REGULATION OF MAMMALIAN 3' SPLICE SITE RECOGNITION

Anna Corrionero Saiz

DOCTORAL THESIS UPF / 2010

Thesis director

Dr. Juan Valcárcel Juárez

Gene Regulation Program, CRG

UNIVERSITAT POMPEU FABRA

A mi padre Ricardo, a la meva mare Rosa Maria,
a mi hermano Ricard y a mi iaia Conchita.

Ao Luís.

THE MYTH OF SISYPHUS

"The gods had condemned Sisyphus to ceaselessly rolling a rock to the top of a mountain, whence the stone would fall back of its own weight. They had thought with some reason that there is no more dreadful punishment than futile and hopeless labor.

(…) If this myth is tragic, that is because its hero is conscious. Where would his torture be, indeed, if at every step the hope of succeeding upheld him? The workman of today works everyday in his life at the same tasks, and his fate is no less absurd. But it is tragic only at the rare moments when it becomes conscious. Sisyphus, proletarian of the gods, powerless and rebellious, knows the whole extent of his wretched condition: it is what he thinks of during his descent. The lucidity that was to constitute his torture at the same time crowns his victory. There is no fate that cannot be surmounted by scorn.

If the descent is thus sometimes performed in sorrow, it can also take place in joy. This word is not too much. Again I fancy Sisyphus returning toward his rock, and the sorrow was in the beginning. When the images of earth cling too tightly to memory, when the call of happiness becomes too insistent, it happens that melancholy arises in man's heart: this is the rock's victory, this is the rock itself. The boundless grief is too heavy to bear.

(…) All Sisyphus' silent joy is contained therein. His fate belongs to him. His rock is a thing. Likewise, the absurd man, when he contemplates his torment, silences all the idols. (…)There is no sun without shadow, and it is essential to know the night. The absurd man says yes and his efforts will henceforth be unceasing. If there is a personal fate, there is no higher destiny, or at least there is, but one which he concludes is inevitable and despicable. For the rest, he knows himself to be the master of his days. At that subtle moment when man glances backward over his life, Sisyphus returning toward his rock, in that slight pivoting he contemplates that series of unrelated actions which become his fate, created by him, combined under his memory's eye and soon sealed by his death. Thus, convinced of the wholly human origin of all that is human, a blind man eager to see who knows that the night has no end, he is still on the go. The rock is still rolling.

I leave Sisyphus at the foot of the mountain! One always finds one's burden again. (…) The struggle itself toward the heights is enough to fill a man's heart. One must imagine Sisyphus happy.

---Albert Camus, 1942

ABSTRACT

Alternative splicing provides the cell the ability to generate, from a single gene, multiple protein isoforms, sometimes with different or even antagonistic functions. This process is tightly regulated and alterations in the accurate balance of alternatively spliced mRNAs are a common cause of disease.

The main objective of this thesis has been to understand the molecular mechanisms underlying disease-causing defective splicing. Skipping of Fas death receptor exon 6 leads to decreased Fas-ligand induced apoptosis. We have studied how this event is promoted by a mutation at the 3' splice site and by the proto-oncogene SF2, leading to Autoimmune Lymphoproliferative Syndrome and possibly contributing to tumor progression, respectively. Moreover, we have determined the mechanism by which an antitumor drug, Spliceostatin A, alters 3' splice site recognition and affects alternative splicing.

This thesis underscores the importance of pre-mRNA splicing in disease and how the study of disease-causing aberrant splicing can be used as a tool to understand splicing mechanisms and vice versa.

RESUM

El processament alternatiu del pre-ARNm proporciona a la cèl·lula l'habilitat de generar, a partir d'un únic gen, proteïnes amb funcions diferents i, fins i tot, antagòniques. Aquest procés està altament regulat i desequilibris en l'abundància de les diferent isoformes són causes comunes de malaltia.

L'objectiu principal d'aquesta tesi ha estat entendre el mecanisme molecular a través del qual problemes en el processament del pre-ARNm causen malalties. L'exclusió de l'exó 6 del receptor de mort cel·lular Fas condueix a una disminució de l'apoptosi en resposta al lligand de Fas. Hem estudiat com una mutació al lloc de processament 3' d'aquest exó i el proto-oncogén SF2 promouen aquest patró, causant el síndrome autoimmune lifoproliferatiu i possiblement contribuint a la progressió tumorogènica, respectivament. A més, hem estudiat el mecanisme pel qual la droga antitumoral Spliceostatin A altera el reconeixement del lloc de processament 3' i causa canvis en el processament alternatiu de diversos gens.

