Role of SR protein modular domains in alternative splicing specificity in vivo

Citation for published version:

Digital Object Identifier (DOI):
10.1093/nar/28.24.4822

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Nucleic Acids Research

Publisher Rights Statement:
This is a publisher-authenticated version of the article. It is also available online at:
http://nar.oxfordjournals.org/content/28/24/4822.full

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Role of SR protein modular domains in alternative splicing specificity in vivo

Willemien van der Houwen van Oordt, Kathryn Newton, Gavin R. Screaton1 and Javier F. Cáceres*

MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK and 1Molecular Immunology Group, Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

Received September 22, 2000; Revised and Accepted October 20, 2000

ABSTRACT

The SR proteins constitute a family of nuclear phosphoproteins which are required for constitutive splicing and also influence alternative splicing regulation. They have a modular structure consisting of one or two RNA recognition motifs (RRMs) and a C-terminal domain, rich in arginine and serine residues. The functional role of the different domains of SR proteins in constitutive splicing activity has been extensively studied in vitro; however, their contribution to alternative splicing specificity in vivo has not been clearly established. We sought to address how the modular domains of SR proteins contribute to alternative splicing specificity. The activity of a series of chimeric proteins consisting of domain swaps between different SR proteins showed that splice site selection is determined by the nature of the RRMs and that RRM2 of SF2/ASF has a dominant role and can confer specificity to a heterologous protein. In contrast, the identity of the RS domain is not important, as the RS domains are functionally interchangeable. The contribution of the RRMs to alternative splicing specificity in vivo suggests that sequence-specific RNA binding by SR proteins is required for this activity.

INTRODUCTION

The SR proteins are a family of structurally and functionally related polypeptides that play an important role in constitutive and regulated pre-mRNA splicing (reviewed in 1–4). They are involved in multiple steps of the constitutive splicing reaction: among other functions they promote the assembly of the earliest pre-spliceosomal complex E (5), bridge the 5′ and 3′ splice sites, recruit the U1snRNP particle to the 5′ splice site and participate in recruitment of the U4/U6-U5 tri-snRNP (for a review see 6). In addition, the discovery that some SR and SR-like proteins remain associated with the mRNA products after the splicing reaction and that a subset of SR proteins shuttle from the nucleus to the cytoplasm suggested that SR proteins may have roles not only in nuclear pre-mRNA splicing, but may also have additional functions such as mRNA transport or be involved in cytoplasmic events (7–9).

The role of SR proteins in alternative splicing regulation is antagonised by members of the hnRNP A/B family of proteins, in such a way that increased levels of SR proteins lead to the selection of proximal 5′ splice sites, whereas an excess of hnRNP A/B proteins promotes the selection of distal 5′ splice sites (10–13). Thus, the relative level and activity of members of the SR and hnRNP A/B families of proteins may represent an important determinant of alternative splicing regulation (9). However, activities that antagonise SR protein function are not restricted to the hnRNP A/B family of proteins, as illustrated by the recent identification in Drosophila of a new antagonist of SR protein function, termed RSF1 (14). In addition, individual SR proteins can sometimes antagonise each other in alternative splice site selection, as in the case of the antagonistic effects of SF2/ASF and SC35 on the regulation of β-tropomyosin (15) and of SF2/ASF and SRp20 on the regulation of SRp20 pre-mRNA alternative splicing (16). In addition to the 10 identified members of the SR family of proteins in mammals, a class of related RS domain-containing proteins, termed SR protein-related polypeptides (SRrp) or SR-like proteins, are also involved in splicing regulation (for review see 1,17).

The SR proteins seem to be functionally redundant in constitutive splicing, as illustrated by the ability of any individual SR protein to complement an otherwise inactive cytosolic HeLa S100 extract (reviewed in 4,18). However, several differences in the ability of these proteins to regulate alternative splicing, as well as the ability of individual SR proteins to commit different pre-mRNAs to the splicing pathway, suggested that individual SR proteins may have unique functions in splicing regulation (19–23). Genetic approaches used to address this question showed that SF2/ASF was essential for cell viability in the DT40 chicken cell line (24). In addition, genetic disruption of SRp55/IB52 in Drosophila (25) and of SRp20 in the mouse (26) caused early embryonic lethality, strongly suggesting essential and unique functions for these genes. However, more recent experiments in which the whole complement of Caenorhabditis elegans SR proteins were individually depleted using RNA interference have demonstrated that certain SR proteins are functionally redundant (27).

