

# Role of the Modular Domains of SR Proteins in Subnuclear Localization and Alternative Splicing Specificity

Javier F. Cáceres,\* Tom Misteli,\* Gavin R. Sreaton,‡ David L. Spector,\* and Adrian R. Krainer\*

\*Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724-2208; and ‡Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

**Abstract.** SR proteins are required for constitutive pre-mRNA splicing and also regulate alternative splice site selection in a concentration-dependent manner. They have a modular structure that consists of one or two RNA-recognition motifs (RRMs) and a COOH-terminal arginine/serine-rich domain (RS domain). We have analyzed the role of the individual domains of these closely related proteins in cellular distribution, subnuclear localization, and regulation of alternative splicing in vivo. We observed striking differences in the localization signals present in several human SR proteins. In contrast to earlier studies of RS domains in the *Drosophila* suppressor-of-white-apricot (SWAP) and Transformer (Tra) alternative splicing factors, we found that the RS domain of SF2/ASF is neither necessary nor suf-

ficient for targeting to the nuclear speckles. Although this RS domain is a nuclear localization signal, subnuclear targeting to the speckles requires at least two of the three constituent domains of SF2/ASF, which contain additive and redundant signals. In contrast, in two SR proteins that have a single RRM (SC35 and SRp20), the RS domain is both necessary and sufficient as a targeting signal to the speckles. We also show that RRM2 of SF2/ASF plays an important role in alternative splicing specificity: deletion of this domain results in a protein that, although active in alternative splicing, has altered specificity in 5' splice site selection. These results demonstrate the modularity of SR proteins and the importance of individual domains for their cellular localization and alternative splicing function in vivo.

**N**UMEROUS protein and ribonucleoprotein components are required to catalyze pre-mRNA splicing, which occurs within a macromolecular complex, the spliceosome. The major components of the spliceosome, which assembles on individual pre-mRNAs, are the small nuclear ribonucleoprotein particles (snRNPs)<sup>1</sup> U1, U2, and U4/6-U5; the polypeptides that associate with hnRNA to form hnRNP particles (hnRNP proteins); and a large set of non-snRNP protein splicing factors, which includes the SR family of proteins (for reviews see Dreyfuss et al., 1993; Fu, 1995; Krämer, 1996; Manley and Tacke, 1996).

The SR proteins are essential splicing factors and also regulate alternative splicing of many pre-mRNAs by affecting the selection of 5' splice sites (Ge and Manley, 1990; Krainer et al., 1990a,b; Zahler et al., 1993; for review see Horowitz and Krainer, 1994). Their activity in alterna-

tive splicing is antagonized by members of the hnRNP A/B family of proteins (Mayeda and Krainer, 1992; Mayeda et al., 1994). Thus, the counteracting activities of these two families of antagonistic factors can regulate alternative splicing, both in vitro and in vivo (Mayeda and Krainer, 1992; Cáceres et al., 1994; Yang et al., 1994). SR proteins are characterized by the presence of one or two copies of the RNA-recognition motif (RRM) and a COOH-terminal domain rich in arginine and serine residues (RS domain) (Zahler et al., 1992; for review see Birney et al., 1993). These proteins are thus very closely related in their domain structure, primary sequence, and functional properties. Individual SR proteins interact with pre-mRNA during the early stages of spliceosome assembly (Krainer et al., 1990b; Fu, 1993; Staknis and Reed, 1994; Zahler and Roth, 1995) and stimulate binding of U1 snRNP particles to the 5' splice site (Eperon et al., 1993; Kohtz et al., 1994) and binding of U2AF to the 3' splice site (Zuo and Maniatis, 1996).

Splicing factors are distributed nonrandomly within the nucleus. Immunofluorescence localization studies showed that snRNP particles and SR proteins are organized in the interphase nucleus in a characteristic speckled pattern. Two morphologically distinct nuclear structures have been identified as constituents of the nuclear speckles by electron microscopy: interchromatin granule clusters and peri-

Please address all correspondence to Adrian R. Krainer, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724-2208. Tel.: (516) 367-8417. Fax: (516) 367-8453. E-mail: krainer@cshl.org

J.F. Cáceres' present address is MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, United Kingdom.

**1. Abbreviations used in this paper:** G-domain, glycine-rich domain; RRM, RNA-recognition motif; RS domain, arginine/serine-rich domain; RT-PCR, reverse transcriptase PCR; snRNP, small nuclear ribonucleoprotein particle; SWAP, suppressor-of-white-apricot; Tra, Transformer.

chromatin fibrils (for review see Spector, 1993). The speckled pattern comprises 20–50 regions that are highly concentrated in splicing factors (Spector, 1990; Huang and Spector, 1992). Immunofluorescence staining and immunoelectron microscopy studies showed that the splicing factor SC35, a member of the SR family of proteins, colocalizes with snRNPs within the nuclear speckle domains (Fu and Maniatis, 1990; Spector et al., 1991). In some cell types, snRNPs can be additionally found in another discrete nuclear structure, the coiled body (for review see Lamond and Carmo-Fonseca, 1993).

The nuclear organization of splicing factors is dynamic, as shown by the fact that inhibition of transcriptional and/or splicing activity causes reorganization of the nuclear speckles (Spector et al., 1983; O'Keefe et al., 1994). The most intense speckled regions are adjacent to, but not coincident with, transcriptionally active sites, as defined by fluorescent labeling of nascent RNAs (Wansink et al., 1993) and by *in situ* hybridization (Huang and Spector, 1991; Xing et al., 1993, 1995). Likewise, the interchromatin granule clusters are not sites of tritiated uridine incorporation (for review see Fakan and Puvion, 1980). The fact that splicing does not appear to take place at the majority of sites where splicing factors are most concentrated (Mattaj, 1994; Zhang et al., 1994; Huang and Spector, 1996) led to the suggestion that splicing factors shuttle between interchromatin granule clusters (sites of storage and/or reassembly) and perichromatin fibrils (sites of active transcription and splicing) (Jiménez-García and Spector, 1993). Many introns, though not all, are spliced as nascent transcripts *in vivo* (Beyer and Osheim, 1988; Bauren and Wieslander, 1994; Kiseleva et al., 1994). An intimate connection between transcription and splicing is further suggested by the observation that a subpopulation of hyperphosphorylated RNA polymerase II localizes to the nuclear speckle domains (Bregman et al., 1995; Mortillaro et al., 1996). The hyperphosphorylated COOH-terminal domain of the polymerase large subunit appears to mediate interactions with components related to splicing (Mortillaro et al., 1996; Yuryev et al., 1996; Du and Warren, 1997; McCracken et al., 1997).

Several protein kinases capable of phosphorylating SR proteins on serine residues *in vitro* have been described: SRPK1 (Gui et al., 1994a,b), Clk/Sty (Colwill et al., 1996), a lamin-B receptor-associated kinase (Nikolakaki et al., 1996), and unexpectedly, DNA topoisomerase I (Rossi et al., 1996). In addition, it has been shown that addition of SRPK1 to permeabilized cells (Gui et al., 1994a) or overexpression of Clk/Sty (Colwill et al., 1996) results in a diffuse distribution of splicing factors, probably because of hyperphosphorylation of their RS domains. It has been proposed that the level of phosphorylation may control the subnuclear distribution of SR proteins in interphase cells and the reorganization of the speckle domains during mitosis (Colwill et al., 1996; Gui et al., 1994a; Misteli and Spector, 1996).

Previous studies showed that the RS domains of two *Drosophila* splicing regulators, suppressor-of-white-apricot (SWAP) and Transformer (Tra), direct these splicing factors to the nuclear speckles (Li and Bingham, 1991; Hedley et al., 1995). We sought to investigate the role of the structural domains of SR proteins in cellular and sub-

nuclear localization and to study the requirement for individual domains of these proteins for modulation of alternative splicing *in vivo*. To this end, we transiently overexpressed in HeLa cells several epitope-tagged SR protein cDNAs encoding either the wild-type proteins or several mutant derivatives thereof. We then determined the cellular distribution of the tagged proteins, as well as their activity in regulating alternative splicing of transcripts expressed from cotransfected reporter genes. We found that SR proteins that have either one or two RRM domains differ in their requirements for the RS domain as a nuclear speckle targeting signal. We also demonstrate that the presence of RRM2 can affect the alternative splicing specificity of SF2/ASF *in vivo*, suggesting that the modular structure of SR proteins is important for regulated splicing.

## Materials and Methods

### Epitope-tagged Expression Plasmids

Oligonucleotide primers were purchased from GIBCO BRL (Gaithersburg, MD). PCR conditions using Vent polymerase (New England Biolabs, Beverly, MA) were as previously described (Krainer et al., 1991). The epitope-tagged, full-length SF2/ASF expression plasmid was constructed by amplifying an SF2/ASF cDNA (Krainer et al., 1991) with specific primers and subcloning of the resulting PCR product as an XbaI-BamHI fragment into the pCGTHCF<sub>FL</sub>T7 expression vector (Wilson et al., 1995). The resulting vector, pCGT7-SF2, like the previously described pCG-SF2 vector, is under the control of the cytomegalovirus enhancer/promoter (Tanaka and Herr, 1990; Cáceres et al., 1994) but also includes an NH<sub>2</sub>-terminal epitope tag, MASMTGGQMG. This epitope tag corresponds to the first 11 residues of the bacteriophage T7 gene 10 capsid protein and is recognized by the T7 tag monoclonal antibody (Novagen, Inc., Madison, WI). The SF2/ASF mutants and domain deletions were previously described (Cáceres and Krainer, 1993) and were subcloned into the same epitope-tagged expression vector as wild-type SF2/ASF. For the SRp40, SC35, SRp20, and hnRNP A1 constructs, PCR products were amplified with specific primers and subcloned into the same expression vector. In the case of the SRp20 constructs, because of 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-BamHI sites of pCGT7-SF2. In the case of the SRp40 constructs, because of the presence of an internal BamHI site in the SRp40 cDNA, the amplified fragments were designed with XbaI and BclI sites and were subcloned into the XbaI-BamHI sites of pCGT7-SF2. The SRp40-ΔRS protein comprises residues 2–183; SC35-ΔRS comprises residues 2–94; and SRp20-ΔRS comprises residues 2–85. Construction of the A1-RS protein was previously described (Mayeda et al., 1994); this chimeric protein comprises amino acids 1–185 from hnRNP A1 and 195–248 from SF2/ASF.

For the nucleoplasm fusion, the vector pCGT7-NPc was constructed by amplifying a *Xenopus* nucleoplasm cDNA (Dingwall et al., 1987) with specific primers and cloning the resulting PCR product into the XbaI and BamHI sites of the pCGTHCF<sub>FL</sub>T7 expression vector. This procedure results in the insertion of amino acids 2–149 of nucleoplasm followed by an XbaI site and stop codon, COOH-terminal to the T7 epitope. This nucleoplasm fragment is also known as the nucleoplasm core domain, NPc. PCR fragments comprising residues 198–248 from the RS domain of SF2/ASF or 88–164 from the RS domain of SRp20 were subcloned downstream of NPc to generate NPc-RS<sub>SF2</sub> and NPc-RS<sub>SRp20</sub>, respectively.

