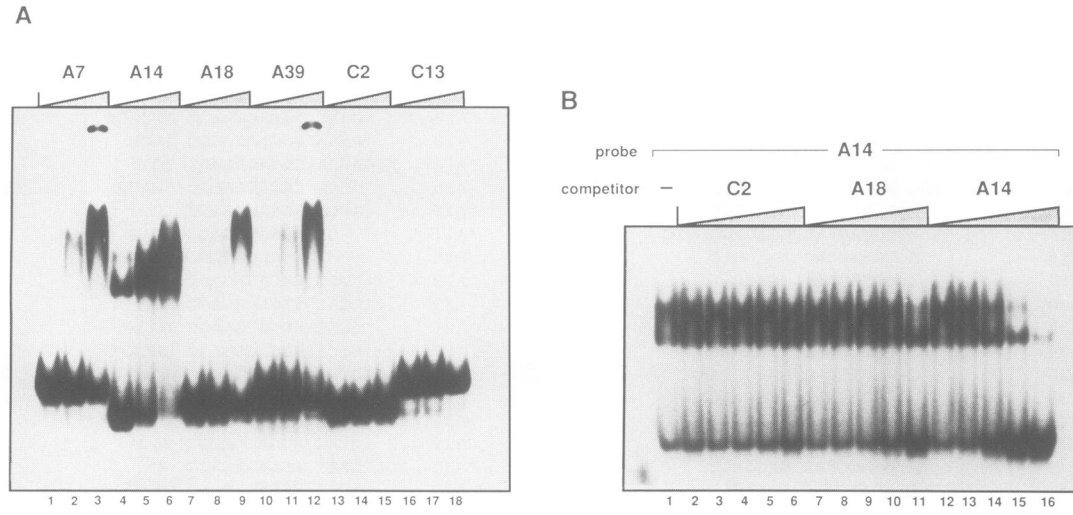

Fig. 3. Mobility shift assay of ASFΔRS. (**A**) Constant amounts of various labeled RNA probes (A7, A14, A18, A39, C2, C13, for sequences see Figure 2) were incubated with increasing amounts (0.3, 1.5 or 7.5 pmol) of ASFΔRS. (**B**) Competition assay with A14 as radiolabeled probe and increasing amounts of three different unlabeled RNA sequences (A14, A18, C2) as competitors. Each competitor was used at 0.5-, 2.5-, 12.5-, 60- or 300-fold molar excess over the probe.

A

S 1	GCAGAAGTA	**AGCGGAGAA** A	2
S 3	GT	**AGGAGTGAC** GCTGACGAG	3
S 7	GA	**CGGAGAGAA** AAGGTGTAA	2
S10	GGAA	**AGCAGAGAT** AGCAG	2
S11		**AGCAGTGTT** CCGGATAGGA	2
S12	G	**AGCAGAGGA** GTAGAAGGC	1
S14	AGTGT	**AGCAGAGGA** AGCA	1
S16	CCCGGAA	**AGCAGAGTA** GGA	0
S17	GCGGAGTGTA	**AGGAGAATA** A	1
S18	TTGCAGAG	**AGGAGAGGT** A	2
S19	AGTTCCAGATGT	**AGCAGAgtt**	1
S20	GT	**AGGAGATAA** AGTACGAG	2
S21	TC	**AGGCGAGTA** GAAT **AGCAGAgtt**	1,1
S24		**tGCAGAGTG** ATAG AGAAGG	2
S25	TG	**AGAAGCGGA** AGAACGA	3
S30	GTA	**AGCCGAGAA** GT **AGCAGAgtt**	2,1
S33	AAGAG	**AGGAGAGGT** GGAG	2
S34		**AGGAGCGTG** TTGCGGAAGGA	2
S35	CGAGGTAGTAG	**ACGAGAgtt**	2
S37	TAGGTT	**AGGCGTGTT** CGA	3
S41	TGTA	**AGGCGAGCA** TGAAG	2
S44	TG	**AGGAGTATC** GTGTGTTA	3
S47	AC	**AGGAGAATG** AGTAGCGGAA	2
S48	GTGCA	**GACGGAGTA** AGGA	3
	consensus:	**AGSAGAGTA**	

B

S 2		**GTTCCAGTT** AGAAGCTGATA	2
S 4	TAGCATGCAGATT	**GTTCCAGgt**	3
S19	A	**GTTCCAGAT** GTAGCAGA	3
S22	CCAGT	**GTTCGAGTA** AAGCGA	0
S27	AA	**GCTCGAGTA** AGTACGCGG	1
S28	GGAGATA	**GTTCGAGTA** AGCA	0
S31	ACA	**GTACGAGGA** AGATGCAG	2
S37	TAGGTTAGGCGT	**GTTCGAgtt**	1
S38	GA	**GATGGAGGA** AAACGCGAA	3
S45	AGAAGGCGTATT	**GTTCGAgtg**	1
S46	CCAGATA	**GTTCGAGTA** AGCA	0
	consensus:	**GTTCGAGTA**	

Fig. 4. Selection/amplification with SC35ΔRS. Sequences of individual clones selected by SC35ΔRS are shown. Motifs resembling either the purine-rich nonamer AGSAGAGTA (S = C or G) (**A**) or the sequence GTTCGAGTA (**B**) are represented in bold type characters. Numbers behind the sequences indicate the number of mismatches between the motifs present in these sequences and the deduced consensus.

an RNA containing the second intron of the α-globin pre-mRNA with a heterologous 5′ splice site derived from an alternative exon of the pre-mRNA coding for the neural cell adhesion molecule (NCAM). This intron is spliced very poorly both *in vivo* (Tacke and Goridis, 1991) and *in vitro* (Figure 6C, lanes 1–4). However, insertion of three copies of the consensus ASF/SF2 octamer (A3 in Figure 6B) 80 nucleotides downstream from the 3′ splice site enhanced *in vitro* splicing dramatically (Figure 6C, lanes 5–8). In contrast, three copies of the purine-rich consensus SC35 nonamer (S3 in Figure 6B) had no effect on splicing (Figure 6C, lanes 9–12), even when nuclear extracts were supplemented with high concentrations (7–35 pmol/reaction) of recombinant SC35 (data not shown). To provide evidence that A3-dependent splicing was due to binding of a *trans*-acting factor to the A3 repeat, we performed competition with isolated copies of A3 and S3. Splicing of the pre-mRNA that contained the A3 repeat (GN-A3) was completely inhibited by a 27-fold molar excess of A3 itself (Figure 7, lane 4). In contrast, the same amount of S3 showed at most only a slight inhibition of splicing (Figure 7, lane 7).