The SR proteins have a modular structure that consists of one or two RNA recognition motifs (RRMs) in the N-terminus.

*To whom correspondence should be addressed. Tel: +44 131 467 8426; Fax: +44 131 343 2620; Email: javier.caceres@hgu.mrc.ac.uk
and a C-terminal domain rich in arginine and serine residues (termed the RS domain). An extensive structure–function analysis of SF2/ASF, which is the prototype member of the SR family of splicing regulators (28,29), indicated that both RRM1s in SF2/ASF are required for efficient binding to RNA (30,31). The use of conventional and functional SELEX protocols allowed the identification of high affinity binding sites for individual SR proteins (32–37) and led to the conclusion that SR proteins are sequence-specific RNA-binding proteins with distinct RNA-binding specificities (for a review see 18). Individual SR proteins have been found to associate with exonic splicing enhancers in many different genes, leading to activation of otherwise inefficient upstream 3′ splice sites (38–42; for reviews see 4,43). In addition, SR proteins can also negatively regulate splicing of an intron, as shown in the adenovirus late transcript, where binding of SF2/ASF to an intronic sequence inhibits splicing of a 3′ splice site (44).

The RS domains have been proposed to mediate protein–protein interactions (45) and to be important determinants of subcellular localisation and nucleo-cytoplasmic shuttling (8,46–48). They are able to activate splicing even when fused to a heterologous RRM, which led to the proposal that they function as activators of splicing (49).

An important issue is how the modular domains of SR proteins contribute to splicing specificity, both in constitutive and regulated splicing, and also how critical is sequence-specific high affinity RNA binding for the function of SR proteins in both splicing processes. It has been shown that both the RRM1 and the RS domain are essential for constitutive splicing in vitro (30,31). The role of the modular domains of SR proteins in constitutive splicing has recently been studied using domain swaps between SC35 and SF2/ASF or deletion mutants of SF2/ASF. Using a functional splicing commitment assay and in vitro splicing it was shown that substrate specificity in constitutive splicing is determined by the nature of the RRM1s and that the identity of the RS domain did not affect splicing specificity, suggesting that the RS domains may have redundant functions in constitutive splicing (23,50). These experiments clearly demonstrated that individual domains in SR proteins function as modules in constitutive splicing.

In contrast, the role of the modular domains of SR proteins in regulated splicing has been less extensively studied. Analysis of SF2/ASF deletion mutant proteins in a nuclear switch assay in vitro established that both RRM1s were necessary and sufficient to influence the selection of proximal 5′ splice sites and, surprisingly, that the RS domain was dispensable for this activity. These experiments suggested that RNA binding is required for alternative splicing (30). The alternative splicing activity of SF2/ASF has also been shown to correlate with its ability to promote multiple occupancy of alternative 5′ splice sites by U1 snRNP (51). It is not clear, however, whether sequence-specific binding to target RNAs play a role in splice site selection.

We sought to investigate the role of the modular domains of different SR proteins in alternative splicing regulation in living cells. To this end, we transiently overexpressed chimeric proteins consisting of individual domains of SF2/ASF and SRp40, SF2/ASF and SC35 and SF2/ASF and SRp20 and assayed changes in the pattern of splicing of two different reporters, the adenovirus E1A pre-mRNA and a fibronectin miniexon. We demonstrate that the RRM1s determine the alternative splicing specificity in vivo and that the RS domains are functionally interchangeable. These results demonstrate the modularity of SR proteins in alternative splicing regulation.