Aquesta tesi posa en evidència la importancia del processament del pre-ARNm en malalties i com l'estudi de mutacions que alteren aquest procés i són causa de malaties pot ser utilitzat con una eina per entendre el mecanisme d'aquest processament i viceversa.

PREFACE

The central dogma of molecular biology, which postulated that DNA was transcribed into RNA and RNA, in turn, translated into protein in a unidirectional flow of information, was announced for the first time more than 50 years ago by Francis Crick (Crick, 1958). He himself re-stated it in 1970 by saying that *"the central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that information cannot be transferred back from protein to either protein or nucleic acid."* (Crick, 1970). But the central dogma, although for many years a keystone in molecular biology, has turned out to be an oversimplified view of what Nature is able to do.

The existence of RNA-dependent RNA- and DNA-polymerases, microRNAs and small interfering RNAs or the ability of proteins to alter DNA composition and function (e. g. DNA methyltransferases) are some of the many exceptions that can be found to the central dogma. Possibly, though, the most underestimated biopolymer (DNA, RNA and protein) in this theory was the RNA. Unimaginably in the 1950's, a hypothetical primordial RNA world, where RNA provided both the genetic information and the function, was suggested by Walter Gilbert in 1986 after the discovery of catalytically competent RNAs, or ribozymes, by Sidney Altman (Guerrier-Takada et al., 1983) and Thomas R. Cech (Kruger et al., 1982). In fact, non-protein coding RNAs play essential roles in many processes in the cell: RNA molecules are core components of ribosomes (rRNAs) and spliceosomes (snRNAs), small nucleolar RNAs (snoRNAs) direct pseudouridylation and ribose methylation of other RNAs; microRNAs (miRNAs) and small interference RNAs (siRNAs) repress translation or mRNA degradation by specifically binding to mRNAs; small double stranded (shRNAs) can induce DNA and histone methylation and promote transcription silencing; and long non-coding RNAs (ncRNAs) can function as a scaffold for the assembly of different cellular structures (reviewed in Mendes Soares and Valcarcel, 2006).

But the central dogma was not only ignorant of the presence of regulatory and catalytically active RNAs but also of the fact that protein coding RNAs have been processed and highly modified from the originally transcribed premature messenger RNA (pre-mRNA).

Pre-mRNAs are co- and post-transcriptionally processed before being exported to the cytoplasm and translated into proteins. Shortly after the initiation of transcription, a guanosine residue is linked through a 5'-5'-triphosphate bond to the first transcribed nucleotide by the Capping Enzyme Complex. Then, a methyltransferase present in the complex methylates the nitrogen 7 of the guanine, generating a 7-methylguanosine cap (reviewed in Shuman, 2001). The 3' end of the pre-mRNA is bound by the polyadenylation machinery, which will cleave the pre-mRNA and add a polyA tail. Although these modifications do not change the coding sequences, they affect many aspects of the RNA metabolism by enhancing mRNA export, preventing mRNA degradation by exonucleases and promoting translation, thus acting as critical check point in preventing aberrant gene expression (reviewed in Moore and Proudfoot, 2009).

A third nuclear modification of the pre-mRNA consists in the removal of non-coding, intervening sequences (introns) in order to join together the coding sequences (exons) by a process called splicing. Splicing was firstly observed in the adenovirus *hexon* gene. During cell infection, several labs identified RNA molecules, originally named mosaics, which contained noncontiguous sites of the viral genome (Berget et al., 1977; Chow et al., 1977). Later on, splicing of many other viral and eukaryotic genes was reported (Darnell, 1978; Knapp et al., 1978) suggesting that this process was an essential step of the maturation of mRNAs. Unlike pre-mRNA capping and polyadenylation, splicing can affect the coding capacity of the mRNA. It also influences many, if not all, steps of mRNA metabolism.

A typical human gene contains between 8 and 9 introns, whose average size is more than 3.4 Kb (Deutsch and Long, 1999), while exon size is on average about 140 nucleotides (Fedorova and Fedorov, 2003). Strikingly, which sequences are included or excluded from the final transcript can be regulated depending on the developmental stage, the cell type or the extracellular signals by the process of alternative splicing, potentially expanding the proteome diversity.