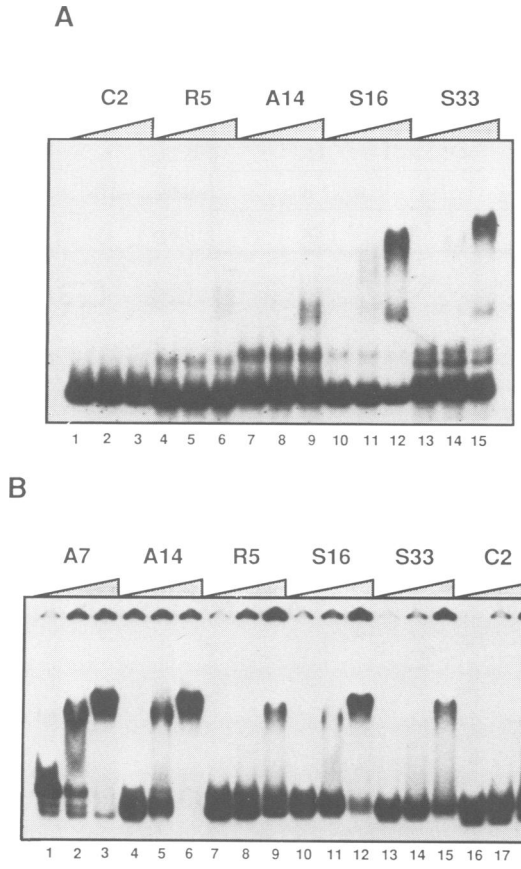

Fig. 5. Mobility shift assays of SC35 (**A**) and ASF/SF2 (**B**). Radiolabeled RNA probes (A7, A14, R5, S16, S33, C2, see Figures 2 and 3 for identification of the sequences) were incubated with increasing amounts of the indicated purified protein. 0.1, 0.5 or 2.5 pmol SC35 and 0.4, 2 or 10 pmol ASF/SF2 were used per assay.

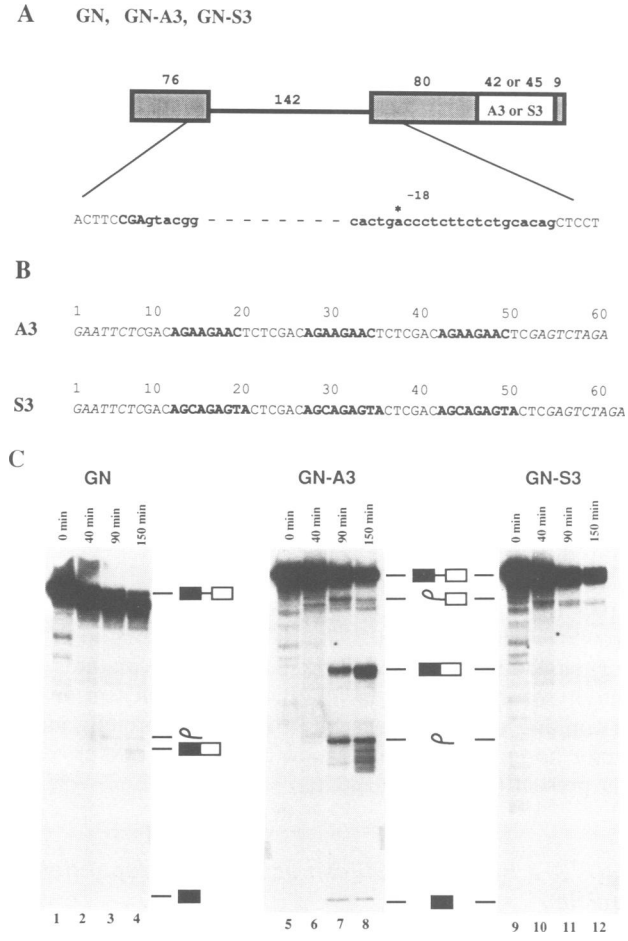

Fig. 6. Three copies of the ASF octamer stimulate splicing of a chimeric α-globin/NCAM intron *in vitro*. (**A**) Schematic representation of the chimeric α-globin/NCAM splicing substrate (GN) and its derivatives (GN-A3 and GN-S3) containing three copies of either the ASF/SF2 consensus octamer (A3) or the purine-rich SC35 consensus nonamer (S3) in the 3′ exon. Numbers on top indicate the length of the corresponding exonic and intronic regions. The sequences around the 5′ and 3′ splice sites are shown below. Exonic sequences are represented in upper case, intronic sequences in lower case letters. The presumed branch point is indicated at position –18 with respect to the 3′ exon, as previously determined in the wild-type intron (Reed and Maniatis, 1985). (**B**) Sequences of A3 and S3 with the ASF/SF2 consensus octamer and the purine-rich SC35 consensus nonamer represented in bold type characters. Sequences also present in GN are shown in italics. (**C**) Time course of *in vitro* splicing reactions using nuclear extracts from HeLa cells and GN, GN-A3 and GN-S3 as splicing substrates. The positions of the autoradiographic bands corresponding to the splicing substrates, intermediates and spliced products are indicated.

SR proteins are required but not sufficient for activation of A3-dependent splicing

The above data indicate that high-affinity binding sites for ASF/SF2, but not for SC35, are sufficient to stimulate *in vitro* splicing of a pre-mRNA with otherwise weak splicing signals, thereby supporting the idea that ASF/SF2 can play a critical role in splicing enhancement. However, although we were able to accomplish splicing of a variety of pre-mRNAs by complementation of HeLa cell S100 extracts with recombinant ASF/SF2, we were unable to do so with GN-A3. Furthermore, recombinant ASF/SF2 could not restore splicing of GN-A3 in nuclear extracts after inhibition with A3 (data not shown). A possible explanation for these results is that the activity of at least one factor in addition to ASF/SF2 is required for activation of GN-A3 splicing and that this factor, like ASF/SF2, is both absent from S100 extracts and also titrated from nuclear extracts by A3. We considered the possibility that this activity might be provided by one or several SR proteins other than ASF/SF2. However, a preparation of SR proteins purified from HeLa cell nuclear extracts that contained all six classical SR proteins (Zahler *et al.*, 1992) was unable to complement S100 for splicing of GN-A3 (Figure 8A, lane 2), although it stimulated efficient splicing of a pre-mRNA with consensus sequence splice sites (data not shown). We proceeded by testing fractions of HeLa cell nuclear extracts for their ability to activate *in vitro* splicing of GN-A3 in S100 supplemented with SR proteins. Such an activity was detected in a 20–40% ammonium sulfate fraction of nuclear extract (NF20–40, see Materials and methods). NF20–40 was unable to activate GN-A3 splicing in S100 by itself (Figure 8A, lane 3), but did so in combination with SR proteins (Figure 8A, lane 4). Western blot analyses using a monoclonal antibody that specifically recognizes SR proteins (mAB 104, Zahler *et al.*, 1992) and anti-ASF/SF2 polyclonal antibodies revealed that NF20–40 contained only very low amounts of SR proteins and barely detectable levels of ASF/SF2 (data not shown). Because similar or higher amounts of SR proteins were also detected in the inactive

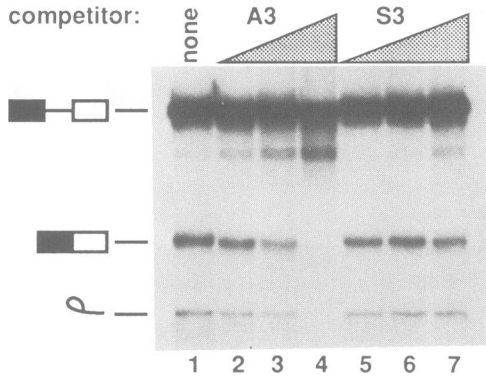

Fig. 7. Splicing of GN-A3 *in vitro* is inhibited by A3 as competitor. The sequences of the competitors A3 and S3 were exactly as shown in Figure 6B. GN-A3 was spliced for 90 min with no competitor present (lane 1), or in the presence of 3-, 9- or 27-fold molar excess of A3 (lanes 2–4) or S3 (lanes 5–7). Only the splicing substrate and the spliced products are shown.