**MATERIALS AND METHODS**

**Epitope-tagged expression plasmids**

The epitope-tagged SR protein expression plasmids were previously described. Briefly, they were constructed by PCR amplification of cDNA clones coding for the respective SR proteins with specific primers and the resulting PCR products were subcloned into the pCGTHCFpT7 expression vector (52), to generate the pCGT7-SR constructs. Transcription is driven by the CMV enhancer/promoter and the coding sequence begins with a N-terminal epitope tag, MASMT-GGQQMGG. This epitope tag corresponds to the first 11 residues of the bacteriophage T7 gene 10 capsid protein and is recognised by the T7 tag monoclonal antibody (Novagen). For the SF2/ASF and the SC35 constructs PCR products were amplified with specific primers and subcloned into the pCGT7 expression vector as XbaI–BamHI fragments. In the case of the SRp20 constructs, due to the presence of an internal XbaI site in the SRp20 cDNA, the amplified fragments were designed with SpeI and BamHI sites and were subcloned into the XbaI and BamHI sites of the pCGT7 expression vector. In the case of the SRp40 constructs, due to the presence of an internal BamHI site in the SRp40 cDNA, the amplified fragments were designed without XbaI and BclI sites and were subcloned into the XbaI and BamHI sites of the pCGT7 expression vector. Construction of the SF2/SRp40 and SF2/SC35 chimeras was described previously (8). Briefly, by swapping domains between the corresponding SR proteins we took advantage of natural SacI and ApaI sites in the SF2/ASF cDNA, between RRM1 and RRM2 and between RRM2 and the RS domain, respectively. PCR products comprising the indicted domains were amplified from cDNA clones using specific primer pairs. The resulting PCR products were assembled using the SacI and ApaI sites and subcloned into the pCGT7 expression vector. The resulting chimeric proteins comprise the following residues: F1F240RS (previously referred to as I35RSF2RS40) comprises residues 3–197 (RRM1 and RRM2) of SF2/ASF followed by residues 184–271 (RS domain) of SF2/ASF; F1402FRS (previously termed I40IISF2RS40) consists of residues 3–197 (RRM1 and RRM2) of SF2/ASF followed by residues 184–271 (RS domain) of SF2/ASF; I35RSSF2 comprises residues 2–114 (RRM) of SC35 followed by residues 184–271 (RS domain) of SC35.

Two new chimeric proteins, F20RS and F20KRS, which comprise the first or second RRM of SF2/ASF followed by the RS domain of SRp20, were first generated and subsequently used as templates to generate the 20F1F20RS and the 20F20KRS chimeric proteins.

In the case of 20F20KRS, SF2/ASF and SRp20 were independently amplified with oligo pairs 1/2 and 3/4, respectively. The two PCR products, which have a 20 bp overlap, were used in equimolar amounts for a second PCR amplification with oligos 1 and 4. The resulting PCR product was purified,
digested with SpeI and BamHI and subcloned into the corresponding sites of the pCG-T7 vector to generate the F₁₂₀RS chimera, which comprises RRMI of SF2/ASF and the RS domain of SRp20. Subsequently, the SRp20 and F₁₂₀RS chimeras were independently amplified with oligo pairs 5/6 and 7/4, respectively. The two PCR products, which have a 20 bp overlap, were used in equimolar amounts for a second PCR amplification with oligos 5 and 4. The resulting PCR product was purified, digested with SpeI and BamHI and subcloned into the corresponding sites of the pCG-T7 vector to generate the F₂₀₀RS chimera, which consists of residues 2–85 (RRM) and 105–164 (RS domain) of SRp20 separated by residues 107–202 (RRM2) of SF2/ASF.

In the case of F₂₀₀RS, SF2/ASF and SRp20 were independently amplified with oligo pairs 8/9 and 10/4, respectively. The two PCR products, which have a 20 bp overlap, were used in equimolar amounts for a second PCR amplification with oligos 8 and 4. The resulting PCR product was purified, digested with SpeI and BamHI and subcloned into the corresponding sites of the pCG-T7 vector to generate the F₂₀₀RS chimera, which comprises RRMI of SF2/ASF and the RS domain of SRp20. Subsequently, the SRp20 and the F₂₀₀RS chimeras were independently amplified with oligo pairs 5/11 and 12/4, respectively. The two PCR products, which have a 20 bp overlap, were used in equimolar amounts for a second PCR amplification with oligos 5 and 4. The resulting PCR product was purified, digested with SpeI and BamHI and subcloned into the corresponding sites of the pCG-T7 vector to generate the F₂₀₀RS chimera, which consists of residues 2–85 (RRM) and 105–164 (RS domain) of SRp20 separated by residues 2–97 (RRM1) of SF2/ASF.

Two steps of PCR amplification were employed to generate the F₁₈₅₆ chimERIC protein expression plasmid. First, SRp20 and SF2 were independently amplified with oligo pairs 5/13 and 14/15, respectively. The two PCR products, which have a 20 bp overlap, were used in equimolar amounts for a second PCR amplification with oligo pairs 5/15. The resulting PCR product was purified, digested with SpeI and BamHI and subcloned into the corresponding sites of the pCG-T7 vector to generate the F₁₈₅₆ chimera, which comprises residues 2–85 (RRM) of SRp20 followed by residues 204–248 (RS domain) of SF2/ASF.