KEY WORDS

Acceptor site

Base pairing

Branch point

Exon

Inclusion

Intron

mRNA

Polypyrimidine tract

Pre-mRNA

RNA recognition motif

SF2

SF3b

Skipping

Spliceosome

Spliceostatin A

Splicing

SR proteins

U2 auxiliary factor

U2 snRNP

3' splice site

5' splice site

ABBREVIATIONS

A: adenosine

AS: alternative splicing

ATP: adenosine triphosphate

bp: base pairs

BP: branch point

bprs: branch point recognition site

C: cytidine

cDNA: mRNA complementary DNA

CMV: cytomegalovirus

C-terminal: carboxi-terminal

CTD: carboxi-terminal domain

CTP: cytosine triphosphate

DMEM: Dulbecco's modified Eagle's medium

DNA: deoxyribonucleic acid

DTT: dithiothreitol

EDTA: ethylenediamine tetraacetic acid

EMSA: electrophoretic mobility shift assay

ESE: exonic splicing enhancer

ESS: exonic splicing silencer

G: guanosine

GTP: guanosine triphosphate

hnRNP: heterogeneous ribonucleoprotein

IP: immunoprecipitation

ISE: intronic splicing enhancer

ISS: intronic splicing silencer

Kb: kilo base pairs

KDa: kilodalton

mRNA: messenger RNA

NE: nuclear extract

nt: nucleotide

N-terminal: amino-terminal

PAGE: polyacrylamide gel electrophoresis

PBS: phophate buffered saline

PCR: polymerase chain reaction

Pol II: RNA polymerase II

pre-mRNA: premature mRNA

PTB: polypyrimidine-tract binding protein

pY tract: polypyrimidine tract

RNA: ribonucleic acid

RT-PCR: reverse transcription-polymerase chain reaction

siRNA: small interfering RNA

SDS: sodium dodecyl sulfate

SF1: splicing factor 1

SF2/ASF: splicing factor 2/alternative splicing factor

snRNP: small nuclear ribonucleoprotein particle

SR: serine/arginine-rich

SRPK: SR protein kinase

ss: splice site

SSA: Spliceostatin A

T: thymidine

T7: T7 bacteriophage

TBE: Tris/Borate/EDTA

tRNA: transfer RNA

U: uridine

U snRNA: uridine-rich small nuclear RNA

UTP: uridine triphosphate

WB: western blot

WT: wild type

TABLE OF CONTENTS

GENERAL INTRODUCTION

The splicing reaction

Defining exon-intron boundaries

Sequences that delimit introns and exons and how these coding sequences are spliced together have been extensively studied, however they are still not well understood.

In higher eukaryotes, introns are delimited by short consensus sequences required for both intron recognition and splicing catalysis (Fig. 1). Despite of its importance, these sequences are extremely variable, making their recognition by the splicing machinery difficult. In contrast, yeast introns are delimited by similar but much more conserved sequences (reviewed in Green, 1986).

The 5' splice site marks the 5' end of the intron and is characterized by an invariable GU dinucleotide flanked by a less conserved consensus sequence: MAG/GURAGU (M = A or C, / = exon-intron junction, R = A or G) (Horowitz and Krainer, 1994). At the other end of the intron, three are important sequences for splicing: the branch point is followed by a polypyrimidine (pY) tract located upstream of the AG dinucleotide that marks the 3' end of the intron. The former is characterized by the presence of a conserved adenosine surrounded by a highly degenerated sequence YNYURAY (Y = U or C and R = G or A) (Zhang, 1998). The pY tract is of variable length and the pyrimidine content is often correlated with the efficiency of 3'ss recognition during the initial steps of the spliceosome assembly pathway (see below) (Ruskin and Green, 1985). Finally, the AG dinucleotide or acceptor site is usually surrounded by a pyrimidine upstream and a G as a first nucleotide of the exon. The acceptor site is usually located 40-18 nucleotides downstream of the branch point (Ruskin et al., 1984; Reed and Maniatis, 1985), although some exceptions can be found where the branch point is located hundreds of nucleotides upstream of the AG dinucleotides (Helfman and Ricci, 1989; Smith and Nadal-Ginard, 1989). The sequences between the branch point and the actual acceptor site are characterized by the absence of AG dinucleotides (Gooding et al., 2006).
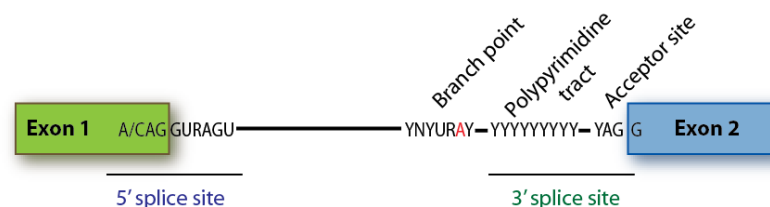
Fig. 1. Introns are flanked by conserved sequences at their 5' and 3' ends. The 5' end is characterized by a GU dinucleotide surrounded by a relatively conserved consensus sequence. The 3' end contains three important sequences: the branch point, the polypyrimidine tract and the AG acceptor site. Exons are represented by boxes and introns by black lines. Y, pyrimidine; R, purine; N, any nucleotide.