S100 (data not shown), and because addition of SR proteins alone could not activate GN-A3 splicing in S100, our data suggest that A3-dependent splicing requires the activity of at least two factors that are not present or are of limiting concentrations in S100: an SR protein and a factor absent in the preparation of classical SR proteins but present in NF20–40.

ASF/SF2 but not SC35 can activate A3-dependent splicing

Our results indicate that SR proteins are required, though not sufficient, for activation of A3-dependent splicing in S100. We next wished to determine whether ASF/SF2 itself could activate *in vitro* splicing of GN-A3 specifically in S100 supplemented with NF20–40. As shown in Figure 8B, recombinant ASF/SF2 was as efficient in activation of splicing as HeLa cell SR proteins, whereas no splicing was observed with similar concentrations of SC35. SC35 showed no activity in this assay over a wide range of concentrations, all of which promoted efficient splicing of a substrate with consensus sequence splice sites in a standard S100 complementation assay (results not shown). We conclude that ASF/SF2, but not SC35, can activate A3-dependent splicing in conjunction with one or more additional factors.

We next carried out UV-crosslinking experiments with A3 RNA as a radiolabeled probe, and exogenously added recombinant ASF/SF2 and SC35 in S100, to determine whether the differences in activity in the GN-A3 splicing assay could be correlated to differences in binding to A3. For comparison, we also performed crosslinking with S3 RNA. As shown in Figure 9A, ASF/SF2 crosslinked much better to A3 than to S3, while the opposite was true for SC35 (Figure 9A, compare lanes 5 and 3, and lanes 2 and 6). In nuclear extracts, UV-crosslinking yielded a strong doublet that comigrated with ASF/SF2 and SC35 and, in addition, two bands corresponding to proteins with apparent molecular weights of ~40 and 25 kDa (Figure 9A). Because the same bands also appeared when crosslinking was performed with A3 in S100 supplemented with SR proteins (results not shown), the corresponding proteins in nuclear extracts are probably SRp30, SRp40

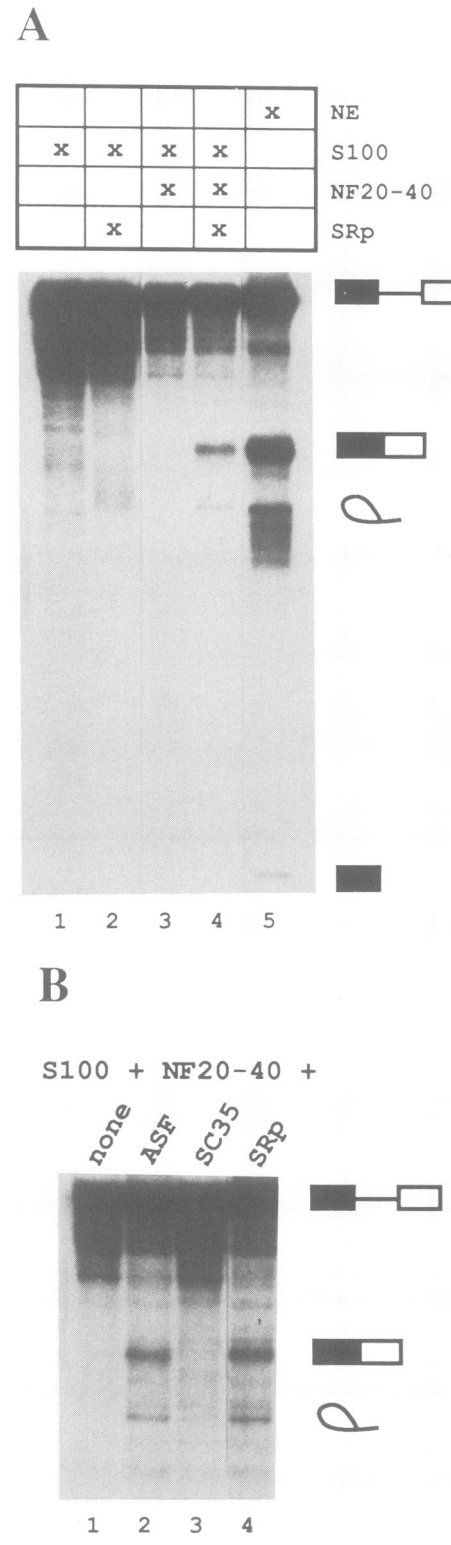

Fig. 8. Activation of splicing of GN-A3 in S100. GN-A3 was incubated for 150 min under splicing conditions in nuclear extract (NE) or various combinations of S100, NF20–40 and SR proteins, as indicated. The amounts of extracts and nuclear fractions used were as indicated in Materials and methods. Recombinant ASF/SF2 and SC35 were used at concentrations of 7 pmol/assay.

and SRp20. In contrast to A3, S3 did not crosslink efficiently to SRp30 in nuclear extracts. Instead, a strong signal was produced by an 80 kDa protein (Figure 9A,