### Cell Culture and Transfections

HeLa cells were grown in DME supplemented with 10% FCS and transfected with 1 μg of plasmid DNA per 60-mm dish of 60–75% confluent cells, in the presence of 20 μg lipofectin (GIBCO BRL) (Cáceres et al., 1994).

### Indirect Immunofluorescence

Cells were fixed for immunofluorescence assays 24 h after transfection. Cells were washed with PBS and incubated with 3% paraformaldehyde,

0.3% Triton (in PBS) for 5 min at room temperature, followed by incubation with 3% paraformaldehyde for 30 min. The fixed cells were incubated for 1 h at room temperature with 1:500 anti-T7 monoclonal antibody (Novagen, Inc.). The cells were washed three times with PBS and incubated for 1 h at room temperature with 1:500 fluorescein-conjugated goat anti-mouse IgG (Cappel Laboratories, Malvern, PA). Double immunofluorescence labeling was performed by incubation for 1 h at room temperature with 1:500 anti-T7 monoclonal antibody (Novagen, Inc.) and 1:2,000 human anti-Sm serum. After washing, the cells were incubated with 1:50 Texas red-conjugated antimouse IgG and 1:50 fluorescein-conjugated antihuman IgG (Cappel Laboratories). Samples were observed on a microscope (model Axiovert 405M; Carl Zeiss, Inc., Thornwood, NY), and images were acquired with a cooled CCD camera (model NU200; Photometrics, Inc., Woburn, MA) using Oncor Image software (Gaithersburg, MD). For confocal microscopy, a confocal laser scanning microscope (model LSM410; Carl Zeiss, Inc.) was used. For localization of the endogenous SF2/ASF protein, a monoclonal antibody against SF2/ASF was used as culture supernatant at a 1:5 dilution.

### Quantitation of Fluorescence Images

Images were acquired under identical conditions, ensuring that the maximal signal was not saturating, and were subjected to contrast stretching, excluding the top and bottom 1% of pixels. Acquisition of images and measurements were performed using Oncor Image software 2.0.04. To measure the intensity of speckles, a modified version of a point-hit method was used (Weibel et al., 1969). Random test lines were drawn over the cell nucleus, several pairs of random test points on the test line were chosen arbitrarily, and the absolute intensity value of the test points was measured ( $X_1$ ,  $X_2$ ). A value  $S$  for the intensity of speckles was determined as the ratio between the sum of differences and the average intensity value of all points on the test line ( $X_{av}$ ):

$$S = \frac{\sum (X_2 - X_1 / X_{av})}{N}$$

$N$  represents the number of pairs of test points used for one sample. Typically, four to six pairs of test points were chosen arbitrarily on a test line, and a minimum of three test lines were used for every cell nucleus examined. For each sample, at least four nuclei were examined. All results represent values obtained from two separate experiments, and the results are average values  $\pm$  standard deviation from the pooled data.

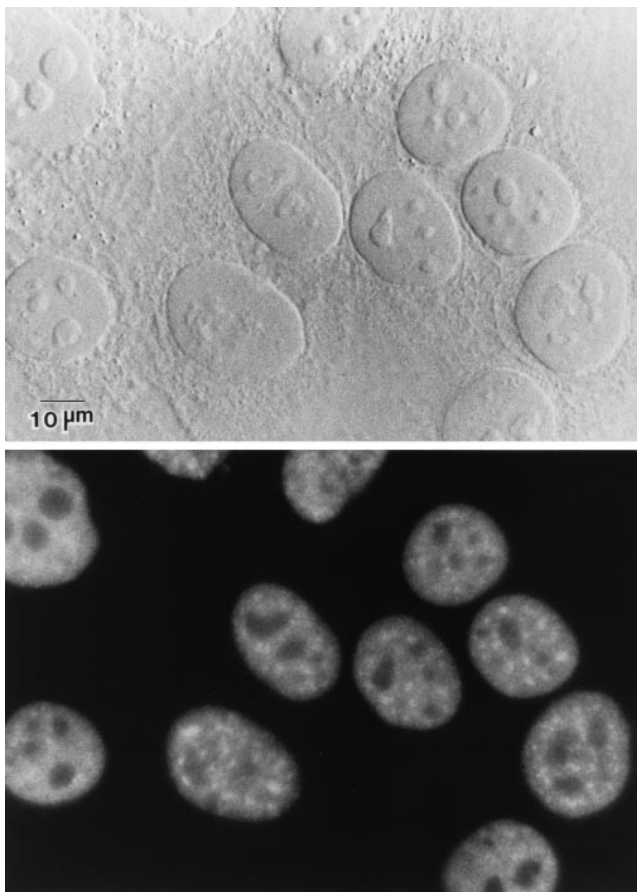
### In Vivo Analysis of Alternative Splicing

Transfections of HeLa cells and purification of total RNA were as previously described (Cáceres et al., 1994). Briefly, 1  $\mu$ g of expression plasmid was cotransfected into HeLa cells with 6  $\mu$ g of the adenovirus E1A reporter plasmid pMTE1A (Zerler et al., 1986) in the presence of 20  $\mu$ g of Lipofectin (GIBCO BRL). The cells were grown to 60–75% confluence in 60-mm dishes and harvested 24 or 48 h after transfection, and 200 ng of total RNA was analyzed by reverse transcriptase PCR (RT-PCR), as described (Screaton et al., 1995). Briefly, first-strand oligo(dT)-primed cDNA synthesized with superscript II (GIBCO BRL) from 200 ng of total RNA was amplified with Taq DNA polymerase (Perkin-Elmer Corp., Norwalk, CT) for 25 cycles, using 5'-end-labeled forward primer. Amplified products separated by urea-PAGE were detected by autoradiography and quantitated by PhosphorImage analysis (Fujix, BAS2000; Fuji Photo Film Co., Tokyo, Japan). The PCR primers were previously described (Cáceres et al., 1994).

## Results

### Role of the Structural Domains of SR Proteins in Cellular Distribution and Subnuclear Localization

First, we analyzed the localization of the endogenous SF2/ASF protein using an anti-SF2/ASF monoclonal antibody (Krainer, A.R., unpublished). This antibody does not cross-react with other SR proteins and detects all phosphorylation variants of SF2/ASF; its detailed characterization will be published elsewhere. We observed a typical nuclear speckled pattern and also a diffuse nuclear signal, but no



**Figure 1.** Localization of endogenous SF2/ASF. HeLa cells were fixed, and the cellular localization of SF2/ASF was determined by indirect immunofluorescence with an anti-SF2/ASF monoclonal antibody (Krainer, A.R., unpublished). The protein localizes in a speckled nuclear pattern and is also diffusely distributed throughout the nucleoplasm. (Top) DIC image; (bottom) fluorescence image of the same field. Bar, 10  $\mu$ m.

nucleolar staining (Fig. 1). This immunofluorescence pattern resembles that of endogenous U2B'' and Sm snRNP proteins (for review see Spector, 1993). No staining of coiled bodies was observed, in agreement with previous data on SC35 localization (Spector et al., 1991).

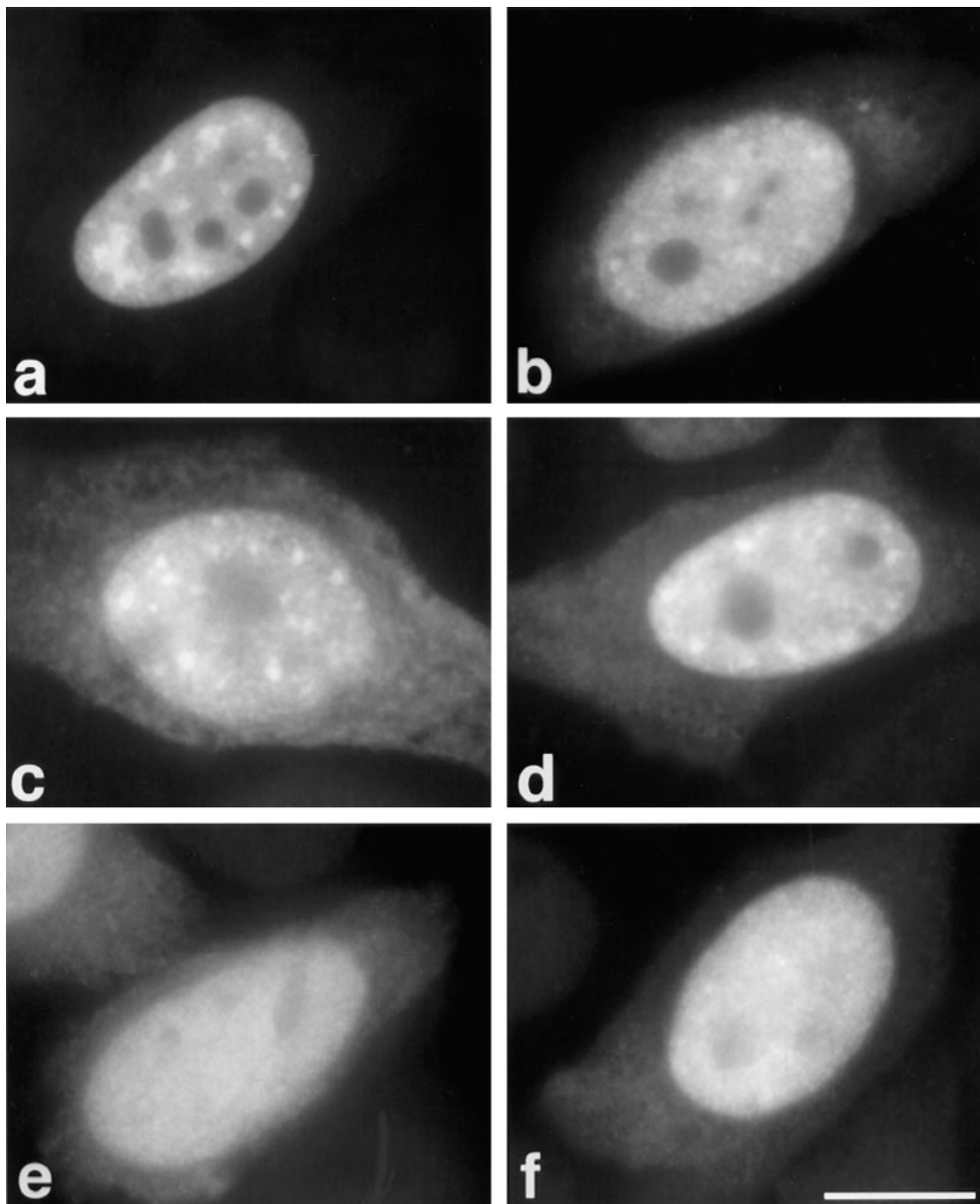
To determine the role of individual domains of SF2/ASF and other SR proteins in nuclear and subnuclear localization, we transiently overexpressed in HeLa cells several epitope-tagged SR protein cDNAs encoding either the wild-type proteins or several mutant derivatives and determined the cellular distribution and subnuclear localization of the tagged proteins by indirect immunofluorescence microscopy. We verified the expression of all the transfected cDNAs by Western blot analysis of whole-cell lysates. All constructs encoded proteins with a bacteriophage T7 gene 10 (T7) epitope tag at their NH<sub>2</sub> termini (see Materials and Methods), allowing detection of the exogenous proteins with a monoclonal antibody that recognizes this epitope. All expressed proteins accumulated to similar levels in transfected HeLa cells and appeared to be full length; in most cases, doublets were detected, which may represent different states of phosphorylation (data not

shown). In the case of SF2/ASF wild-type and mutant proteins, the use of an anti-SF2/ASF monoclonal antibody allowed us to compare the levels of exogenous and endogenous SF2/ASF in the cell population. After normalizing for transfection efficiency, we estimate that the transiently expressed proteins accumulate in the transfected cells at 5–10-fold higher levels than endogenous SF2/ASF, which is in good agreement with our previous results (Cáceres et al., 1994; and data not shown).