Splicing biochemistry

The excision of the introns from the pre-mRNAs occurs through two transesterification reactions carried out by the spliceosome (Fig. 2). In the first step, the 2' hydroxyl group of the conserved adenosine at the branch point carries out a nucleophilic attack of the 3'-5' phosphodiester bond between the 5' splice site and the exon. This reaction results in a 2'-5' phosphodiester bond between the branching adenosine and the 5' terminal phosphate of the intron, giving rise to a lariat-shaped intermediate, which contains the intron and the downstream exon, and the free upstream exon (Ruskin and Green, 1985; Konarska et al., 1985). During the second transesterification reaction, the 3' hydroxyl group of the free exon attacks the phosphodiester bond at the 3' splice site of the lariat intermediate, joining together the two exons and releasing the lariat intron (Konarska et al., 1985; Moore and Sharp, 1993).
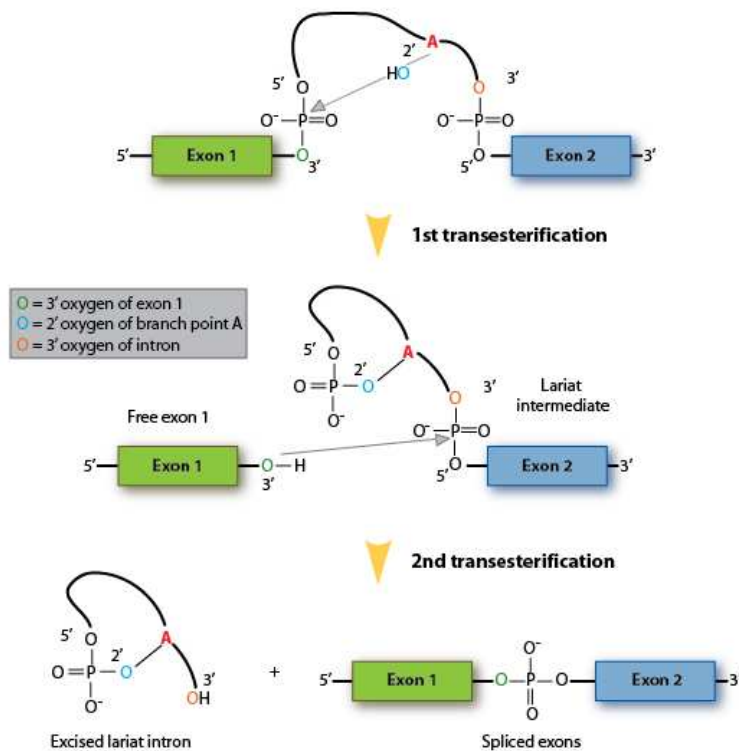
Fig. 2. Splicing takes place in two transesterification reactions. The first catalytic reaction consists of a nucleophilic attack by the 2'-OH (in blue) of the branch point adenosine on the 5' splice site phosphate. Subsequently, the 3'OH of the free 5' exon (in green) attacks the 3' splice site phosphate, splicing together both exons and releasing the intron in a lariat structure (adapted from [Griffiths, 1999]).

This reaction is chemically identical to the one carried out by the class II of self-splicing introns (reviewed in Michel and Ferat, 1995; Sharp, 1987). As pre-mRNA splicing, group II introns are spliced out via two transesterification reactions, the first one consisting of a nucleophilic attack of the 2'OH of an adenosine to the 5' end of the intron generating a lariat structure, followed by a second nucleophilic attack on the 3'ss by the 3' OH of the upstream

exon. Although protein components are required for pre-mRNA splicing, the intron itself provides the reactive groups in the splicing reaction as well as nucleates a network of RNA-RNA interactions with small nuclear RNAs that is believed to form the catalytic core of the spliceosome (reviewed in Valadkhan, 2005).