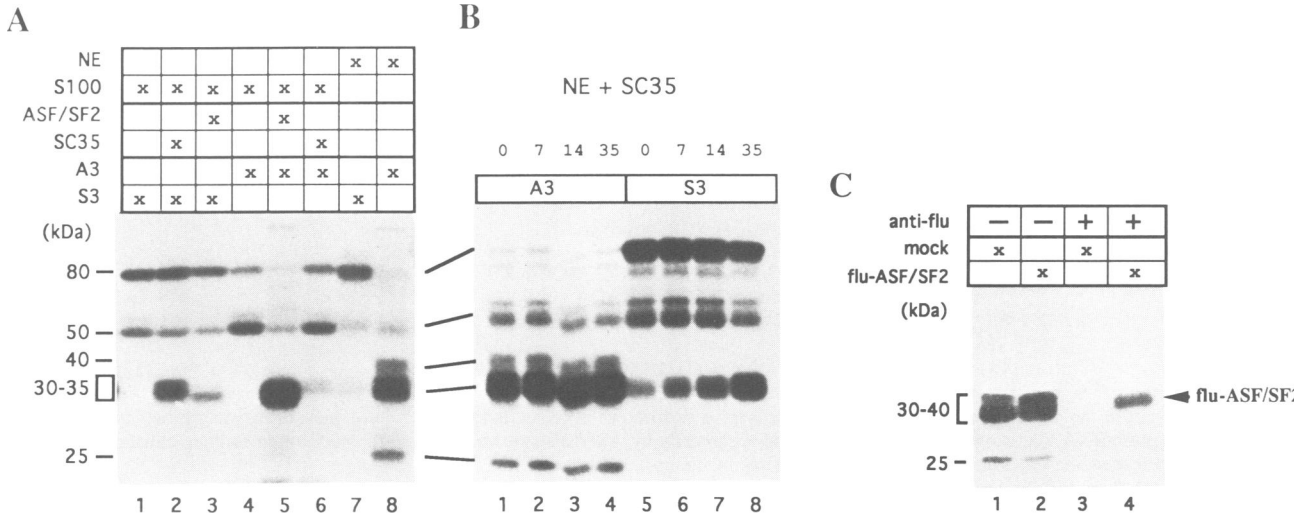

Fig. 9. UV-crosslinking of SR proteins to A3 and S3. Prior to irradiation with UV, radiolabeled probes were incubated for 30 min under splicing conditions, except that polyvinyl alcohol was omitted. (**A**) The combinations of nuclear extract (NE), S100, SR proteins, ASF/SF2, SC35 and radiolabeled probe (A3 or S3) used for crosslinking are indicated at the top. Recombinant ASF/SF2 and SC35 were used at concentrations of 25 pmol/assay. (**B**) The concentrations of recombinant SC35 (in pmol/reaction) added to NE are indicated at the top. (**C**) Crosslinking was performed with A3 as radiolabeled probe and nuclear extracts from HeLa cells transfected with an expression plasmid that contained either no insert (mock) or a cDNA for flu–ASF/SF2 (flu-ASF/SF2). Shown are crosslinked proteins before (–) and after (+) immunoprecipitation with anti-flu antibodies.

lane 7), which was also detected in S100 (Figure 9A, lanes 1–3). To determine whether this protein might interfere with binding of SC35 to S3, we tested the effect of supplementing nuclear extracts with recombinant SC35 (7–35 pmol/reaction). This resulted in a strong increase in crosslinking of SRp30 to S3 and a corresponding decrease in the 80 kDa crosslinking (Figure 9B, lanes 5–8). Crosslinking of SRp30 to A3 was not affected by SC35 (Figure 9B, lanes 1–4) and, at the highest level of SC35, SRp30 crosslinking to A3 and S3 was in fact equivalent (Figure 9B, compare lanes 4 and 8). Thus the inability of S3 to function as a splicing enhancer (see above) was not due to an inability to bind SC35.

The finding that, in S100, ASF/SF2 crosslinked much better to A3 than did SC35 (Figure 9A, compare lanes 5 and 6) strongly supports the idea that crosslinking of SRp30 to A3 in nuclear extracts (Figure 9A, lane 8) was due to ASF/SF2. To establish this directly, we performed crosslinking experiments with nuclear extracts prepared from HeLa cells expressing recombinant ASF/SF2 fused to an N-terminal influenza virus hemagglutinin epitope (flu–ASF/SF2; Kohtz *et al.*, 1994). Flu–ASF/SF2 was present in these extracts at approximately the same concentration as endogenous ASF/SF2, from which it could be distinguished by its slower migration in SDS–polyacrylamide gels (results not shown). Nuclear extracts from HeLa cells transfected with the expression vector lacking a cDNA insert yielded the same crosslinked products with A3 RNA as extracts from untransfected cells (Figure 9C, lane 1) and none of these products was immunoprecipitated with anti-flu epitope antibodies (Figure 9C, lane 3). In contrast, extracts from cells expressing flu–ASF/SF2 revealed an additional strong band following crosslinking to A3 (Figure 9C lane 2), which was also detected after immunoprecipitation (Figure 9C, lane 4). Together, our results provide strong evidence that endogenous ASF/SF2 participates directly in activation of GN-A3 splicing by binding to its experimentally determined consensus recognition motif.

Discussion

In this paper we have presented several lines of evidence suggesting that the two essential splicing factors, ASF/SF2 and SC35, are sequence-specific RNA binding proteins with different RNA binding specificities. First, selection/amplification from random sequence pools identified different consensus motifs as high-affinity binding sites for ASFΔRS and SC35ΔRS. Second, mobility shift data obtained with the full-length proteins were consistent with the results of the selection/amplification experiments. The fact that one SC35ΔRS-selected sequence (S16) was recognized efficiently by ASF/SF2 can probably be attributed to the fortuitous presence of a degenerate ASF/SF2 binding site. On the other hand, it is remarkable that the S33 sequence, despite its extremely high purine content—it contains only a single pyrimidine—bound ASF/SF2 much less efficiently than did the ASFΔRS-selected sequences. This observation underscores the fact that ASF/SF2 is not merely a polypurine binding protein, but in fact recognizes short, distinct sequences. Third, three copies of ASF/SF2 or SC35 binding sites display specific crosslinking activities with respect to these two proteins.

Implications of the RNA binding specificities of ASF/SF2 and SC35 for their function as essential splicing factors

Our results offer, in part, an explanation for previously published data indicating that ASF/SF2 and SC35 may commit different pre-mRNAs to the splicing pathway (Fu, 1993). It was shown that SC35, but not ASF/SF2, was sufficient to allow *in vitro* splicing of a human β-globin pre-mRNA when nuclear extract was subsequently added

Table I. Sequences similar to the ASF/SF2 consensus octamer RGAAGAAC are present in naturally occurring splicing enhancers and in an HIV tat pre-mRNA

Gene (exon)	Sequence	Mismatches with **RGAAGAAC**
Human FN (exon ED1)[a]	...TGG**TGAAGAAG**ACACT...	2
Cardiac TNT (exon 5)[b]	AAGA**GGAAGAAT**GGCTTGA**GGAAGACG**ACG	1, 2
Human hprt (exon 3)[c]	...GAGATGTGATGA**TGAAGGAG**ATGGGAGG	3
Mouse IgM (exon M2)[d]	...**GGAAGGAC**AGCAGACCAAGAG...	1
ASLV (env 3′ exon)[e]	...CGAGCA**AGAAGGAC**TCCA**AGAAGAAG**CCGCCAGCAACAAGCA**AGAAGGAC**CCC...	1, 1, 1
Bovine GH (exon 5)[f]	...TCTCCTGCTTCC**GGAAGGAC**CTGCATAAGA...	1
HIV tat (exon 1)[g]	...TTACTCGACA**GGAGAGC**AAGAAATGGAGCCATGGCA**GGAAGAAG**CGGAGAC	
	AGCGA**CGAAGACC**...	2, 1, 2
HIV tat (exon 2)[g]	...**CGAAGGAATAGAAGAAG**AAGGTGGAG**AGAGAGAC**AGAT...	3, 1, 3