**Oligonucleotides**

#1, TCGACTATGTGCTCCCCGGAGGTGCTATGGC; #2, TCTCCTACAGTACAGTCGGGGCTACG; #3, CCCGACGCTATCTGAAGAGGACCTC; #4, TCAGGATCCCTATTCTTTCTTATGTA; #5, TCAGCAGGTTTTTGCTACGCTTCCGTTCA; #6, CTGGGGAGGCTTTCCTTCACATTCCG; #7, TGTTGAAAAGCTCCTCCCAGGTTCAG; #8, TCGACAGGATTGGGTCAGATT; #9, TTCTCCTACGCTCAGCTTCC; #10, AACAGGCCGAGCTGAG GAGGATGCTC; #11, TACACTCCCATTCTCATCCCATC; #12, TCTGTTAATGATCGGAGGTGTTG; #13, TGGGTACCGTTCCTTCACATTCCG; #14, TGGTTGAA AATATCGCAGGCAATCG; #15, TCTTGACTCCTTAGGTCGGA.

**Cell culture and transfections**

HeLa cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal calf serum (FCS). Transfections of HeLa cells and purification of total RNA were as previously described (12,53). Briefly, 1 µg of expression plasmid was co-transfected into HeLa cells with 6 ng of the adenosine E1A reporter plasmid pMTE1A (54) in the presence of 20 µg of Lipofectin (Gibco BRL). The E1A gene plasmid pMTE1A used in the alternative splicing assays was described previously (12,54) and was kindly provided by B. Moran. The cells were grown to 60–75% confluence in 60 mm dishes, plasmid DNA was removed 12–16 h later and DMEM containing 10% FCS was added for an additional 24 h. RNA was extracted using the Total RNA Isolation Reagent (Advanced Biotechnologies). Total RNA (5 µg) was analysed by RT–PCR with Superscript II reverse transcriptase (Life Technologies) and AmpliTaq DNA polymerase (Perkin Elmer), as described previously (53). E1A mRNA detection was carried out with the exon 1 forward primer 5’-GTTT-TCTCCTCCGAGCCGCTCCA-3’ and the 5’-end-labelled exon 2 reverse primer 5’-CTCAAGCTCTTACCCAGAACA AGG-3’. Amplified products separated by urea–PAGE were detected by autoradiography and quantitated by PhosphorImager analysis (BAS2000; Fujix).

For analysis of fibronectin EDI splicing in vivo, 60–70% confluent Hep3B cells in 60 mm dishes were transfected with 2 µg of the fibronectin minigene driven by a mutant fibronectin promoter and 400 ng of SR expression plasmids using the Fugene transfection reagent (Roche Molecular Biochemicals) according to the manufacturer’s conditions. Analysis of the in vivo splicing reaction was carried out as previously described (55).

**Western analysis**

For western blot analysis cell lysates were made with RIPA buffer as described previously (53). The electrophoretically separated proteins were transferred onto Hybond P membranes (Amersham Pharmacia Biotech). Non-specific binding was blocked by incubating the blot with 5% non-fat dry milk in TBST (20 mM Tris, pH 7.6, 137 mM NaCl and 0.1% Tween 20). Proteins were detected by subsequent incubation with the primary monoclonal antibody anti-T7,tag (Novagen) in TBST, which recognises the epitope tag at the N-terminus of the proteins. After extensive rinsing with TBST the blots were incubated with secondary antibodies conjugated to horseradish peroxidase at a 1:7000 dilution. After further rinsing in TBST the blots were developed using ECL.

**RESULTS**

**Domain requirements for alternative splicing specificity**

In order to investigate the role of individual domains of SR proteins in alternative splicing, we analysed the activity of a series of chimeric SR proteins in regulated splicing of two reporters, adenovirus E1A pre-mRNA and a fibronectin minigene. The first series of SR chimeric proteins were created by substituting individual domains of SF2/ASF with those of SRp40 and by replacing the RS domain of SC35 with that of SF2/ASF. These chimeras are depicted in Figure 1 and have been previously described (8).