Indirect immunofluorescence using the T7 tag antibody showed that 24 h after transfection, the transiently expressed wild-type SF2/ASF protein localized exclusively in the nucleus, with a typical speckled pattern (Fig. 2 *a*). 48 h after transfection, SF2/ASF overexpression resulted in the formation of nuclear aggregates: the speckles were fewer in number and appeared enlarged (data not shown), in agreement with previous reports (Hedley et al., 1995; Romac and Keene, 1995). The remaining localization data

were obtained at the shorter time point, when the nuclear staining pattern of the transiently expressed wild-type protein coincides with that of the endogenous protein in untreated cells. This concordance suggests that transient expression is a valid approach to study the localization signals of SR proteins. Although the transfected proteins are expressed at higher levels, we note that the natural abundance of individual SR proteins fluctuates considerably in different cell types and that modulation of the levels of these proteins appears to be the key to their role in alternative splicing (Ge and Manley, 1990; Krainer et al., 1990a; Zahler et al., 1993; Cáceres et al., 1994).

The localization of several transiently expressed SF2/ASF domain-deletion mutants was then compared to that of the wild-type protein (SF2-WT). The structures of the mutant proteins, designated  $\Delta$ RS, RRM1/RS, RRM2/RS, RRM1, and RRM2, have been previously described (Cáceres and Krainer, 1993). (RRM2 has sometimes been



**Figure 2.** Role of SF2/ASF structural domains in cellular localization and distribution in the nuclear speckles. The structure of the SF2/ASF domain-deletion mutants was previously described (Cáceres and Krainer, 1993). (*a*) SF2-WT, (*b*) SF2- $\Delta$ RS, (*c*) RRM1/RS, (*d*) RRM2/RS, (*e*) RRM1, (*f*) RRM2. HeLa cells were transfected with plasmids encoding the respective T7-tagged proteins and fixed 24 h after transfection. The localization of the expressed proteins was determined by indirect immunofluorescence with anti-T7 monoclonal antibody and FITC-conjugated secondary antibody. Bar, 5  $\mu$ m.

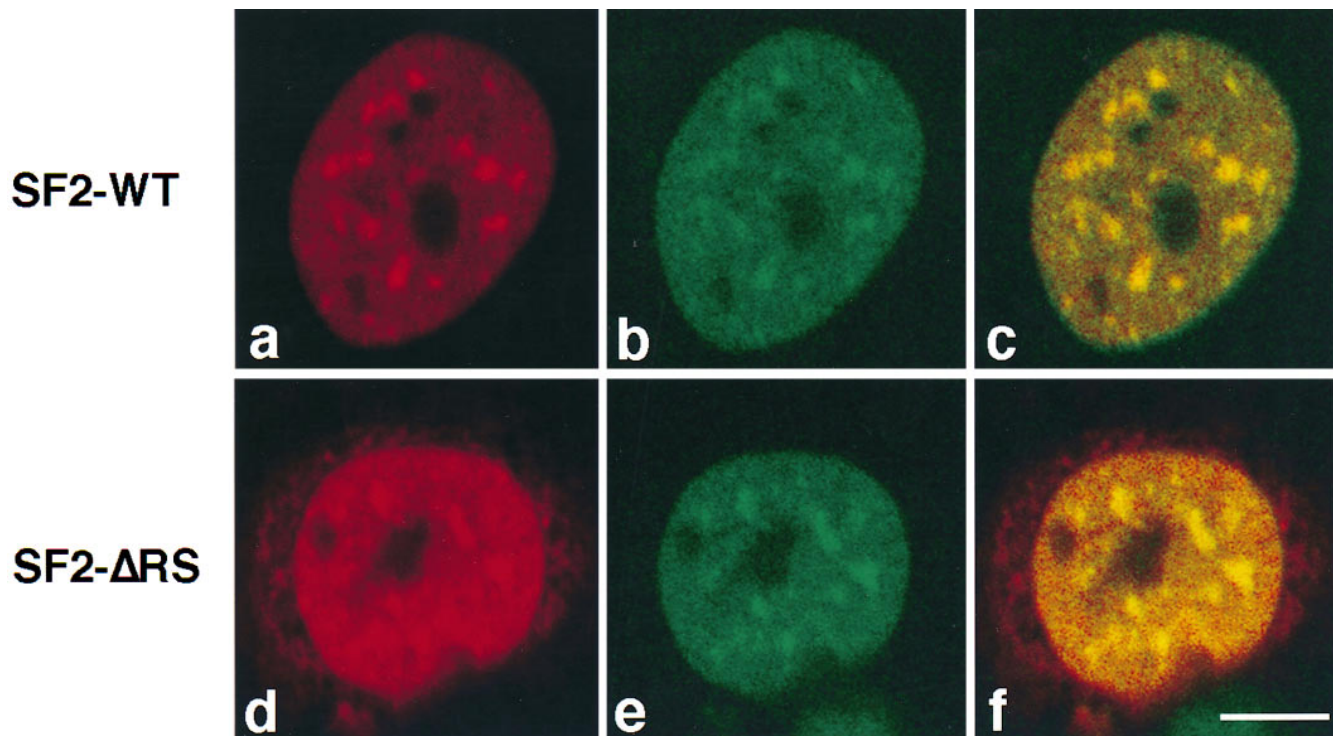
referred to as pseudo-RRM,  $\Psi$ RRM, or RRM-homology, because of its lack of canonical RNP-1 and RNP-2 submotifs; see Birney et al., 1993.) In contrast to wild-type SF2/ASF, the  $\Delta$ RS mutant protein localized mostly in the nucleus (excluding the nucleoli) but also in the cytoplasm; unexpectedly, nuclear speckles were still clearly detectable (Fig. 2 *b*). This observation demonstrates that the RS domain of SF2/ASF, although required for exclusive nuclear localization, is not essential for subnuclear localization to the speckles. The RRM1/RS and RRM2/RS mutant proteins displayed a similar cellular distribution to that observed with the  $\Delta$ RS mutant protein, and they also localized to speckles (Fig. 2, *c* and *d*). In contrast, when individual domains were expressed (RRM1 or RRM2), the mutant proteins localized throughout the cell, and no nuclear speckles were detected (Fig. 2, *e* and *f*). Expression of just the epitope-tagged RS domain from SF2/ASF was very inefficient and could not be detected by immunofluorescence (data not shown).

The transiently expressed SF2-WT and SF2- $\Delta$ RS proteins colocalized with snRNP particles in the endogenous speckles, as shown by confocal laser scanning microscopy using the anti-T7 tag mouse monoclonal antibody (red) (Fig. 3, *a* and *d*) and a human autoimmune serum specific for the Sm core proteins of snRNP particles (green) (Fig. 3, *b* and *e*). Colocalization of SF2/ASF proteins and snRNPs results in a yellow color (Fig. 3, *c* and *f*). We con-

clude that transiently expressed SF2-WT and a fraction of SF2- $\Delta$ RS proteins localize properly to nuclear speckles containing endogenous snRNPs.

These experiments demonstrate that all three structural domains of SF2/ASF are required for exclusive nuclear localization since deletion of the RS domain, or of either RRM, resulted in proteins with both nuclear and cytoplasmic distribution. However, there appears to be functional redundancy in the speckle localization signals within SF2/ASF. At least two domains are necessary for SF2/ASF localization to nuclear speckles: either two copies of the RRM or one of the two RRMs together with the RS domain.

Next, we analyzed the effect of deleting the RS domain from another SR protein that has two RRMs (SRp40), and from two SR proteins that have a single RRM (SC35 and SRp20). To this end, these epitope-tagged proteins (either wild-type proteins or mutant proteins lacking the RS domain) were transiently expressed in HeLa cells, and their cellular localization was analyzed by indirect immunofluorescence. Wild-type SRp40 localized exclusively to the nucleus and showed a typical speckled pattern (Fig. 4 *a*). Deletion of its RS domain resulted in localization both in the nucleus (excluding the nucleoli) and in the cytoplasm; however, nuclear speckles were clearly detected (Fig. 4 *b*), as was the case with SF2/ASF lacking its RS domain. Both SC35 and SRp20 wild-type proteins also localized exclu-