Shown are all or part of the purine-rich splicing enhancer elements of various pre-mRNAs (a–f), and purine-rich sequences in the two exons of a modified HIV tat pre-mRNA that can be committed to splicing specifically by ASF/SF2 (g). Mutations in the purine-rich splicing enhancers caused skipping of the corresponding exon (a, b, c) or decreased splicing of the upstream intron (d, e, f).
References: [a]Mardon *et al.* (1987), Lavigueur *et al.* (1993); [b]Cooper and Ordahl (1989), Cooper (1992); [c]Steingrimsdottir *et al.* (1992); [d]Watakabe *et al.* (1993); [e]Katz and Skalka (1990), Fu *et al.* (1991); [f]Hampson *et al.* (1989); [g]Fu (1993).
Abbreviations used: (FN) fibronectin; (TNT) troponin T; (hprt) hypoxanthine–guanine phosphoribosyltransferase; (IgM) immunoglobulin μ; (ASLV) avian sarcoma-leukosis virus; (GH) growth hormone; (HIV) human immunodeficiency virus.

together with excess of competitor RNA. With a human immunodeficiency virus (HIV) tat RNA splicing substrate, the situation was the reverse, with ASF/SF2 but not SC35 functioning to commit the RNA to splicing. Interestingly, the HIV tat pre-mRNA, but not the β-globin pre-mRNA, contains several sequence motifs in its 5′ and 3′ exons that show strong resemblance to the ASF/SF2 consensus octamer (see Table I). In line with a previously proposed model, ASF/SF2 may first bind stably to high-affinity binding sites within this pre-mRNA and then recruit U1 snRNP to the 5′ splice site through interaction with the U1 snRNP 70 kDa protein (Kohtz *et al.*, 1994). We cannot speculate on the basis for commitment of the β-globin pre-mRNA by SC35, since no potential high-affinity binding sites for SC35 could be detected in this splicing substrate.

Several studies that have addressed the RNA binding properties of multiple RBD-containing proteins, such as the U1 snRNP A protein and the poly(A) binding protein (PABP), indicated that RBDs can interact with RNA independently (Lutz-Freyermuth *et al.*, 1990; Nietfeld *et al.*, 1990; Burd *et al.*, 1991; Lutz and Alwine, 1994). This may not be the case for the two RBDs of ASF/SF2, as sequences selected by ASF RBD1 were not selected by ASFΔRS and did not bind ASF/SF2 efficiently. Although we have not performed SELEX with the second, non-canonical ASF/SF2 RBD, it seems unlikely that this RBD alone would determine the specificity of the intact molecule. Therefore, the sequence specificity of ASF/SF2 is likely to be determined by cooperation between the two RBDs, each of which by itself recognizes a distinct sequence. It is noteworthy that a similar conclusion was drawn from a recent analysis of the RNA binding specificities of hnRNP A1 and its two individual RBDs (Burd and Dreyfuss, 1994b). Conceivably, the presence of two RBDs rather than one may ensure a wider range of RNA binding affinities and specificities. The fact that ASF RBD1 selected different sequences than ASFΔRS may be functionally relevant in the light of previous findings suggesting the existence of ASF/SF2 isoforms that do not contain a complete RBD2, i.e. ASF2 and ASF3 (Ge *et al.*, 1991). ASF3 has been shown to act as a dominant inhibitor of ASF/SF2 in *in vitro* splicing reactions (Zuo and Manley, 1993) and the results presented here raise the possibility that this may reflect at least in part an altered RNA binding specificity.

ASF/SF2 has previously been shown to recognize 5′ splice sites in two different pre-mRNAs (Zuo and Manley, 1994). Although our present work argues against 5′ splice sites being the highest affinity targets of ASF/SF2, it is consistent with the observation that binding of ASF/SF2 to the SV40 small t intron was affected most severely by mutations decreasing the purine content of the 5′ splice site (Zuo and Manley, 1994). Because most introns are unlikely to contain or have in their vicinity high-affinity ASF/SF2 binding sites, recognition of 5′ splice sites by ASF/SF2 may play an important role in spliceosome assembly and may account for the protein's effect on 5′ splice site selection. Interestingly, the SC35 consensus nonamer AGSAGAGTA also bears some resemblance to the consensus 5′ splice site. It is tempting to speculate that SC35, like ASF/SF2, may have limited affinities for 5′ splice sites, which could explain at least in part the functional similarities between these two splicing factors. Recognition of 5′ splice sites may possibly also underly the apparent ability of SR proteins to substitute for U1 snRNP function (Crispino *et al.*, 1994; Tarn and Steitz, 1994).

The role of ASF/SF2 in the activation of purine-rich splicing enhancers

Several studies have recently shown that purine-rich sequences within 3′ exons can activate or enhance splicing of upstream introns (Lavigueur *et al.*, 1993; Watakabe *et al.*, 1993, and references therein; Xu *et al.*, 1993; Tanaka *et al.*, 1994). Strikingly, in all of these cases, the splicing enhancer elements contain motifs closely matching the consensus octamer RGAAGAAC suggested by our SELEX data to be a high-affinity binding site for ASF/SF2 (see Table I). Indeed, we found that three copies of this sequence constitute a powerful splicing enhancer *in vitro* (ASF/SF2 splicing enhancer, ASE) and that ASF/SF2 can specifically activate ASE-dependent splicing. A previous study with bovine growth hormone pre-mRNA

provides additional support for the physiological significance of our results. ASF/SF2 was shown to stimulate splicing of the last intron of this pre-mRNA through interaction with a 115 nucleotide enhancer element (Sun *et al.*, 1993). Subsequent mutational analyses indicated that a purine-rich pentamer sequence (GGAAG) was essential for function of the element (Dirksen *et al.*, 1994). This pentamer is in fact part of the sequence GGAAGGAC, which we predict to be a high-affinity ASF/SF2 binding site because of its almost perfect match to the ASF/SF2 consensus octamer. It is also noteworthy that a study examining the capacity of polypurine stretches to function as splicing enhancers *in vitro* found that the exact sequence is important. Notably, poly(A) or poly(G) sequences proved to be completely inactive in this assay, while $(AAG)_4$ and $(AAG)_8$ were most effective (Tanaka *et al.*, 1994).

At least one factor in addition to ASF/SF2 that is not present in S100 is required for ASE function. Because SR proteins failed to activate ASE-dependent splicing in S100, this factor is probably not a classical SR protein. The factor could not be detected by UV-crosslinking with A3 (R.T. and J.L.M., unpublished observation), thus raising the possibility that it does not function as an RNA binding protein, but instead provides essential protein–protein interactions. In this context, we note that, although SR proteins have been shown to be associated with various purine-rich splicing enhancers (Lavigueur *et al.*, 1993; Sun *et al.*, 1993; Staknis and Reed, 1994), activation of enhancer-dependent splicing with SR proteins in S100 has not, to our knowledge, been documented previously.

In contrast to the ASE, three copies of a high-affinity binding site for SC35 (S3) failed to stimulate splicing of the GN pre-mRNA, even when nuclear extracts were supplemented with high concentrations of SC35. Since efficient binding of SC35 to S3 was observed under these conditions, SC35 may be unable to stimulate splicing of introns with weak splicing signals by binding to exon sequences. Consistent with this, all known purine-rich enhancer sequences contain motifs that match the ASF/SF2, but not the SC35, binding site consensus. We cannot, however, exclude the alternative possibility that the 80 kDa protein in nuclear extracts that crosslinked to S3 interfered with SC35 function in our assay. Further experiments are required to assess a potential role of SC35 in splicing enhancement.