We first evaluated the effects of these chimeric SR proteins on alternative splicing of the EDI fibronectin exon in vivo. We used a well-characterised system, the α-globin/fibronectin
minigene, in which the human fibronectin EDI exon, its flanking introns and part of its neighbouring exons are inserted in the third exon of the human α-globin gene. The expression of this minigene is driven by a mutant version of the human fibronectin promoter (Fig. 2A) (56–58). This alternative splicing event involves inclusion or skipping of the alternative EDI exon, which encodes a facultative type III repeat of fibronectin (59). Inclusion of the EDI exon, also known as EDA or EIIIA, depends within the central region of the alternative exon (38,57,60) and is highly regulated during embryo development (reviewed in 61). Binding of SR proteins to this splicing cis-acting sequence mediates the SF2/ASF effect on the most proximal 5′ splice site, whereas SC35 did not significantly affect the extent of EDI inclusion (55). Two main conclusions could be extracted from the activity of the chimeric SR proteins used in this assay. First, the identity of RRM2 is of crucial importance, as a chimera in which the second RRM of SF2/ASF was replaced by that of SRp40 showed a reduction in EDI inclusion, comparable to the effect observed with wild-type SRp40 (F1,F2,F4RS, Fig. 2B and C). Likewise, the reciprocal chimera in which RRM2 of SRp40 was replaced with that of SF2/ASF was even more active than wild-type SF2/ASF in promoting EDI inclusion (40,F2,F4RS, Fig. 2B and C). In contrast, replacing the RS domain of SF2/ASF with that of SRp40 did not affect the activity of the resultant chimeric protein (F1,F2,F4RS), which behaved as wild-type SF2/ASF, suggesting that RS domains are interchangeable and do not affect alternative splicing activity and/or specificity. Likewise, the C1FRS mutant, in which the RS domain of SC35 was replaced with that of SF2/ASF, displays similar activity to wild-type SC35, again demonstrating that the RS domains are interchangeable for alternative splicing. In every case we observed that the specificity of splice site selection was determined by the RRMs and that the RS domains were functionally interchangeable. The SF2/SRp40 and SF2/SC35 chimeric proteins localise to the nuclear speckles, despite differences in their ability to shuttle from the nucleus to the cytoplasm, which was dependent on the presence of the RS domain of SF2/ASF (8). When the activities of the same chimeric proteins were analysed with the E1A splicing reporter similar conclusions were obtained. As previously reported, SF2/ASF strongly promotes selection of the most proximal 5′ splice site, giving rise to the 13S isoform, and SRp40 has a similar activity, although the switch towards the most proximal site is less pronounced (20). We found that the switch to the most proximal 5′ splice site in the E1A adenovirus reporter was more or less pronounced, depending on whether RRM2 was from SF2/ASF or from SRp40, respectively (data not shown).

To further evaluate the role of the RRMs in splice selection, and in particular of RRM2 of SF2/ASF, we made two additional chimeric constructs in which either the first (RRM1) or the second (RRM2) RRM of SF2/ASF was inserted between the natural RRM and RS domain of SRp20, converting this one RRM-into a two RRM-containing SR protein (Fig. 3A). In addition, we constructed a chimeric protein in which the RS domain of SRp20 was replaced with that of SF2/ASF. We over-expressed wild-type SF2/ASF, deletion mutants of this protein, wild-type SRp20 and the chimeric proteins described above and assayed changes in the patterns of alternative splicing of an adenovirus E1A splicing reporter. A purine-rich bidirectional splicing enhancer that is located upstream of the 12S 5′ splice site and bound by a subset of SR proteins, among them SF2/ASF, has recently been identified in the E1A pre-mRNA, but it is not clear whether this cis-acting sequence mediates the SF2/ASF effect on the most proximal 5′ splice site (62). All expressed proteins accumulated to similar levels in transfected HeLa cells, as verified by western blot analysis of whole cell lysates (Fig. 3B). Insertion of either RRM1 or RRM2 of SF2/ASF into SRp20 did not alter the subcellular localisation of the resulting chimeric proteins, which localise to the nuclear speckles (not shown).

In agreement with previous results, wild-type SF2/ASF and the mutant lacking RRM1 (F2,F4RS) strongly activated the 13S 5′ splice site, whereas SRp20 and the SF2/ASF mutant that lacks