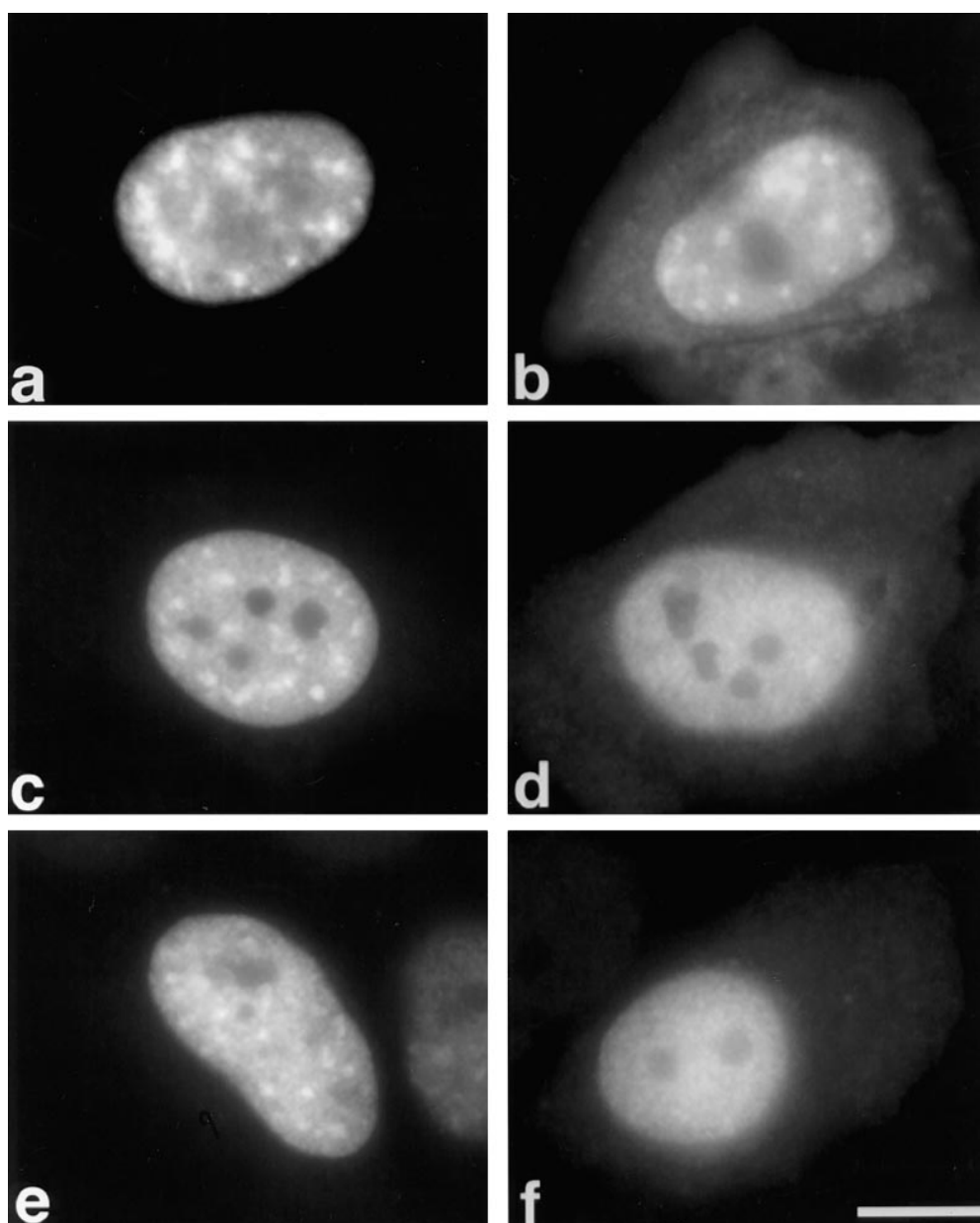
**Figure 3.** Colocalization of transiently expressed SF2/ASF, with or without its RS domain, with endogenous snRNPs. HeLa cells were transiently transfected with plasmids expressing T7-tagged, wild-type SF2/ASF (SF2-WT; top row) or a derivative lacking the RS domain (SF2- $\Delta$ RS; bottom row). The cells were fixed 24 h after transfection and analyzed by double-label immunofluorescence using laser scanning confocal microscopy. SF2/ASF or SF2- $\Delta$ RS were detected using an anti-T7 mouse monoclonal antibody followed by Texas red-conjugated secondary antibody (*a* and *d*). Endogenous snRNPs were detected in the same cell using an anti-Sm human serum followed by FITC-conjugated secondary antibody (*b* and *e*). Both SF2-WT and SF2- $\Delta$ RS colocalized with endogenous snRNPs in nuclear speckles, as shown by the yellow color when the two signals were superimposed (*c* and *f*). Bar, 5  $\mu$ m.

sively to the nucleus and displayed a typical nuclear speckled pattern (Fig. 4, *c* and *e*, respectively). In contrast, SC35 $\Delta$ RS and SRp20 $\Delta$ RS localized throughout the cytoplasm and nucleoplasm, and no nuclear speckles were detected (Fig. 4, *d* and *f*); this cellular distribution is very similar to that observed with the RRM1 and RRM2 single-domain mutants of SF2/ASF (Fig. 2). These observations indicate a significant difference in the localization signals present in two SR proteins that have two copies of the RRM (SF2/ASF and SRp40), and two SR proteins that have only one RRM (SC35 and SRp20), since only in the latter case is the RS domain absolutely required for localization to the speckle domains. Quantitation of the fluorescence signals confirmed these qualitative observations. The relative intensity of nuclear speckles was reduced by 30% in constructs lacking one of the three domains and further reduced by 80% in those constructs lacking two domains. Loss of one

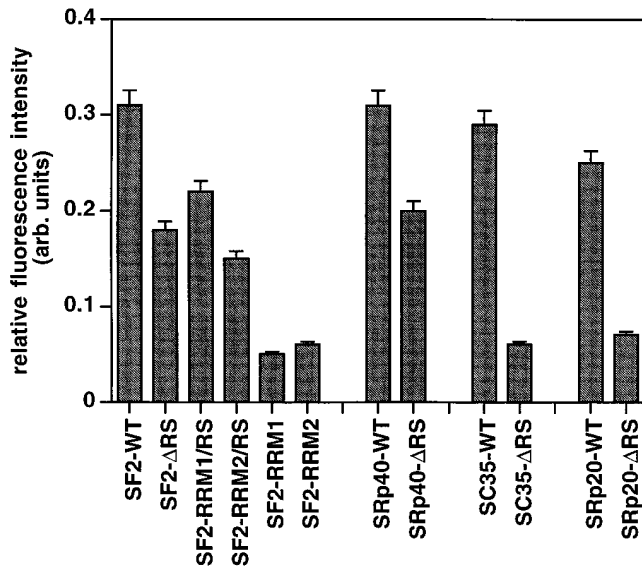
domain from two-domain SR proteins had the same effect as loss of two domains from a three-domain SR protein. The relative intensity values were similar for wild-type SRp40, SC35, and SRp20 as compared to SF2/ASF, and were similarly decreased for SC35- $\Delta$ RS and for SRp20- $\Delta$ RS, as compared to SF2/ASF RRM1 or RRM2 mutants (Fig. 5).

The finding that the RS domain of SF2/ASF is not required for speckle localization was unexpected since previous work showed that the RS domains of *Drosophila* Tra and SWAP (which have no RRMs) are necessary and sufficient for targeting to speckles (Li and Bingham, 1991; Hedley et al., 1995).

Although the RS domain of SF2/ASF is not required to target SF2/ASF to the speckles, we tested whether it can target a different protein to this subnuclear domain. We assayed the subnuclear localization of a chimeric protein,



**Figure 4.** Cellular localization of other transiently expressed SR proteins with or without their RS domains. HeLa cells were transfected with plasmids encoding the following epitope-tagged proteins: (a) wild-type SRp40, (b) SRp40- $\Delta$ RS, (c) wild-type SC35, (d) SC35- $\Delta$ RS, (e) wild-type SRp20, and (f) SRp20- $\Delta$ RS. The cells were fixed 24 h after transfection and the localization of the expressed proteins was determined as in Fig. 1. Bar, 5  $\mu$ m.



**Figure 5.** Quantitative analysis of localization of transfected proteins in nuclear speckles. The relative fluorescence intensity was calculated as described in Materials and Methods. Note that the relative intensity of speckles decreased by  $\sim 30\%$  for SF2/ASF lacking one of its three domains, and by  $\sim 80\%$  when two of the three domains were missing. Results represent average values  $\pm$  standard deviation from the pooled data of at least two experiments.

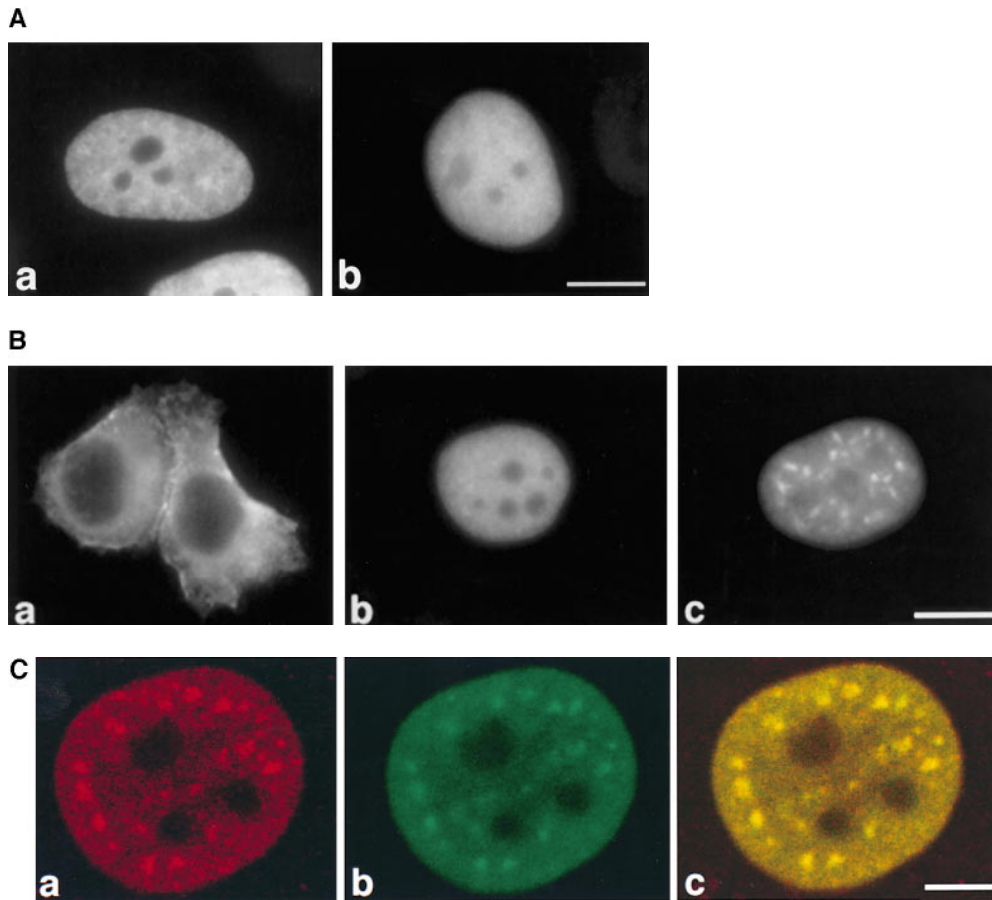
A1-RS, which consists of both RRMs from hnRNP A1 and the RS domain from SF2/ASF replacing the natural COOH-terminal glycine-rich domain (G-domain) of hnRNP A1 (Mayeda et al., 1994). This domain of hnRNP A1 contains a short signal for bidirectional transport between the nucleus and the cytoplasm (Michael et al., 1995; Siomi and Dreyfuss, 1995; Weighardt et al., 1995). The hnRNP A1 wild-type protein gave a diffuse nucleoplasmic staining rather than a speckled pattern (Fig. 6 A, a), in agreement with published data (for review see Piñol-Roma and Dreyfuss, 1993). Replacing the G-domain of hnRNP A1 by the RS domain of SF2/ASF did not target the chimeric protein to the nuclear speckles (Fig. 6 A, b), demonstrating that, at least in this context, the RS domain of SF2/ASF is not sufficient to direct a heterologous protein to the nuclear speckles. On the other hand, the A1-RS protein localized exclusively in the nucleus, whereas hnRNP A1 lacking a COOH-terminal domain is known to localize throughout the cell (Siomi and Dreyfuss, 1995). Therefore, the RS domain of SF2/ASF is a nuclear localization signal, but not a subnuclear speckle localization signal.