In summary, we have shown that ASF/SF2 and SC35 possess different RNA binding specificities and that ASF/SF2 can specifically activate *in vitro* splicing of a pre-mRNA that contains high-affinity binding sites for this splicing factor. However, because ASF/SF2 and SC35 can also activate splicing of pre-mRNAs that do not contain sequences with significant similarity to such sites, it follows that these motifs are not absolutely required for these proteins to function as essential splicing factors *in vitro*. Pre-mRNA splicing that does not rely on such sites may involve lower affinity interactions, such as with 5′ splice sites. If this is generally true for SR proteins, it may account for the redundancy of SR protein function in this respect. In contrast, high-affinity binding sites for ASF/SF2 and possibly other SR proteins may be required for splicing of pre-mRNAs with otherwise weak splicing signals, and may allow for alternative splicing events to be regulated by specific SR proteins or by combinations of these and other proteins.

Materials and methods

Oligonucleotides

Synthetic oligodeoxynucleotides used for SELEX experiments were identical (SELREV) or similar to those used by Tsai *et al.* (1991): SELREV, CCCGACACCCGCGGATCCATGGGCACTATTTATATCAAC; SELT7, CGCGAATTCTAATACGACTCACTATAGGGGCCACCAACGACATT; SELN20, TGGGCACTATTTATATCAACN_{20}GTTGATATAAATAGTGCCC (N_{20} standing for 20 bases of random sequence). Three oligodeoxynucleotides were used for amplification of the SC35 coding region by RT-PCR: SCREV (for reverse transcription), TACACTGCTTGCCGATACATC; SCU (upstream PCR primer), CGCGGATCCATGAGCTACGGCCGCCCCCCT; and SCD (downstream antisense PCR primer), CGCAAGCTTAAGAGGACACCGCTCCTC. Oligodeoxynucleotides used for construction of A3 and S3, respectively were: OCTA1, TCGACAGAAGAACTC; OCTA2, TCGAGAGTTCTTCTG; NONA1, TCGACAGCAGAGTAC; and NONA2, TCGAGTACTCTGCTG.

Constructs

pDS-SC35ΔRS (for expression of SC35ΔRS in Escherichia coli). First, the coding region of human SC35 was amplified by RT-PCR of total RNA from HeLa cells using a primer for reverse transcription and two PCR primers that introduced artificial restriction sites upstream (*Bam*HI) and downstream (*Hin*dIII) of the amplified cDNA. The *Bam*HI–*Hin*dIII fragment was subcloned into the bacterial expression vector pDS56-6H. Subsequently, the 3′ portion coding for the RS domain was removed by digestion with *Avr*II and *Hin*dIII and religation after filling-in.

pGN, pGN-A3, pGN-S3. A PCR fragment containing part of the human α-globin gene with a mutant 5′ splice site derived from exon 18 of the NCAM gene (Tacke and Goridis, 1991) was cloned into the *Sma*I site of pGEM2 (Promega Biotec), such that the *Eco*RI site of the polylinker was placed downstream. Subsequently, the *Hin*dIII–*Eco*RI fragment containing the second intron was transferred into pGEM7Z to produce pGN. To construct pGN-A3 and pGN-S3, the corresponding *Xho*I–*Sal*I fragments constituting, respectively, the A3 or S3 repeat were inserted into the *Xho*I site of pGN. These fragments were obtained by employing a previously described procedure (Watakabe *et al.*, 1993). Oligodeoxynucleotides OCTA1 and OCTA2, or NONA1 and NONA2, were phosphorylated, annealed and ligated. The ligation products were digested with *Sal*I and *Xho*I and the desired fragments purified by electroelution from non-denaturing polyacrylamide gels. For transient expression of flu–ASF/SF2 in HeLa cells, a cDNA for ASF/SF2 with an additional N-terminal influenza hemagglutinin epitope (Kohtz *et al.*, 1994) was cloned into the *Bam*HI site of a modified version of pCMV-Neo-Bam (Baker *et al.*, 1990), in which the transcription unit for neomycin had been removed.

Recombinant proteins

SC35ΔRS was expressed in JM101 from pDS-SC35ΔRS. Likewise ASFΔRS and ASF RBD1 were expressed using pDS56-6H-derived expression plasmids (Zuo and Manley, 1993). Expression and purification by Ni^{2+} agarose chromatography was performed as described previously (Ge *et al.*, 1991), except that SC35ΔRS was finally dialyzed against 10 mM HEPES, pH 7.9, 200 mM KCl, 20% glycerol, and ASF derivatives were dialyzed against the same buffer supplemented with 1 M urea. ASF/SF2 and SC35 were expressed from SF21 cells infected with the corresponding recombinant baculoviruses (Fu and Maniatis, 1992a; Tian and Maniatis, 1993), kindly provided by Xiang-Dong Fu. Monolayer cultures of SF21 cells were harvested 48 h after infection and the recombinant proteins purified as described (Tian and Maniatis, 1993).

In vitro transcription

Depending on the promoter present in the template, *in vitro* transcription was performed with SP6 or T7 polymerases (New England Biolabs). Capped substrates for *in vitro* splicing were synthesized in the presence of 2 mM m7GpppG, 100 μM GTP and 330 nM [α-^{32}P]GTP (3000 Ci/mmol). The corresponding specific activities applied to uncapped RNA probes in gel mobility shift and UV-crosslinking assays were, respectively, two times lower and two times higher. RNA used in SELEX experiments was synthesized with all nucleotides present at 1 mM final concentration,

and 330 nM [α-^{32}P]GTP or [α-^{32}P]UTP added for trace labeling. All *in vitro* transcripts were purified by denaturing PAGE.