**Figure 1.** The domain structures of SF2/ASF, SRp40, SC35 and chimeric proteins consisting of domain swaps are shown schematically. SF2/ASF (first row, A and B) and SRp40 (second row, A) are composed of two RRMs and a C-terminal RS domain, whereas SC35 (second row, B) has only one N-terminal RRM. The nomenclature is based on the origin of each domain: F indicates a domain from SF2/ASF; 40 indicates a domain from SF2/ASF; 40 indicates a domain from SRp40; C, indicates a domain from SC35. The subscript 1 or 2 indicates whether the domain derives from RRM1 or RRM2, whereas the subscript RS indicates an RS domain.
RRM2 (F1F2RS) strongly and reproducibly stimulated the 12S 5′ splice site (Fig. 4) (20,48). The rationale behind the design of the SF2/SRp20 chimeras was to test whether the presence of RRM2 of SF2/ASF in the context of the SRp20 protein could act in a dominant manner and confer SF2/ASF specificity to the chimeric protein. This was indeed the case, whereas insertion of RRM1 of SF2/ASF between the RRM and RS domain of SRp20 does not change the splicing specificity of the resulting chimera (201F120RS), insertion of RRM2 produced a drastic shift in the splice site selection since the mutant protein, 201F220RS, was as active as SF2/ASF wild-type and strongly activated the most proximal 5′ splice site that gives rise to the 13S isoform. This experiment confirmed that the nature of the RRM determines the selection of a particular 5′ splice site. The RS domain of SF2/ASF is not required for alternative splicing in vivo, as shown by the activity of the SF2/ASF mutant, which lacks the RS domain (F1F2 in Fig. 4; 48). The 20,F2RS mutant in which the RS domain of SRp20 was replaced with the corresponding domain of SF2/ASF has similar activity to wild-type SRp20, strongly suggesting that the RS domains are interchangeable. Quantitation of the relative use of the E1A 5′ splice sites upon overexpression of the different proteins is shown in Figure 4B.

Figure 3. (A) The domain structures of SF2/ASF, SRp20 and chimeric proteins consisting of domain swaps are shown schematically. SF2/ASF (first row) is composed of two RRMs and a C-terminal RS domain, whereas SRp20 has only one N-terminal RRM. The nomenclature is based on the origin of each domain: F indicates a domain from SF2/ASF; 20 indicates a domain from SRp20. The subscript 1 or 2 indicates whether the domain derives from RRM1 or RRM2, whereas the subscript RS denotes an RS domain. (B) Western blotting analysis of wild-type and chimeric SR proteins. All expressed proteins accumulated to similar levels in transfected Hep3B cells (this figure) and HeLa cells (data not shown).
Next, we analysed the effects of the SF2/SRp20 chimeric proteins described above on alternative splicing of the fibronectin splicing reporter. Hep3B cells were co-transfected with a version of the fibronectin minigene driven by the mutant fibronectin promoter and plasmids expressing cDNAs for the SR proteins SF2/ASF and SRp20 in their wild-type versions or the SF2/SRp20 chimeric proteins. Overexpression of wild-type SF2/ASF greatly enhanced EDI inclusion, whereas wild-type SRp20 had a slightly inhibitory effect. Replacing the RS domain of SRp20 with that of SF2/ASF (Fig. 5, 201FRS mutant) did not confer SF2/ASF activity to the resulting chimeric protein, confirming that the RS domains do not determine specificity in alternative splicing and are functionally interchangeable. The construct that inserts RRM1 of SF2/ASF into SRp20 did not significantly affect inclusion of EDI (Fig. 5, 201F1120RS). In contrast, insertion of RRM2 of SF2/ASF into SRp20 resulted in a shift to inclusion of EDI, but the effects were modest (Fig. 5, 201F1220RS).

In summary, we have shown that the specificity of splice site selection in living cells is determined by the nature of the RRM modules, which function as modules, and that the RS domains do
not contribute to alternative splicing specificity and are functionally interchangeable. We also found that SF2/ASF specificity in alternative splicing resides in RRM2, as was demonstrated for constitutive splicing in *vitro* (50), which suggests that the basic mechanisms of binding to targets are common in both processes. More importantly, the functions of the two RRM modules can be separated, as shown by the dominant role of RRM2 of SF2/ASF in different contexts.

**DISCUSSION**

The SR proteins function in both constitutive and alternative splicing, but the role of individual domains of SR proteins has only been analysed for the constitutive splicing reaction. Although both the RRMs and the RS domain of SR proteins are required for constitutive splicing, the substrate specificity of SR proteins in constitutive splicing is determined by the RRMs and the RS domains are functionally interchangeable (23,50).