The Spliceosome, the splicing machinery

Spliceosome components: snRNPs biogenesis

While the biochemical steps of the splicing reaction are well known, how splice sites are unequivocally recognized and which elements carry out both transesterification reactions are not completely understood.

The splicing machinery, or spliceosome, recognizes the consensus sequences present at the exon-intron boundaries. It is composed of five snRNPs (U1, U2, U4, U5 and U6), each containing a uridine-rich snRNAs with a complex secondary structure, a common set of Sm proteins and a unique set of snRNP-specific proteins (reviewed in Wahl et al., 2009). U1, U2, U4 and U5 snRNAs are transcribed by RNA pol II and are characterized by the presence of a 5'-trimethylguanosine (TMG) cap, an Sm-protein-binding site (Sm site) and a 3' stem-loop structure. On the other hand, U6 snRNA is transcribed by RNA pol III and contains a 5'-monomethylphosphate cap and a 3' stem-loop followed by a stretch of uridine residues (Lsm site) that is bound by the Lsm core (reviewed in Matera et al., 2007).

U snRNAs must undergo a complex maturation process before being functional in the spliceosome. First, Sm-class snRNAs (U1, U2, U4 and U5) are capped and exported to the cytoplasm where the SMN complex assembles the Sm core (composed by seven common Sm proteins B/B', D1, D2, D3, E, F, andG) on the Sm site (Raker et al., 1996]. Second, the m7G cap is hypermethylated to form the TMG cap structure (Mouaikel et al., 2002)and the 3'end is trimmed by a 3'-to-5' exonuclease (Huang et al., 1999). The presence of the TMG cap and the SMN complex promotes the assembly of an import complex that facilitates re-entering of the particle in the nucleus and targeting to Cajal bodies, where snRNA site-specific modifications, such as pseudouridylation and methylation, take place (reviewed in Matera et al., 2007; Will and Luhrmann, 2001). The association of snRNP-specific proteins is thought to occur in the nucleus, consistent with the snRNA-independent nuclear import of U2B" and U1A (Hetzer and Mattaj, 2000).

U2 snRNP assembly has been particularly well studied. This snRNA is initially bound by the Sm core and two U2-specific proteins, A' and B", giving rise to a 12S snRNP (Will and Luhrmann, 1997). Later on, the SF3b subcomplex, consisting of 8 subunits (SF3b155, SF3b145, SF3b130, SF3b125, SF3b49, SF3b14b, SF3b10 and p14) (Behrens et al., 1993; Das et al., 1999; Kramer et al., 1999; Will et al., 2002) associates to form the 15S U2 snRNP which will be ultimately bound by the SF3a subcomplex, composed by SF3a60, SF3a66 and

SF3a120 (Brosi et al., 1993b), giving rise to the mature, splicing competent 17S snRNP. Moreover, many other non-snRNP proteins are associated with U2 snRNP (Will et al., 2002). Where in the nucleus, whether in the Cajal bodies or in the nuclear speckles, all these proteins associate to the snRNA is still largely controversial (Nesic et al., 2004).

The splicing cycle

Although a spliceosome composed by the five snRNP could be purified under low-salt conditions in yeast extracts (Stevens et al., 2002), data coming from classical in vitro assays (e. g. gel electrophoresis or gel filtration) (Konarska and Sharp, 1986; Pikielny et al., 1986), or more recent chromatin immunoprecipitation against splicing factors experiments (Kotovic et al., 2003; Lacadie and Rosbash, 2005; Listerman et al., 2006) strongly suggest that the spliceosome assembles onto the RNA in a stepwise manner. The spliceosome cycle is highly conserved along the eukaryotic evolutionary tree, reason why human and the yeast *Saccharomyces cerevisiae* have been the main organism models used to study how the spliceosome assembles on the pre-mRNA.

The 5'ss is initially recognized by U1 snRNP. This interaction relies on both protein-RNA and RNA-RNA interactions, the latter consisting of canonical base pairing of the 5' end of U1 snRNA with the last three nucleotides of the exon and the first six of the intron (reviewed in Madhani and Guthrie, 1994). However, alternative patterns of base pairing have been described (Roca and Krainer, 2009).The U1 snRNP-specific protein U1C interacts with the RNA stabilizing the U1-5'ss base pairing (Du and Rosbash, 2002; Heinrichs et al., 1990; Pomeranz Krummel et al., 2009).