SELEX

For selection, ~10 μg each of recombinant ASFΔRS and SC35ΔRS were immobilized on Ni^{2+} agarose and exposed to ~20 μg of random sequence RNA. SC35ΔRS was incubated with 15 μl of packed Ni^{2+} agarose beads in a total volume of 200 μl buffer (10 mM HEPES, pH 7.9, 200 mM KCl, 10% glycerol) for 30 min at 4°C. The beads were subsequently washed three times with 0.5 ml of the same buffer. ASFΔRS and ASF RBD1 were immobilized in a similar fashion, except that 1 M urea was present during binding, and 1 M, 0.5 M and no urea during the subsequent washes. To obtain random sequence RNA required for the first cycle of selection, 10 ng SELN20 was amplified by PCR using 600 ng each of primers SELT7 and SELREV. One μg of the amplification products was digested with *Bam*HI and transcribed with T7 polymerase to produce ~20 μg of trace-labeled RNA. Binding of the RNA to immobilized protein was carried out in 200 μl binding buffer (10 mM HEPES, pH 7.9, 200 mM KCl, 10% glycerol, 0.025% NP-40, 400 μg/ml tRNA). After 30 min incubation at room temperature, the beads were spun down and washed several times with binding buffer containing increasing concentrations of KCl and tRNA. Washing was continued until the amount of bound RNA became less than 10% of the input RNA. In subsequent cycles, increasingly stringent conditions for binding and washes were employed. The most stringent condition consisted of a combination of 400 mM KCl and 1 mg/ml tRNA. Selected RNA was eluted with 200 μl 500 mM NaCl, 0.1% SDS, extracted with phenol/chloroform and precipitated with ethanol. A fifth of it was used for reverse transcription with 5 U AMV reverse transcriptase (Promega Biotec) and 100 ng SELREV as primer. After annealing of RNA and primer for 5 min at 50°C, the reaction was allowed to proceed for 1 h at 42°C. The resulting cDNA was recovered by phenol/chloroform extraction and ethanol precipitation, and a fraction was amplified by PCR to enter the next cycle of selection. PCR reactions consisted of 25 cycles: 5× (30 s at 95, 40, 72°C) + 20× (30 s at 95, 60, 72°C). PCR products were cloned as *Eco*RI–*Bam*HI fragments into pGEM2 (ASFΔRS-selected sequences) or pSP64polyA (SC35ΔRS-selected sequences).

Transfections

Monolayer cultures of HeLa cells were transfected using the calcium phosphate precipitation method. Briefly, 10 μg of expression vector or flu–ASF/SF2 expression plasmid were added to each of six 150 mm plates, and a glycerol shock was performed 16 h later. At 48 h post-transfection, cells were harvested and nuclear extracts were prepared by a small-scale procedure (Lee *et al.*, 1988).

Gel mobility shift assays

Binding of recombinant proteins to radiolabeled RNA (~100 fmol/20 μl assay) was carried out as described in Zuo and Manley (1993), except that the concentration of KCl was raised to 400 mM. RNA–protein complexes were separated from free RNA by electrophoresis on 5% native polyacrylamide gels using either TBE (Figure 3) or Tris-glycine (Figure 5) buffer systems.

Fractionation of HeLa cell extracts

Nuclear extract and cytoplasmic S100 was prepared from HeLa cells by the method of Dignam *et al.* (1983). To obtain NF20–40, ammonium sulfate (AS) was added to nuclear extracts to 20% saturation and precipitated material removed by centrifugation at 30 000 *g* for 20 min. After addition of AS to 40% saturation and stirring of the extract for 30 min, centrifugation was repeated. The pellet was resuspended in a minimal volume of buffer D (Dignam *et al.*, 1983) and residual AS removed by gel filtration. SR proteins were prepared from the supernatant following essentially the procedure of Zahler *et al.* (1992).

In vitro splicing and UV-crosslinking

In vitro splicing was performed as described previously (Harper and Manley, 1991). In general, 10 μl of nuclear extracts or S100 were used per 25 μl reaction. In cases where S100 was supplemented with NF20–40, 5 μl of NF20–40 were added to 7 μl of S100. SR proteins were used at a concentration of ~0.4 μg per reaction. For competition experiments, cold competitor RNA and labeled splicing substrate were added simultaneously to the extract. The sequences of the radiolabeled RNA probes A3 and S3 used in UV-crosslinking studies were exactly as shown in Figure 6B. The probes were first incubated with nuclear extracts or S100 plus recombinant proteins for 30 min under splicing conditions, and then irradiated for 20 min with UV light (254 nm) as described previously (Zuo and Manley, 1994). After incubation with 1 μg ribonuclease A for another 30 min, an equal amount of 2× SDS gel loading buffer was added to each sample. Crosslinked proteins were separated on SDS/11% polyacrylamide gels. Immunoprecipitation of flu–ASF/SF2 with anti-flu antibodies was performed as described previously (Kohtz *et al.*, 1994).

Acknowledgements

We would like to thank X.-D.Fu for supplying ASF/SF2- and SC35-expressing baculovirus stocks and J.D.Kohtz, J.Prasad and Y.Takagaki for providing nuclear extracts. We also thank K.G.K.Murthy and J.D.Kohtz for helpful discussions during the course of this work and Y.Chen for excellent technical assistance. This work was supported by a grant from the National Institutes of Health (GM 48259).

References

Amrein,H., Gorman,M. and Nothiger,R. (1988) The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell*, **55**, 1025–1035.

Baker,S.J., Markowitz,S., Fearon,E.R., Willson,J.K.V. and Vogelstein,B. (1990) Suppression of human colorectal carcinoma cell growth by wild-type p53. *Science*, **249**, 912–915.

Burd,C.G. and Dreyfuss,G. (1994a) Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.

Burd,C.G. and Dreyfuss,G. (1994b) RNA binding specifity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.*, **13**, 1197–1204.

Burd,C.G., Matunis,E.L. and Dreyfuss,G. (1991) The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. *Mol. Cell. Biol.*, **11**, 3419–3424.

Caceres,J.F. and Krainer,A.R. (1993) Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J.*, **12**, 4715–4726.

Cooper,T.A.(1992) *In vitro* splicing of cardiac troponin T precursors. *J. Biol. Chem.*, **267**, 5330–5338.

Cooper,T.A. and Ordahl,C.P. (1989) Nucleotide substitutions within the cardiac troponin T alternative exon disrupt pre-mRNA alternative splicing. *Nucleic Acids Res.*, **17**, 7905–7921.

Crispino,J.D., Blencowe,B.J. and Sharp,P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*, **265**, 1866–1869.

Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Accurate transcription by RNA polymerase II in a soluble extract from isolated nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.

Dirksen,W.P., Hampson,R.K., Sun,Q. and Rottman,F.M. (1994) A purine-rich exon sequence enhances alternative splicing of bovine growth hormone pre-mRNA. *J. Biol. Chem.*, **269**, 6431–6436.

Fu,X.-D. (1993) Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature*, **365**, 82–85.

Fu,X.-D. and Maniatis,T. (1992a) Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science*, **256**, 535–538.

Fu,X.-D. and Maniatis,T. (1992b) The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3′ splice site. *Proc. Natl Acad. Sci. USA*, **89**, 1725–1729.

Fu,X.-D., Katz,R.A., Skalka,A.M. and Maniatis,T. (1991) The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. *Genes Dev.*, **5**, 211–220.

Fu,X.-D., Mayeda,A., Maniatis,T. and Krainer, A.R. (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. and Manley,J.L. (1990) A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA *in vitro*. *Cell*, **62**, 25–34.

Ge,H., Zuo,P. and Manley,J.L. (1991) Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. *Cell*, **66**, 373–382.

Goralski,T.J., Edstrom,J.E. and Baker,B.S. (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.

Hampson,R.K., La Follette,L. and Rottman,F.M. (1989) Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. *Mol. Cell. Biol.*, **9**, 1604–1610.

Harper,J.E. and Manley,J.L. (1991) A novel protein factor is required for use of distal alternative 5′ splice sites *in vitro*. *Mol. Cell. Biol.*, **11**, 5945–5953.