Next, we compared the nuclear localization signals as well as the subnuclear targeting properties of both types of RS domain, i.e., those present in SR proteins with either one or two copies of the RRM. To this end, we constructed chimeric proteins in which the RS domain of either SF2/ASF or SRp20 was fused to a protein reporter, the nucleoplasmic core domain. This protein fragment by itself displays a cytoplasmic distribution (Fig. 6 B, a) since it lacks the nuclear localization signal present in wild-type nucleoplasmic (Dingwall et al., 1987). Both types of RS domain act as nuclear localization signals, as demonstrated

by the fact that both NPC-RS<sub>SF2</sub> and NPC-RS<sub>SRp20</sub> localized to the nucleus (Fig. 6 B, b and c). However, a clear difference was evident in the ability of these RS domains to target a protein reporter to the nuclear speckles. Whereas a diffuse nuclear staining was observed with the SF2/ASF RS domain (Fig. 6 B, b), the RS domain of SRp20 was sufficient to direct the NPC reporter to the nuclear speckles (Fig. 6 B, c). The nuclear speckles observed with the NPC reporter fused to the RS domain of SRp20 colocalized with endogenous human Sm (Fig. 6 C), as shown by confocal laser microscopy using the mouse anti-T7 tag monoclonal antibody (red) (Fig. 6 C, a) and human autoimmune serum specific for the Sm core proteins of snRNP particles (green) (Fig. 6 C, b). Colocalization of NPC-RS<sub>SRp20</sub> protein and snRNPs results in a yellow color (Fig. 6 C, c). We conclude that the transiently expressed NPC-RS<sub>SRp20</sub> protein localizes properly in nuclear speckles containing endogenous snRNPs, demonstrating that this RS domain is sufficient to determine subnuclear localization in the speckle domain. These experiments demonstrate the existence of two distinct pathways for localization of different SR proteins to the nuclear speckles and also show the unique properties of individual RS domains.

### Role of SF2/ASF Domains in Alternative Splicing Regulation In Vivo

In addition to their function as general splicing factors, SR proteins regulate alternative splicing in a concentration-dependent manner (Ge and Manley, 1990; Krainer et al., 1990a; Zahler et al., 1993). We sought to investigate the role of individual domains of SF2/ASF in the regulation of alternative splicing in vivo. To this end, we overexpressed several of the constructs described above and assayed changes in the patterns of alternative splicing of an adenovirus E1A splicing reporter. It was previously shown that SF2/ASF strongly activates the use of the proximal 13S E1A 5' splice site in vivo (Cáceres et al., 1994; Wang and Manley, 1995) and that overexpression of hnRNP A1 results in activation of the distal 9S 5' splice site (Cáceres et al., 1994; Yang et al., 1994). Compared to wild-type SF2/ASF, which gave the expected activation of the 13S 5' splice site (Fig. 7 B, lane 2), the  $\Delta$ RS protein led to only slightly lower relative levels of 13S mRNA (Fig. 7 B, lane 3), demonstrating that the RS domain of SF2/ASF is not required for alternative splicing modulation in vivo. This result is consistent with previous findings (Wang and Manley, 1995), except that in the earlier study the 9S mRNA levels remained constant for both wild-type and  $\Delta$ RS proteins, whereas in our hands they decreased, as expected from in vitro studies. The SF2/ASF mutants with a single RRM and the RS domain had strikingly different functional properties: both proteins influenced alternative splicing, but whereas RRM2/RS strongly activated the 13S 5' splice site (Fig. 7 B, lane 5), RRM1/RS strongly and reproducibly stimulated the 12S 5' splice site (Fig. 7 B, lane 4). The unexpected altered specificity observed upon deletion of RRM2 demonstrates that the nature of the RRM can influence the selection of a particular 5' splice site. Activation of the E1A 12S 5' splice site by the RRM1/RS mutant is similar to the effect of SRp20, which naturally lacks RRM2 (Fig. 7 B, lane 6; Sreaton et al., 1995). Quantitation



**Figure 6.** Cellular localization of transiently expressed chimeric proteins. HeLa cells were transfected with plasmids encoding the following epitope-tagged proteins: (A) (a) hnRNP A1; (b) A1-RS, a chimeric protein with both RRMs from hnRNP A1 fused to the COOH-terminal RS domain of SF2/ASF. (B) (a) NPc, the nucleoplasmic core domain; (b) NPc-RS<sub>SF2</sub>, a chimera consisting of NPc fused to the RS domain from SF2/ASF; (c) NPc-RS<sub>SRP20</sub>, NPc fused to the RS domain of SRp20. The cells were fixed 24 h after transfection and the localization of the expressed proteins was determined as in Fig. 1. (C) Colocalization of transiently expressed NPc-RS<sub>SRP20</sub> with endogenous snRNPs. HeLa cells were transiently transfected with plasmid expressing T7-tagged NPc-RS<sub>SRP20</sub>. The cells were fixed 24 h after transfection and analyzed by double-label immunofluorescence using confocal laser scanning microscopy. NPc-RS<sub>SRP20</sub> was detected using an anti-T7

mouse monoclonal antibody followed by Texas red-conjugated secondary antibody (a). Endogenous snRNPs were detected in the same cell using an anti-Sm human serum followed by FITC-conjugated secondary antibody (b). NPc-RS<sub>SRP20</sub> colocalized with endogenous snRNPs in nuclear speckles, as shown by the yellow color when the two signals were superimposed (c). Bar, 5  $\mu$ m.

of the relative use of the E1A 5' splice sites upon overexpression of the different proteins is shown in Fig. 7 C.

Both constructs with a single RRM and the RS domain (RRM1/RS and RRM2/RS) were also active in regulating alternative splicing of a  $\beta$ -thalassemia splicing reporter, leading in both cases to activation of the most proximal cryptic 5' splice site (data not shown); this is a similar pattern as that obtained with wild-type SF2/ASF (Cáceres et al., 1994). We and others previously reported that deletion of RRM1 or RRM2 in SF2/ASF abolished the alternative splicing activity of this protein in vitro (Cáceres and Krainer, 1993; Zuo and Manley, 1993). This apparent discrepancy with our present results may reflect incorrect folding of the mutant proteins in vitro, but not in vivo, and/or the different pre-mRNA substrates used in the two types of assays. Wang and Manley (1995) reported that small deletions within RRM2 inactivated the protein in vivo; since we find that deletion of the entire domain imparts a novel specificity on the protein in a similar assay, it is likely that the small deletions disrupted the folding of the RRM, which has a highly conserved tertiary structure (for review see Birney et al., 1993).

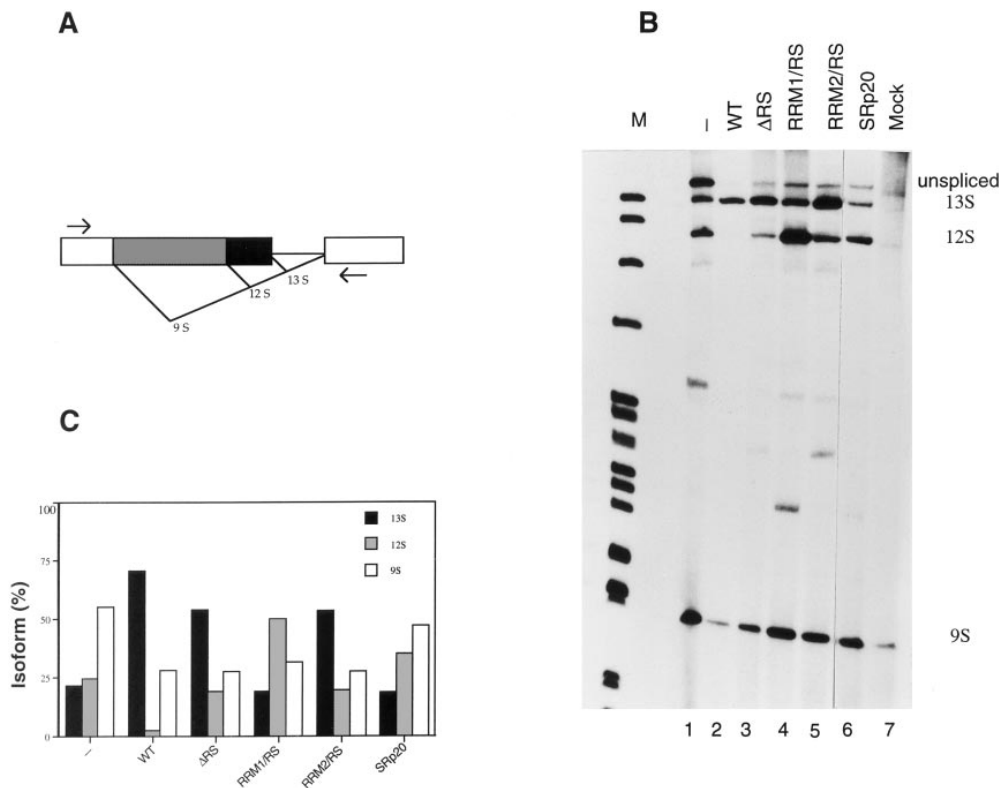
Taken together, the cotransfection results suggest that RRM2 of SF2/ASF plays a critical role in the specificity of

alternative splicing with certain substrates, such as the E1A pre-mRNA. Thus, all the functional SR protein constructs that include RRM2 (SF2/ASF WT,  $\Delta$ RS, and RRM2/RS) favored the 13S 5' splice site, whereas those that lack a second RRM (RRM1/RS and SRp20) selected the 12S 5' splice site (Fig. 8). We conclude that there is a different functional requirement for individual domains of SR proteins in the regulation of alternative splicing in living cells. Whereas the RS domain is dispensable for this activity, as was previously demonstrated in vitro (Cáceres and Krainer, 1993; Zuo and Manley, 1993), the presence of a particular RRM can promote selection of a specific splice site and also influence substrate specificity.