It has not been clearly established whether general and/or sequence-specific RNA binding by SR proteins is required for their activity in alternative splicing regulation. In fact, contradictory results have been obtained using different experimental systems. Using *in vitro* assays it was shown that mutants of SF2/ASF that lack either RRM1 or RRM2 were inactive in alternative splicing, suggesting a correlation between RNA binding and alternative splicing activity (30,31). It is possible, although unlikely, that the protein tags used in these various *in vitro* and *in vivo* systems are affecting protein function. Our previous observations with wild-type untagged SR expression vectors show that these constructs give the same effects as the T7 epitope-tagged versions (20). In contrast, the same SF2/ASF mutant proteins were shown to be active in E1A pre-mRNA alternative splicing *in vivo*, although with an altered splice site specificity, suggesting that sequence-specific interactions may play a role in selecting particular splice sites for activation (48). In addition, similar SF2/ASF mutant proteins lacking either RRM1 or RRM2 were still able to facilitate U1 snRNP binding to a pre-mRNA containing a functional 5′ splice site (63). It was proposed that SF2/ASF mediates alternative splicing by promoting the indiscriminate binding of U1 snRNP to alternative 5′ splice sites and it has recently been shown that hnRNP A1 antagonises SF2/ASF action by causing reduced U1 snRNP binding to the same sites (51,64). More importantly, mutations in hnRNP A1 and SF2/ASF showed that the opposite effects of the proteins on 5′ splice site choice are correlated with their effects on U1 snRNP binding (64).

In this paper, we have sought to address how the modular domains of SR proteins contribute to alternative splicing specificity and whether sequence-specific binding to high affinity sites in the RNA is required for the function of SR proteins in regulated splicing.

Adenovirus E1A pre-mRNA generates three major mRNA isoforms that represent alternative 5′ splice site utilisation and it has been previously shown that SF2/ASF and hnRNP A1 have antagonistic functions *in vivo*, with SF2/ASF promoting the selection of the most proximal 13S 5′ splice site and hnRNP A1 selecting the most distal 95 5′ splice site (12,13). It was also shown that a mutant SF2/ASF protein that lacks RRM2, as well as wild-type SRp20, which naturally lacks a second RRM, selected the 12S 5′ splice site in transient transfection assays (48).

Here we show that insertion of RRM2 of SF2/ASF between the RRM and the RS domain of SRp20 rendered a chimeric protein that was as active as wild-type SF2/ASF in promoting selection of the most proximal 5′ splice site, whereas insertion of RRM1 of SF2/ASF did not alter the activity of SRp20. This experiment clearly demonstrates that the nature of the RRM determines splice site selection and, in particular, that RRM2 of SF2/ASF has a dominant role. These conclusions were further confirmed by the activity of the SF2/SPR40 chimeric proteins on alternative splicing of the fibronectin ED1 exon, which is dependent on binding of SR proteins to an exonic splicing enhancer. Thus, the findings in this paper show that splicing specificity is determined by the RRMs and, in particular, the identity of RRM2 is important.

These experiments demonstrate that the presence of RRM2 of SF2/ASF determines specificity of splice site selection. One likely explanation for these results is that the RNA binding specificity of the chimeric proteins was altered in such a way that SF2/ASF specificity of binding was conferred to those chimeric proteins that contain RRM2 of SF2/ASF. However, it is not clear how the RRMs in the SR proteins bind to RNA. It has been shown that both RRMs of SF2/ASF together bind with strong synergy to RNA. In addition, the high affinity binding sites selected by SELEX with an SF2/ASF protein comprising both RRMs are distinct from the sites selected with RRM1 alone, suggesting that both RRMs of SF2/ASF act together in binding RNA (32). In contrast, in other RNA-binding proteins, such as the U1snRNP-associated U1A protein, individual RRMs have their own RNA-binding specificity and interact independently with distinct RNA elements in pre-mRNA (65).

In the case of the hnRNP A1 protein, structural and functional analysis of variant proteins comprising duplications, deletions or swaps of the RRMs showed that the two RRMS have naturally evolved as part of a double RRM structure (66,67). It is highly likely that each RRM in the two RRM-containing SR proteins function as a bipartite RNA-binding domain while retaining significant modular character, similar to what has been shown for hnRNP A1 (67).

An alternative explanation for the dominant role of RRM2 of SF2/ASF in alternative splicing may lie in the establishment of new protein–protein rather than protein–RNA contacts. In such a scenario RRM2 of SF2/ASF may help recruit additional splicing factors to the spliceosome and modulate splicing specificity.

It has been shown that individual SR proteins interact with each other, with the small subunit of U2AF (U2AF35) and with the U1snRNP protein U170K (45,68). This has led to the suggestion that the SR proteins may act as bridging factors to facilitate functional interactions between the 5′ and 3′ splice sites. In addition, a subset of SR proteins interacts with a protein that binds to S/MAR regions, known as scaffold attachment factor B or SAF-B (69), and SRp20 has been shown to bind RBMp, a mammalian germ cell-specific RNA-binding protein (70).