At the other end of the intron, the branch point (BP) is recognized by the branch point binding protein (BBP)/SF1 through its maxi KH domain (Kramer and Utans, 1991; Berglund et al., 1997, Liu et al., 2001b). The pY tract and the AG acceptor site are bound by the 65 KDa (Ruskin et al., 1988; Zamore et al., 1992) and the 35 KDa (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999) subunits of the U2 snRNP auxiliary factor (U2AF), respectively. U2AF65 interacts through its C-terminal UHM domain with SF1, significantly improving the affinity of SF1/BBP for the branch point (Berglund et al., 1998; Selenko et al., 2003) (Fig. 3A).

The interaction of both U2AF heterodimer and SF1 at the 3'ss and U1 snRNP at the 5'ss, together with possible additional interactions across the intron, constitute the ATP-independent E complex (Michaud and Reed, 1991). A precursor complex, E', which only contains U1 snRNP and SF1 was identified in U2AF65-depleted extracts (Kent et al., 2005). At this early steps of the spliceosome assembly pathway, both 5' and 3'ss are already found in close proximity of each other (Kent et al., 2005) and the pre-mRNA is committed to the splicing pathway (Seraphin and Rosbash, 1989; Michaud and Reed, 1991).

U2 snRNP is already recruited, although unstably, in E complex (Hong et al., 1997), possibly through protein-protein interactions involving the UHM domain of U2AF65 and the unstructured N-terminal domain of SF3b155 (Gozani et al., 1998; Spadaccini et al., 2006; Thickman et al., 2006). Upon ATP hydrolysis, U2 snRNP stably associates with the branch point, displacing SF1 and leading to the formation of the pre-spliceosome or A complex. U2 snRNA base pairs with this sequence through the branch point recognition sequence corresponding to the hexanucleotide +33-GUAGUA- +38 (Nelson and Green; Wu and Manley, 1989; Zhuang and Weiner, 1989), where the U at position +34 is pseudouridilated, modification that has been proposed to induce a unique structure in the U2 snRNA-branch point helix (Newby and Greenbaum, 2002) (Fig. 4B). The most characteristic feature of this duplex is the presence of a single unpaired residue, the branch adenosine, that bulges out of the helix facilitating the nucleophilic attack on the 5'ss (Query et al., 1994).
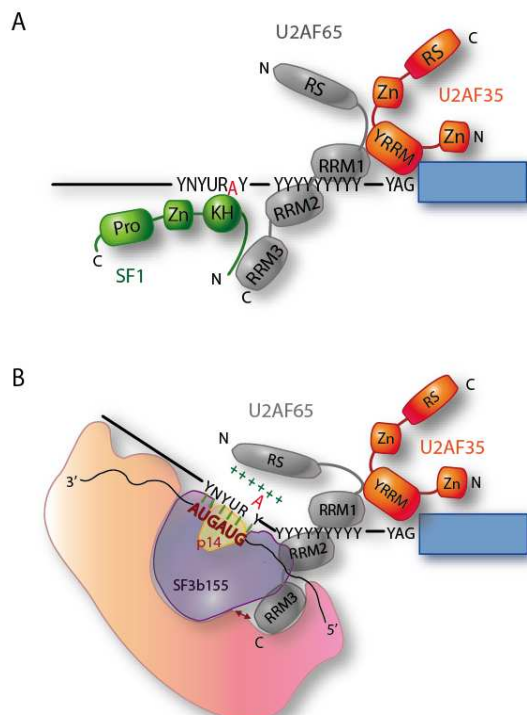
Fig. 3. 3' splice site recognition. (A) 3'ss is recognized initially in an ATP-independent manner by SF1, that will bind the branch point, and U2AF35 and U2AF65 that interact with the AG acceptor site and the polypyrimidine tract respectively. **(B)** Upon ATP hydrolysis, U2 snRNP is stably recruited to the branch point. U2AF65 promotes this binding through RNA-U2AF65 RS domain and SF3b155-U2AF65 RRM3 interactions.

U2 snRNP protein components also associate with the branch point and its neighboring sequences, stabilizing the RNA duplex. p14 has been described to interact directly to the branch point adenosine (Query et al., 1997; Will et al., 2001) while other components of the SF3b and SF3a complexes can be crosslinked either upstream (anchoring site) or downstream of the branch point in a sequence independent manner (Gozani et al., 1996). Interestingly, SF3b155, the most highly conserved SF3b subunit (Wang et al., 1998a) which tightly interacts with p14 (Will et al., 2001), binds at both sites of this sequence and it has been proposed to tether U2 snRNP to near the branch point (Gozani et al., 1998). U2 snRNP-

branch point intraction is also stabilized by the RS domain of U2AF65, which contacts the branch point (Valcarcel et al., 1996) (Fig. 3A).