## Discussion

We have studied the role of individual domains of SR proteins in cellular distribution and subnuclear localization, as well as in alternative splicing activity in vivo. Unexpected differences were uncovered in the localization pathways for different SR proteins, despite the close conservation of structure and biochemical properties among members of this protein family. SR proteins with a single RRM require an RS domain for proper localization in the nuclear speck-





**Figure 7.** Role of SF2/ASF structural domains in regulating alternative splicing of adenovirus E1A pre-mRNA. (A) Diagram of the E1A reporter gene. The alternative 5' splice sites and splicing events that generate 13S, 12S, and 9S mRNAs are shown schematically. The location of the exon primers used for RT-PCR analysis is shown. (B) Alternative splicing activity of the SF2/ASF domain-deletion mutants. Each of the indicated SF2/ASF mutant and control proteins was overexpressed from plasmids cotransfected with the E1A reporter gene. RNA was harvested 24 h after transfection and analyzed by RT-PCR with a labeled forward primer, denaturing PAGE, and autoradiography, as described in Materials and Methods. The positions of the unspliced pre-mRNA, and of 13S, 12S, and 9S spliced mRNAs are indicated on the right. The back-

ground products obtained from mock-transfected cells are shown in lane 7. End-labeled DNA size markers are shown in lane M. (C) Quantitation of E1A mRNA isoforms in transfected cells. The relative amounts of 13S, 12S, and 9S E1A mRNAs were calculated from the data in B, using a PhosphorImager, and the percentage of each isoform is shown.

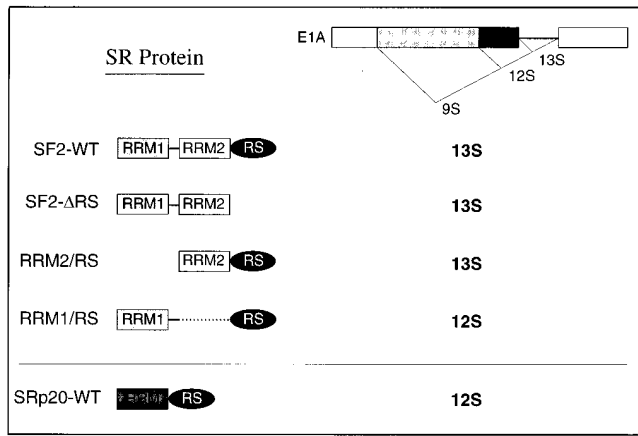
les, whereas SR proteins with two RRM domains do not; in the latter, weak speckle targeting signals are present in each of the three constituent domains and function additively. The fact that single-domain deletion variants of SF2/ASF localized properly made it possible to determine the role of each domain in alternative splicing *in vivo*. Each of the three domains could be deleted individually without abrogating the ability of SF2/ASF to modulate alternative splice site selection. Remarkably, however, deletion of RRM2 imparted a distinct activity on the protein, such that it promoted the selection of a different alternative 5' splice site in the adenovirus E1A pre-mRNA.

### Nuclear Localization of SR Proteins

The SR protein SF2/ASF has a modular structure, consisting of two RRM domains and a COOH-terminal RS domain. Here we showed that each of these three domains contributes additively to the nuclear localization of the protein since deletion of individual domains results in proteins that, unlike the wild-type protein, no longer have an exclusively nuclear distribution (Fig. 9). In particular, the RS domain of SF2/ASF contributes to its nuclear localization since deletion of this domain causes nuclear and cytoplasmic distribution of the resulting mutant protein. In agreement with this finding, when the RS domain of SF2/ASF was fused to the NPC protein reporter (which on its own localizes in the cytoplasm), the fusion protein (NPC-RS<sub>SF2</sub>) lo-

calized in the nucleus (Table I). Thus, the SF2/ASF RS domain is a nuclear targeting signal.

We do not know at present whether the nuclear and cytoplasmic distribution observed with several mutants represents incomplete nuclear import and/or incomplete retention of these proteins in the nucleus. Two alternative explanations, which are not mutually exclusive, can be proposed on the basis of our findings. In the first model, there are multiple nuclear localization signals distributed throughout SF2/ASF. These signals have additive effects, such that deletion of either RRM, or of the RS domain, reduces the overall steady-state accumulation in the nucleus. Thus, when all three structural domains are present, the protein has an exclusively nuclear distribution. When only two of the three domains are present, a fraction of the protein accumulates in the cytoplasm, and when only one domain is present, the proportion of cytoplasmic protein greatly increases. In the second model, partition between the nucleus and the cytoplasm can be attributed to incomplete nuclear retention of SF2/ASF mutants lacking either RRM repeat or the RS domain. In support of this model, we have shown that fusing the RS domain to a cytoplasmic reporter, NPC, directs the fusion protein exclusively to the nucleus. NPC is thought to have a propensity to be retained in the nucleus, once it gets there, because of oligomerization (Michael et al., 1995). In contrast, the SF2/ASF mutant proteins that lack one of the two RRM domains (RRM1/RS, RRM2/RS) display mostly nuclear but also cytoplasmic lo-



**Figure 8.** Summary of the role of SR protein domains on E1A alternative splicing specificity in vivo. The domain structures of SF2/ASF and domain-deletion mutants thereof, and that of SRp20, are shown schematically. The predominant E1A isoform generated upon overexpression of each protein is shown on the right. The presence of RRM2 of SF2/ASF correlates with use of the proximal 13S 5' splice site, whereas proteins that lack this atypical RRM cause preferential use of the 12S 5' splice site.

calization, and they have been shown to have reduced RNA binding, compared to the wild-type protein. SF2/ASF RRM domains expressed individually distribute evenly throughout the whole cell (Fig. 2) and bind RNA very poorly (Cáceres and Krainer, 1993; Zuo and Manley, 1993).

In the case of SR proteins that in their natural form have a single RRM (e.g., SC35 and SRp20), deletion of the RS domain resulted in mutant proteins with nuclear and cytoplasmic distribution (Fig. 9). Therefore, the nuclear localization mechanism appears to be different for the two SR protein subfamilies: in the case of SC35 and SRp20, a single RRM together with the RS domain resulted in exclusive nuclear localization. In contrast, in the case of SF2/ASF, either RRM together with the RS domain resulted in nuclear and cytoplasmic distribution (RRM1/RS and RRM2/RS; Fig. 2), and exclusive nuclear localization was only achieved when both RRMs were present together with the RS domain.

### Localization of SR Proteins to Nuclear Speckles

The process of nucleo-cytoplasmic transport has been extensively studied, and several components of this pathway have been identified (for reviews see Silver, 1991; Dingwall and Laskey, 1991; Görlich and Mattaj, 1996). In contrast, the mechanisms involved in localization of proteins within specific subnuclear regions are poorly understood. They may involve active transport mediated by subnuclear targeting signals, or alternatively, passive diffusion and binding to the nuclear matrix or its components. It has been proposed that the speckle domains are anchored directly or indirectly to the nuclear matrix (Spector et al., 1983). In agreement with this hypothesis, a role for the nuclear matrix in splicing has been postulated (Zeitlin et al., 1987), and antibodies raised against components of the nuclear matrix cross-react with SR proteins (Blencowe et al., 1994, 1995).

The RS domain is the most prominent feature shared by splicing factors that localize in a speckled pattern. As such, it is the best candidate domain to mediate this subnuclear localization, as first proposed and tested with *Drosophila* proteins by Bingham and coworkers (Li and Bingham, 1991). In the case of human SF2/ASF, we have shown that the RS domain is neither necessary nor sufficient for localization to the speckles. In marked contrast, deleting the RS domain of two human SR proteins that have a single RRM caused the mutant proteins to distribute throughout the cell, and accumulation in the speckled region was no longer observed. A similar result was obtained when expressing individual domains of SF2/ASF (RRM1 and RRM2 proteins). In addition, by fusing the RS domain of SRp20 to a cytoplasmic protein reporter, we demonstrated that this RS domain is necessary and sufficient to target a protein to the nuclear speckle domains (Table I). These results demonstrate the existence of two different mechanisms for localization to the speckle domains, based on the different behavior of two types of RS domains.

The lack of speckle-targeting signals in the RS domain of human SF2/ASF was unexpected, in light of previous findings with the RS domains of the *Drosophila* SWAP and Tra splicing regulators, which are necessary and sufficient to target reporter proteins to the speckled region (Li and Bingham, 1991). A recent study further defined the specific elements within the Tra RS domain required for localization in the speckled region (Hedley et al., 1995). In addition to a classical bipartite nuclear localization sequence, a novel motif was defined, which is necessary and sufficient for subnuclear localization. The subnuclear targeting signal of Tra and homologous motifs present in *Drosophila* Tra2 and SWAP, and in human U1-70K and SC35, consist of three or four basic amino acids (generally arginine and histidine) followed by a run of RS dipeptides (Hedley et al., 1995). Interestingly, this motif is present in human SR proteins that have a single RRM (SC35, SRp20 and 9G8) (Ayane et al., 1991; Fu and Maniatis, 1992; Vellard et al., 1992; Cavaloc et al., 1994), but it is absent from some, though not all, of the SR proteins with two RRMs (SF2/ASF, SRp40 and SRp30c) (Ge et al., 1991; Krainer et al., 1991; Srean et al., 1995).

The *Drosophila* Tra and SWAP alternative splicing factors have RS domains, but they lack RRMs. In contrast, the modular structure of SR proteins, with a COOH-terminal RS domain of variable length and one or two RRMs, is likely to allow multiple protein-protein and protein-RNA interactions. The mechanism of nuclear speckle localization may be a complex process, by analogy to the pathways described for localization in coiled bodies or in nucleoli (Schmidt-Zachmann and Nigg, 1993; Bohmann et al., 1995; Scheer and Weisenberger, 1995), which involve complex signals rather than simple sequence motifs.

Studies with the *Drosophila* splicing regulators suggested a mechanism for subnuclear localization based on protein-protein interactions mediated by the RS domains. Recent studies pointed to a role for the RS domains of *Drosophila* and/or human Tra, Tra2, U2AF<sup>35</sup>, U1-70K, and those of several SR proteins in mediating specific protein-protein interactions, which may be modulated by phosphorylation of many of the serine residues (Wu and Maniatis, 1993; Amrein et al., 1994; Kohtz et al., 1994;

SR Protein	Cellular Location	Nuclear Speckles
SF2-WT	N	+++
SF2-ΔRS	N + C	++
RRM1/RS	N + C	++
RRM2/RS	N + C	++
RRM1	N(diffuse) + C	-
RRM2	N(diffuse) + C	-
SRp40-WT	N	+++
SRp40-ΔRS	N + C	++
SC35-WT	N	+++
SC35-ΔRS	N(diffuse) + C	-
SRp20-WT	N	+++
SRp20-ΔRS	N(diffuse) + C	-

Figure 9. Summary of the localization of SR protein variants. The domain structures of the SF2/ASF wild-type and mutant proteins are shown schematically above the line, and those of SRp40, SC35, and SRp20, with or without an RS domain, are shown below the line. The observed cellular location of the expressed proteins is indicated as N (nuclear) or C (cytoplasmic). In all cases, the expressed proteins were excluded from nucleoli. Subnuclear accumulation in speckles is indicated by + or - signs, with the number of + signs reflecting the quantities shown in Fig. 5. Proteins that did not accumulate in speckles gave a diffuse nuclear pattern.