Most of the SR protein–protein interactions have been shown to be dependent on the presence of the RS domain in the interacting proteins (45,68,71). However, the RBM–SRp20 interaction is mediated via the SRp20 RRM (70), which suggests that protein–protein interaction domains are present...
in the RRM2. It is possible that RRM2 could help mediate homodimeric interactions recruiting wild-type SF2/ASF protein or, alternatively, sequestering endogenous factors such as SRp20. This activity would explain why the sole presence of RRM2 of SF2/ASF results in EDI inclusion in the fibronectin gene and selection of the 13S 5′ splice site in the E1A adenovirus pre-mRNA, both activities that are strongly favoured by wild-type SF2/ASF protein. It should be taken into account that the alternative splicing assays, both in vivo and in vitro, are carried out in the presence of multiple endogenous SR proteins which could be recruited to the spliceosome via protein–protein interactions with the exogenously expressed proteins. Thus, it is not possible at this stage to exclude the possibility that, in addition to directly binding the substrate pre-mRNA, the RRM2 domain can also function as a protein–protein interaction domain to recruit other SR proteins involved in alternative splicing of the substrate pre-mRNA.

The RS domain is a distinctive feature of the SR family of proteins and is also present in a family of related splicing factors, the so-called SR-like proteins. It has been proposed as being a domain that directs protein–protein interactions and has also been shown to be an important determinant of subcellular localisation, but it is not clear whether individual RS domains have redundant or unique functions. In fact, contradictory results have been obtained using different experimental systems. Specific roles for the RS domains of individual SR proteins have been suggested by the high phylogenetic conservation of specific sequences within the RS domains (72). The activities of chimeric proteins containing different RS domains fused to the RRM of the bacteriophage MS2 protein in the splicing of substrates containing a single MS2-binding site correlated directly with the number of RS dipeptides present in the RS domain, suggesting that different RS domains have unique activities (73). Individual RS domains have also been shown to have unique properties in directing subcellular localisation and influencing the ability of SR proteins to shuttle from the nucleus to the cytoplasm (8,46–48). However, the RS domains are functionally interchangeable in vivo, as shown by the ability of chimeric proteins consisting of the RRM2 of SF2/ASF fused to RS domains of different SR proteins to rescue cell viability in a chicken B cell line, DT40 (74). In contrast, it was recently shown that RS domains are not all functionally equivalent in vivo in Drosophila, as different RS domains varied considerably in their ability to restore Tra2 function (75).

The RS domains are essential for constitutive splicing activity, but dispensable for regulated splicing, both in vitro and in vivo (30,31,48). When its function was assayed in constitutive splicing it was shown that RS domains are functionally interchangeable and do not contribute to splicing specificity (23,50). We show here that replacing the RS domain of SF2/ASF by that of SRp40 or that of SC35 by that of SF2/ASF did not affect the activity of the resultant chimeric proteins (F1F2A0Ks and C1F8S, respectively) in alternative splicing of the fibrinogen report, suggesting that these RS domains are interchangeable and do not affect alternative splicing activity and/or specificity (Fig. 2). Similarly, replacing the RS domain of SRp20 with that of SF2/ASF (20)F8S mutant) did not alter the activity of the chimeric protein with either the E1A or the fibrinogen splicing reporters, confirming that the RS domains are interchangeable (Figs 4 and 5).

In summary, this present study has shown that substrate specificity in alternative splicing is determined by the nature of the RRM2 and that the RS domains are interchangeable, as has been demonstrated for the constitutive splicing reaction. These results demonstrate the modularity of SR proteins for their alternative splicing function in vivo. Our previous results showing different splice site activation specificities of the SF2/ASF mutants depending on which of the two RRM2s was deleted, together with the results shown here, strongly suggest that sequence-specific interactions may play a role in selecting particular splice sites for activation. The dominant role of RRM2 of SF2/ASF could be explained by two non-mutually exclusive models. It is possible that the presence of this domain confers SF2/ASF RNA binding specificity to the chimeric proteins. Alternatively, RRM2 of SF2/ASF could mediate protein–protein interactions, bringing wild-type SF2/ASF or other SR proteins into play.

ACKNOWLEDGEMENTS

We thank Alberto Kornblihtt (Buenos Aires) for the generous gift of the FN/EDI constructs. We are grateful to Ian Eperon (Leicester) and Andres Muro (Trieste) for critical reading of the manuscript. We are indebted to Douglas Stuart (MRC HGU Photography Department) for help in the preparation of figures. J.F.C. was supported by the Medical Research Council and W.v.d.H.v.O by the Wellcome Trust.

REFERENCES