Subsequently, the preformed U4/U6.U5 tri-snRNP is recruited to the spliceosome to form the catalytically inactive B complex (Konarska and Sharp, 1987). Upon important structural and compositional rearrangements (Bessonov et al., 2008), U4 snRNP, which is base paired with U6 snRNA through two helices (Fig. 4B), leaves the spliceosome allowing U6 snRNA to base pair with the 5'ss through the conserved ACAGAGA box, displacing U1 snRNP (Lamond et al., 1988; Wassarman and Steitz, 1992). U6 snRNA also forms a new intramolecular stem-loop involved in metal binding (Fortner et al., 1994; Huppler et al., 2002) and extensive base pairing interactions with U2 snRNA (Sun and Manley, 1995) (Fig. 4C).
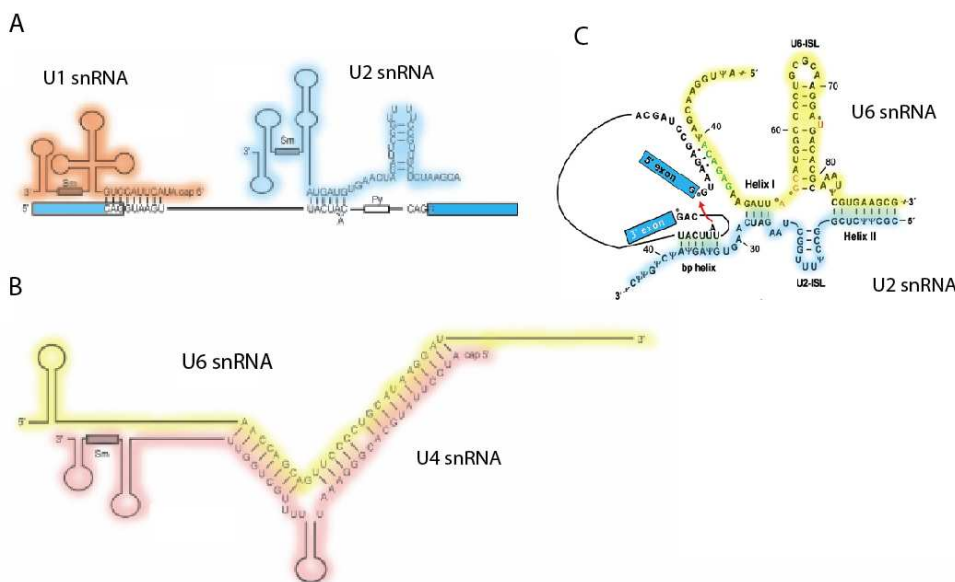
Fig. 4. RNA-RNA interactions in the splicing cycle. (A) RNA-RNA interactions during A complex. U1 snRNA (red) and U2 (blue) snRNA base pair with the 5'ss and branch point, respectively. **(B)** U6 (yellow) and U4 (pink) snRNAs are found in the cell as a di-snRNP, due to extended base pairing interactions. **(C)** U2 (blue) and U6 (yellow) snRNAs perform complex RNA-RNA base pairings among them and with the pre-mRNA in the activated spliceosome (modified from Rhode et al., 2006).

Concomitantly, Prp19p associated complex (NineTeen Complex, NTC, in yeast; CDC5 complex in metazoans) is recruited. NTC is required for stable association of U5 and U6 snRNP with the spliceosome and is involved in several conformational changes during spliceosome activation (reviewed in Hogg et al., 2010). Subsequently, U5 snRNA contacts nucleotides at both the 5' and 3' ss. These rearrangements bring together the 5'ss and the branch point and generate the catalytically active B* complex, characterized by SF3b155 hyperphosphorylation (Wang et al., 1998a) and release of SF3 from the intron branch point (Lardelli et al., 2010), which carries out the first transesterification reaction. After additional conformational and compositional changes, the C complex is formed and the second catalytic