Xiao and Manley, 1997). In particular, the RS domains of SC35 and SF2/ASF are thought to mediate specific interactions with the RS domains of both the U1-70K polypeptide and with the 35-kD subunit of the splicing factor U2AF (Wu and Maniatis, 1993; Kohtz et al., 1994). These interactions are thought to be involved in defining and bridging the splice sites during spliceosome assembly, but in principle they could also play a role in specifying subnuclear localization. Although both snRNP polypeptides and SR proteins colocalize in a speckled distribution, their localization appears to involve different molecular interactions since the speckled distribution of snRNP proteins is sensitive to RNase A treatment, while the SC35 distribution is not affected (Spector et al., 1991).

Proteins that lack RS domains, such as the splicing factor PSF, are thought to localize to the speckled region by interacting with snRNPs or with RS domain-containing proteins (Patton, J., personal communication; Hedley et al., 1995). Deletion mutants of Tra lacking the RS domain

can still localize in the speckles, provided that they are able to interact with Tra2 (Hedley et al., 1995). Thus, two mechanisms may operate to target splicing components to the nuclear speckles: a direct one, in which the presence of a targeting signal determines the proper subnuclear localization, and an indirect one, in which proteins lacking a targeting signal localize to the speckle domains by binding to splicing components that have a targeting signal. If SF2/ASF is targeted to the speckles indirectly, by protein-protein or RNA-protein interactions, the interaction regions must be redundant since we showed that each of the three constituent domains can be deleted individually without a complete loss of localization in the speckle domains. Currently, the RS domain of SF2/ASF is the only region of the protein known to be involved in protein-protein interactions, but this domain is neither necessary nor sufficient for speckle localization of SF2/ASF. Targeting by interaction with RNA is a distinct possibility; to explain the present localization data, this model would require that a single RRM derived from an SR protein be unable to interact stably with RNA, and that either a second RRM or an RS domain be required to stabilize interactions with RNA. In vitro RNA-binding experiments showed that these are indeed the properties of SF2/ASF (Cáceres and Krainer, 1993; Zuo and Manley, 1993; Jamison et al., 1995). The direct targeting model is more likely for single-domain SR proteins, such as SC35, whose localization is RNase resistant (Spector et al., 1991), and which appear to have short sequence motifs that are sufficient for correct subnuclear localization (Hedley et al., 1995).

#### Alternative Splicing Activity of SR Proteins

We analyzed the role of the structural domains of SF2/

Table I. Summary of Localization Of Chimeric Proteins

Protein	Cellular location	Nuclear speckles
A1-RS	N	-
NPc	C	-
NPc-RS <sub>SF2</sub>	N	-
NPc-RS <sub>SRp20</sub>	N	+

The A1-RS protein consists of the two RRMs from hnRNP A1 fused to the COOH-terminal RS domain of SF2/ASF. NPc is the nucleoplasmic core domain (Dingwall et al., 1987). This domain was fused to the SF2/ASF RS domain (NPc-RS<sub>SF2</sub>) or to the SRp20 RS domain (NPc-RS<sub>SRp20</sub>). The observed cellular distribution was either nuclear (N) or cytoplasmic (C). Subnuclear accumulation in speckles is indicated by + or - signs.

ASF in regulation of alternative splicing in living cells (Fig. 8). We found that the RS domain of SF2/ASF is not required for regulation of alternative splicing *in vivo*, since a mutant protein lacking this domain regulates alternative splice site selection of an E1A splicing reporter in a manner very similar to the wild type. This finding is consistent with previous observations made *in vitro* and *in vivo* (Cáceres and Krainer, 1993; Zuo and Manley, 1993; Wang and Manley, 1995). Whereas the *in vitro* studies also showed that the RS domain of SF2/ASF is required for its full activity in constitutive splicing, the observation that it is not required for concentration-dependent effects on alternative splicing may suggest that for this particular function, the lack of the RS domain can be compensated by interactions with other SR proteins. As noted above, however, the only protein-protein interaction region of the protein known so far is, in fact, the RS domain (Wu and Maniatis, 1993; Amrein et al., 1994; Kohtz et al., 1994).

We showed that the presence of a particular domain, RRM2 of SF2/ASF, confers selectivity for a specific alternative 5' splice site. Thus, all the constructs we analyzed that included RRM2 favored the most proximal 5' splice site in E1A pre-mRNA (13S), whereas natural (SRp20) or mutant (SF2-RRM1/RS) proteins lacking RRM2 selected the middle 5' splice site (12S). This remarkable difference in splice site selection demonstrates that the RRMs of SF2/ASF function as modules that contribute to specificity in alternative splicing, and that the functions of the two RRM modules can be separated. Moreover, this effect is substrate specific since both the mutant SF2/ASF lacking RRM2 and SRp20 had altered specificity (compared to the other proteins) with the E1A pre-mRNA but showed normal specificity with a  $\beta$ -thalassemia pre-mRNA (Screaton et al., 1995; data not shown). We conclude that the selection of alternative splice sites depends on both the properties of the pre-mRNA as well as on the presence of particular domains in the SR proteins. However, not all SR proteins that possess or lack RRM2 will necessarily have the same alternative splicing specificities as SF2/ASF and SRp20, respectively. For example, overexpression of SC35, which has a single RRM, fails to activate 12S splicing (Wang and Manley, 1995; data not shown). These substrate differences among SR proteins are consistent with the notion that they may function as global regulators of alternative splicing for distinct classes of pre-mRNAs *in vivo* (Screaton et al., 1995).

The mechanistic relation between general or specific RNA binding by SR proteins and their activity in alternative splicing regulation has not been established. Although both RRMs in SF2/ASF are required for efficient binding to RNA (Cáceres and Krainer, 1993; Zuo and Manley, 1993) and to high affinity sites (Tacke and Manley, 1995), our present data show that mutants of SF2/ASF lacking either RRM1 or RRM2 are nevertheless active in alternative splicing *in vivo*. The fact that alternative splicing activity can be observed *in vivo* with SF2/ASF mutants lacking any one of its three constituent domains suggests that this activity of the protein does not require highly sequence-specific binding to RNA. On the other hand, the different splice site activation specificity of the SF2/ASF mutants, depending on which of the two RRMs was deleted, sug-

gests that sequence-specific interactions may play a role in selecting particular splice sites for activation. The alternative splicing activity of SF2/ASF has been shown to correlate with its ability to promote multiple occupancy of alternative 5' splice sites by U1 snRNP (Eperon et al., 1993). Interestingly, SF2/ASF mutants lacking either RRM are still able to form a ternary complex with U1 snRNP and pre-mRNA containing a functional 5' splice site (Jamison et al., 1995). The activity of similar mutants in alternative splicing *in vivo* is consistent with this finding.

In summary, we have shown that the modular structure of SR proteins has profound consequences for their cellular localization, alternative splicing activity, and also their specificity in alternative 5' splice site selection.

We thank Angus Wilson for the T7-expression vector, C. Dingwall for the nucleoplasm cDNA, and Mario Gimona and Sui Huang for helpful discussions.

T. Misteli was supported by the Swiss National Science Foundation and the Human Frontiers Science Program. G.R. Screaton was supported by the Wellcome Trust and the Arthritis and Rheumatism Council. D.L. Spector is supported by a grant from NIH/NIGMS (GM 42694). J.F. Cáceres and A.R. Krainer were supported by grants GM42699 from NIH/NIGMS and CA13107 from NCI and by the Pew Charitable Trusts.

Received for publication 21 February 1996 and in revised form 16 May 1997.

#### References

- Amrein, H., M.L. Hedley, and T. Maniatis. 1994. The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by Transformer 2. *Cell*. 76:735-746.
- Ayane, M., U. Preuss, G. Köhler, and P.J. Nielsen. 1991. A differentially expressed murine RNA encoding a protein with similarities to two types of nucleic acid binding motifs. *Nucleic Acids Res.* 19:1273-1278.
- Bauren, G., and L. Wieslander. 1994. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell*. 76:183-192.
- Beyer, A.L., and Y.N. Osheim. 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev.* 2:754-765.
- Birney, E., S. Kumar, and A.R. Krainer. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* 21:5803-5816.
- Blencowe, B.J., J.A. Nickerson, R. Issner, S. Penman, and P.A. Sharp. 1994. Association of nuclear matrix antigens with exon-containing splicing complexes. *J. Cell Biol.* 127:593-607.
- Blencowe, B.J., R. Issner, J. Kim, P. McCaw, and P.A. Sharp. 1995. New proteins related to the Ser-Arg family of splicing factors. *RNA*. 1:852-865.
- Bohmann, K., J. Ferreira, and A.I. Lamond. 1995. Mutational analysis of p80 coilin indicates a functional interaction between coiled bodies and the nucleolus. *J. Cell Biol.* 131:817-831.
- Bregman, D.B., L. Du, S. van der Zee, and S.L. Warren. 1995. Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J. Cell Biol.* 129:287-298.
- Cáceres, J.F., and A.R. Krainer. 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO (Eur. Mol. Biol. Organ.) J.* 12:4715-4726.
- Cáceres, J.F., S. Stamm, D.M. Helfman, and A.R. Krainer. 1994. Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. *Science (Wash. DC)*. 265:1706-1709.
- Cavaloc, Y., M. Popielarz, J.-P. Fuchs, R. Gattoni, and J. Stévenin. 1994. Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:2639-2649.
- Colwill, K., T. Pawson, B. Andrews, J. Prasad, J.L. Manley, J.C. Bell, and P.I. Duncan. 1996. The Clk/Sty protein kinase phosphorylates SR protein factors and regulates their intracellular distribution. *EMBO (Eur. Mol. Biol. Organ.) J.* 15:265-275.
- Dingwall, C., and R.A. Laskey. 1991. Nuclear targeting sequences—a consensus? *Trends Biochem. Sci.* 16:478-481.
- Dingwall, C., S.M. Dilworth, S.J. Black, S.E. Kearsey, L.S. Cox, and R.A. Laskey. 1987. Nucleoplasm cDNA reveals polyglutamic acid tracts and a cluster of sequences homologous to putative nuclear localization signals. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:69-74.
- Dreyfuss, G., M.J. Matunis, S. Piñol-Roma, and C.G. Burd. 1993. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* 62:289-321.
- Du, L., and S.L. Warren. 1997. A functional interaction between the carboxy-