Hoshijima,K., Inoue,K., Higuchi,I., Sakamoto,H. and Shimura,Y. (1991) Control of doublesex alternative splicing by transformer and transformer-2 in *Drosophila*. *Science*, **252**, 833–836.

Huh,G.S. and Hynes,R.O. (1994) Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. *Genes Dev.*, **8**, 1561–1574.

Jamison,S.F., Crow,A. and Garcia-Blanco,M.A. (1992) The spliceosome assembly pathway in mammalian extracts. *Mol. Cell. Biol.*, **12**, 4279–4287.

Katz,R.A. and Skalka,A.M. (1990) Control of retroviral RNA splicing through maintenance of suboptimal processing signals. *Mol. Cell. Biol.*, **10**, 696–704.

Kenan,D.J., Query,C.C. and Keene,J.D. (1991) RNA recognition: towards identifying determinants of specificity. *Trends Biochem. Sci.*, **16**, 214–220.

Kim,Y.-J., Zuo,P., Manley,J.L. and Baker,B.S. (1992) A *Drosophila* RNA binding protein RBP1 is localized to transcriptionally active sites of chromosomes and shows a functional similarity to human splicing factor ASF/SF2. *Genes Dev.*, **6**, 2569–2579.

Kohtz,J.D., Jamison,S., Will, C.L., Zuo,P., Luhrmann,R., Garcia-Blanco,M.A. and Manley, J.L. (1994) Protein–protein interactions and 5′-splice site recognition in mammalian mRNA precursors. *Nature*, **368**, 119–124

Krainer,A.R. and Maniatis,T. (1985) Multiple factors including the small ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing *in vitro*. *Cell*, **42**, 725–736.

Krainer,A.R., Conway,G.C. and Kozak,D. (1990a) Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.*, **4**, 1158–1171.

Krainer,A.R., Conway,G.C. and Kozak,D. (1990b) The essential pre-mRNA splicing factor SF2 influences 5′ splice site selection by activating proximal sites. *Cell*, **62**, 35–42.

Krainer,A.R., Mayeda,A., Kozak,D. and Binns,G. (1991) Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, *Cell*, **66**, 383–394.

Lavigueur,A., La Branche,H., Kornblihtt,A.R. and 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.

Lee,K.A.W., Bindereif,A. and Green,M. (1988) A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. *Gene Anal. Techn.*, **5**, 22–31.

Lutz,C. and Alwine,J. (1994) Direct interaction of the U7 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev.*, **8**, 576–586.

Lutz-Freyermuth,C., Query,C.C. and Keene,J. (1990) Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small nuclear ribonucleoprotein complex binds with high affinity to stem–loop II of U1 RNA. *Proc. Natl Acad. Sci. USA*, **87**, 6393–6397.

Mardon,H.J., Sebastio,G. and Baralle,F.F. (1987) A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic Acids Res.*, **15**, 7725–7733.

Michaud,S. and Reed,R. (1991) An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev.*, **5**, 2534–2546.

Michaud,S. and Reed,R. (1993) A functional association between the 5′ and 3′ splice sites is established in the earliest prespliceosome complex (E) in mammals. *Genes Dev.*, **7**, 1008–1020.

Moore,J.M., Query,C.C. and Sharp,P.A. (1993) Splicing of precursors to messenger RNAs by the spliceosome. In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–358.

Nietfeld,W., Mentzel,H. and Pieler,T. (1990) The *Xenopus laevis* poly(A) binding protein is composed of multiple functionally independent RNA binding domains. *EMBO J.*, **9**, 3699–3705.

Reed,R. and Maniatis,T. (1985) Intron sequences involved in lariat formation during pre-mRNA splicing. *Cell*, **41**, 95–105.

Sosnowski,B.A., Belote,J.M. and McKeown,M. (1989) Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. *Cell*, **58**, 449–459.

Staknis,D. and Reed,R. (1994) SR proteins promote the first specific recognition of pre-mRNA and are present together with U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol. Cell. Biol.*, **14**, 7670–7682.

Steingrimsdottir,H., Rowley,G., Dorado,G., Cole,J. and Lehmann,A.R. (1992) Mutations which alter splicing in the human hypoxanthine–guanine phosphoribosyltransferase gene. *Nucleic Acids Res.*, **20**, 1201–1208.

Sun,Q., Mayeda,A., Hampson,R.K., Krainer,A.R. and Rottman,F.M. (1993) General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes Dev.*, **7**, 2598–2608.

Tacke,R. and Goridis,C. (1991) Alternative splicing in the neural cell adhesion molecule pre-mRNA: regulation of exon 18 skipping depends on the 5′-splice site. *Genes Dev.*, **5**, 1416–1429.

Tanaka,K., Watakabe,A. and Shimura,Y. (1994) Polypurine sequences within a downstream exon function as a splicing enhancer. *Mol. Cell. Biol.*, **14**, 1347–1354.

Tarn,W.-Y. and Steitz,J. (1994) SR proteins can compensate for functions of the U1 snRNP. *Genes Dev.*, **8**, 2704–2717.

Tian,M. and Maniatis,T. (1992) Positive control of pre-mRNA splicing. *Science*, **256**, 237–240.

Tian,M. and Maniatis,T. (1993) A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell*, **74**, 105–114.

Tian,M.and Maniatis,T. (1994) A splicing enhancer exhibits both constitutive and regulated activities. *Genes Dev.*, **8**, 1703–1712.

Tsai,D.E., Harper,D.S. and Keene,J.D. (1991) U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids. Res.*, **19**, 4931–4936.

Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.

Watakabe,A., Tanaka,K. and Shimura,Y. (1993) The role of exon sequences in splice site selection. *Genes Dev.*, **7**, 407–418.

Wu,J.Y. and Maniatis,T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, **75**, 1061–1070.

Wu,J. and Manley,J.L. (1989) Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev.*, **3**, 1553–1561.

Xu,R., Teng,J. and Cooper,T.A. (1993) The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. *Mol. Cell. Biol.*, **13**, 3660–3674.

Zahler,A.M., Lane,W.S., Stolk,J.A. and Roth,M.B. (1992) SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.*, **6**, 937–847.

Zahler,A.M., Neugebauer,K.M. Lane,W.S. and Roth,M.B. (1993) Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science*, **260**, 219–222.

Zamore,P.D. and Green,M.R. (1989) Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc. Natl Acad. Sci. USA*, **86**, 9243–9247.

Zamore,P.D., Patton,J.G. and Green,M.R. (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*, **355**, 609–614.

Zhuang,Y. and Weiner,A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell*, **46**, 827–835.

Zhuang,Y. and Weiner,A.M. (1989) A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.*, **3**, 1545–1552.

Zuo,P. and Manley,J.L. (1993) Functional domains of the human splicing factor ASF/SF2. *EMBO J.*, **12**, 4727–4737.

Zuo,P. and Manley,J.L. (1994) The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5′ splice sites. *Proc. Natl Acad. Sci.*, **91**, 3363–3367.

Received on March 13, 1995; revised on April 25, 1995