- terminal domain of RNA polymerase II and pre-mRNA splicing. *J. Cell Biol.* 136:5–18.
- Eperon, I.C., D.C. Ireland, R.A. Smith, A. Mayeda, and A.R. Krainer. 1993. Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO (Eur. Mol. Biol. Organ.) J.* 12:3607–3617.
- Fakan, S., and E. Puvion. 1980. The ultrastructural visualization of nucleolar and extranucleolar RNA synthesis and distribution. *Int. Rev. Cytol.* 65:255–299.
- Fu, X.-D. 1993. Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature (Lond.)*. 365:82–85.
- Fu, X.-D. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA*. 1: 663–680.
- Fu, X.-D., and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature (Lond.)*. 343: 437–441.
- Fu, X.-D., and T. Maniatis. 1992. Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science (Wash. DC)*. 256:535–538.
- Ge, H., and J.L. Manley. 1990. A protein factor, ASF, controls alternative splicing of SV40 early pre-mRNA *in vitro*. *Cell*. 62:25–34.
- Ge, H., P. Zuo, and J.L. Manley. 1991. Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. *Cell*. 66:373–382.
- Görllich, D., and I.W. Mattaj. 1996. Nucleocytoplasmic transport. *Science (Wash. DC)*. 271:1513–1518.
- Gui, J.-F., W.S. Lane, and X.-D. Fu. 1994a. A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature (Lond.)*. 369:678–682.
- Gui, J.-F., H. Tronchere, S.D. Chandler, and X.-D. Fu. 1994b. Purification and characterization of a kinase specific for the serine and arginine-rich pre-mRNA splicing factors. *Proc. Natl. Acad. Sci. USA*. 91:10824–10828.
- Hedley, M.L., H. Amrein, and T. Maniatis. 1995. An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine rich splicing factor. *Proc. Natl. Acad. Sci. USA*. 92:11524–11528.
- Horowitz, D.S., and A.R. Krainer. 1994. Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing. *Trends Genet.* 10:100–106.
- Huang, S., and D.L. Spector. 1991. Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. *Genes Dev.* 5:2288–2302.
- Huang, S., and D.L. Spector. 1992. U1 and U2 small nuclear RNAs are present in nuclear speckles. *Proc. Natl. Acad. Sci. USA*. 89:305–308.
- Huang, S., and D.L. Spector. 1996. Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. *J. Cell Biol.* 133:719–732.
- Jamison, S.F., Z. Pasman, J. Wang, C. Will, R. Lührmann, J.L. Manley, and M. Garcia-Blanco. 1995. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. *Nucleic Acids Res.* 23:3260–3267.
- Jiménez-García, L.F., and D.L. Spector. 1993. *In vivo* evidence that transcription and splicing are coordinated by a recruiting mechanism. *Cell*. 73:47–59.
- Kiseleva, E., T. Wurtz, N. Visa, and B. Daneholt. 1994. Assembly and disassembly of spliceosomes along a specific pre-messenger RNP fiber. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:6052–6061.
- Kohtz, J.D., S.F. Jamison, C.L. Will, P. Zuo, R. Lührmann, M.A. Garcia-Blanco, and J.L. Manley. 1994. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature (Lond.)*. 368:119–124.
- Krainer, A.R., G.C. Conway, and D. Kozak. 1990a. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell*. 62:35–42.
- Krainer, A.R., G.C. Conway, and D. Kozak. 1990b. Purification and characterization of SF2, a human pre-mRNA splicing factor. *Genes Dev.* 4:1158–1171.
- Krainer, A.R., A. Mayeda, D. Kozak, and G. Binns. 1991. Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. *Cell*. 66:383–394.
- Krämer, A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* 65:367–409.
- Lamond, A.I., and M. Carmo-Fonseca. 1993. The coiled body. *Trends Cell Biol.* 3:198–204.
- Li, H., and P.M. Bingham. 1991. Arginine/Serine rich domains of the su(wa) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell*. 67:335–342.
- Manley, J.L., and R. Tacke. 1996. SR proteins and splicing control. *Genes Dev.* 10:1569–1579.
- Mattaj, I.W. 1994. Splicing in space. *Nature (Lond.)*. 372:727–728.
- Mayeda, A., and A.R. Krainer. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*. 68:365–372.
- Mayeda, A., S.H. Munroe, J.F. Cáceres, and A.R. Krainer. 1994. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO (Eur. Mol. Biol. Organ.) J.* 13:5483–5495.
- McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S.D. Patterson, M. Wickens, and D.L. Bentley. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature (Lond.)*. 385:357–361.
- Michael, W.M., M. Choi, and G. Dreyfuss. 1995. A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell*. 83:415–422.
- Misteli, T., and D.L. Spector. 1996. Protein phosphorylation and nuclear organization of pre-mRNA splicing. *Trends Cell Biol.* 7:135–138.
- Mortillaro, M.J., B.J. Blencowe, X. Wei, H. Nakayasu, L. Du, S.L. Warren, P.A. Sharp, and R. Berezney. 1996. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc. Natl. Acad. Sci. USA*. 93:8253–8257.
- Nikolakaki, E., G. Simos, S.D. Georgatos, and A.T. Giannakouros. 1996. A nuclear envelope-associated kinase phosphorylates arginine-serine motifs and modulates interactions between the lamin B receptor and other nuclear proteins. *J. Biol. Chem.* 271:8365–8372.
- O'Keefe, R.T., A. Mayeda, C.L. Sadowski, A.R. Krainer, and D.L. Spector. 1994. Disruption of pre-mRNA splicing *in vivo* results in reorganization of splicing factors. *J. Cell Biol.* 124:249–260.
- Piñol-Roma, S., and G. Dreyfuss. 1993. hnRNP proteins: localization and transport between the nucleus and the cytoplasm. *Trends Cell Biol.* 3:151–155.
- Romac, J.M.-J., and J.D. Keene. 1995. Overexpression of the arginine-rich carboxy-terminal region of U1 snRNP 70K inhibits both splicing and nucleocytoplasmic transport of mRNA. *Genes Dev.* 9:1400–1410.
- Rossi, F., E. Labourier, T. Forné, G. Divita, J. Derancourt, J.F. Riou, E. Antoine, G. Cathala, C. Brunel, and J. Tazi. 1996. Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature (Lond.)*. 381:80–82.
- Scheer, U., and D. Weisenberger. 1994. The nucleolus. *Curr. Opin. Cell Biol.* 6: 354–359.
- Schmidt-Zachmann, M.S., and E.A. Nigg. 1993. Protein localization to the nucleolus: a search for targeting domains in nucleolin. *J. Cell Sci.* 105:799–806.
- Screation, G.R., J.F. Cáceres, A. Mayeda, M.V. Bell, M. Plebanski, D.G. Jackson, J.I. Bell, and A.R. Krainer. 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:4336–4349.
- Silver, P.A. 1991. How proteins enter the nucleus. *Cell*. 64:489–497.
- Siomi, H., and G. Dreyfuss. 1995. A nuclear localization domain in the hnRNP A1 protein. *J. Cell Biol.* 129:551–560.
- Spector, D.L. 1990. Higher order nuclear organization: three-dimensional distribution of small nuclear ribonucleoprotein particles. *Proc. Natl. Acad. Sci. USA*. 87:147–151.
- Spector, D.L. 1993. Macromolecular domains within the cell nucleus. *Annu. Rev. Cell Biol.* 9:265–315.
- Spector, D.L., W.H. Schrier, and H. Busch. 1983. Immunoelectron microscopic localization of snRNPs. *Biol. Cell* 49:1–10.
- Spector, D.L., X.-D. Fu, and T. Maniatis. 1991. Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3467–3481.
- Staknis, D., and R. Reed. 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol. Cell Biol.* 14:7670–7682.
- Tacke, R., and J.L. Manley. 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO (Eur. Mol. Biol. Organ.) J.* 14:3540–3551.
- Tanaka, M., and W. Herr. 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. *Cell*. 60:375–386.
- Vellard, M., A. Sureau, J. Soret, C. Martinier, and B. Perbal. 1992. A potential splicing factor is encoded by the opposite strand of the trans-spliced c-myc exon. *Proc. Natl. Acad. Sci. USA*. 89:2511–2515.
- Wang, J., and J.L. Manley. 1995. Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing *in vivo* in diverse ways. *RNA*. 1:335–346.
- Wansink, D.G., W. Schul, I. van der Kraan, B. van Steensel, R. van Driel, and L. de Jong. 1993. Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J. Cell Biol.* 122:283–293.
- Weibel, E.R., W. Stäubli, H.R. Gnägi, and F.A. Hess. 1969. Correlated morphometric and biochemical studies on the liver cell. I: Morphometric model, stereological methods, and normal morphometric data for rat liver. *J. Cell Biol.* 42:69–91.
- Weighardt, F., G. Biamonti, and S. Riva. 1995. Nucleo-cytoplasmic distribution of human hnRNP proteins: a search for the targeting domains in hnRNP A1. *J. Cell Sci.* 108:545–555.
- Wilson, A.C., M.G. Peterson, and W. Herr. 1995. The HCF repeat is an unusual proteolytic cleavage signal. *Genes Dev.* 9:2445–2458.
- Wu, J.Y., and T. Maniatis. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*. 75:1061–1070.
- Xiao, S.H., and J.L. Manley. 1997. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev.* 11:334–344.
- Xing, Y., C.V. Johnson, P.R. Dobner, and J.B. Lawrence. 1993. Higher level organization of individual gene transcription and RNA splicing. *Science (Wash. DC)*. 259:1326–1330.
- Xing, Y., C.V. Johnson, P.T. Moen, J. McNeil, and J.B. Lawrence. 1995. Non-random gene organization: structural arrangements of specific pre-mRNA transcription and splicing with SC35-domains. *J. Cell Biol.* 131:1635–1647.
- Yang, X., M.R. Bani, S.-J. Lu, S. Rowan, Y. Ben-David, and B. Chabot. 1994. The A1 and A1B proteins of heterogeneous nuclear ribonucleoproteins modulate 5' splice site selection *in vivo*. *Proc. Natl. Acad. Sci. USA*. 91:6924–6928.

- Yuryev A., M. Patturajan, Y. Litingtung, R.V. Joshi, C. Gentile, M. Gebara, and J.L. Corden. 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. USA*. 93:6975-6980.
- Zahler, A.M., and M.B. Roth. 1995. Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites. *Proc. Natl. Acad. Sci. USA*. 92:2642-2646.
- Zahler, A.M., W.S. Lane, J.A. Stolk, and M.B. Roth. 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* 6:837-847.
- Zahler, A.M., K.M. Neugebauer, W.S. Lane, and M.B. Roth. 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science (Wash. DC)*. 260:219-222.
- Zeitlin, S.A., A. Parent, S. Silverstein, and A. Efstratiadis. 1987. Pre-mRNA splicing and the nuclear matrix. *Mol. Cell. Biol.* 7:111-120.
- Zerler, B., B. Moran, K. Maruyama, J. Moomaw, T. Grodzicker, and E.H. Ruly. 1986. Adenovirus E1A coding sequences that enable ras and pmt oncogenes to transform cultured primary cells. *Mol. Cell. Biol.* 6:887-899.
- Zhang, G., K.L. Taneja, R.H. Singer, and M.R. Green. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature (Lond.)*. 372:809-812.
- Zuo, P., and T. Maniatis. 1996. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes Dev.* 10:1356-1368.
- Zuo, P., and J.L. Manley. 1993. Functional domains of the human splicing factor ASF/SF2. *EMBO (Eur. Mol. Biol. Organ.) J.* 12:4727-4737.