step takes place, concomitant with SF3b155 dephosphorylation (Shi et al., 2006). Finally, the spliceosome is dissociated and recycled for other rounds of splicing, the mRNA is released as an mRNP particle and the lariat intron is debranched and degraded by exonucleases (Fig. 5) (reviewed in Wahl et al., 2009).
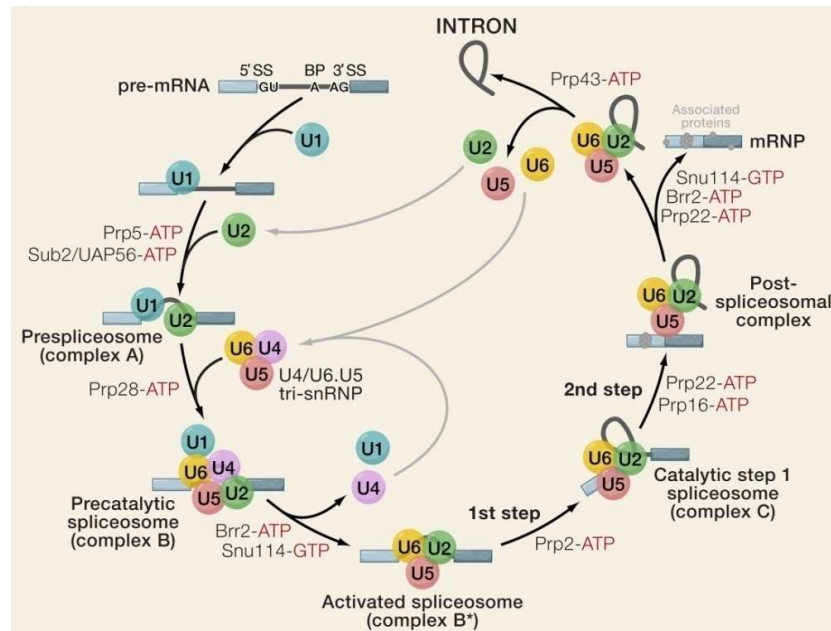
Fig. 5. The splicing cycle. The spliceosome assembles in a stepwise manner, undergoing extensive compositional and conformational rearrangements (extracted from Wahl et al., 2009).

Although the splicing machinery disassembles from the spliced mRNA before it is exported to the cytoplasm, the spliceosome leaves a permanent trace of its activity that will only be removed after the first, pioneer round of translation. This vestige of the spliceosome activity consists in a multiprotein exon junction complex (EJC) deposited about 24 nucleotides upstream of the exon junction (Le Hir et al., 2000). The EJC promotes mRNA export and translation and is essential for the surveillance system of non-sense mediated decay (NMD), which promotes the degradation of premature stop codon-containing mRNAs after the pioneer round of translation (reviewed in Lejeune and Maquat, 2005). Therefore, the splicing process not only modifies the coding sequence of an mRNA but also the composition of the RNP, which determines the fate of the mRNA, including its stability, translation efficiency and localization (reviewed in Glisovic et al., 2008).

Splicing fidelity

Functional sequences in the pre-mRNA are highly degenerated and poorly conserved. Then, how does the spliceosome recognize bona fide signals among many similar others present in the pre-mRNA and, thus, ensures the fidelity of the splicing reaction?

One simple mechanism is the recognition of the exon-intron boundaries multiple times during the splicing cycle through a network of RNA-RNA and RNA-protein interactions. A decreased affinity of the different factors that recognize functional sequences present in the pre-mRNA would render the splicing process less efficient. Also, competition between splicing factors and non-specific RNA binding proteins to interact with the pre-mRNA might be a simple way to ensure splicing fidelity and prevent the use of cryptic splice sites.

However, the most prominent mechanism to ensure the fidelity of the reaction is established by the activity of numerous RNA-dependent DExH/D-box ATPases/helicases along the splicing cycle (reviewed in Wahl et al., 2009). This family of proteins is present all along the spliceosome assembly pathway and the splicing reaction, coupling ATP hydrolysis with conformational rearrangements of the spliceosome. Much of the knowledge on the contribution of this family of ATPases comes from elegant genetic studies in the yeast *S. cerevisiae*. For example, in this organism it has been shown that Prp5 proofreads the U2 snRNA-branch point base pairing (Xu and Query, 2007), Prp16 facilitates the transition of the spliceosome from the first-step to the second-step conformations (Villa and Guthrie, 2005) and Prp22 regulates fidelity of the exon ligation by promoting release of the products (Mayas et al., 2006). The basis for the control of fidelity by these proteins consists of kinetic competition between ATP hydrolysis, and hence remodeling of RNA-RNA, protein-protein and/or RNA-protein interactions, and substrate rejection. Reduced ATPase activity or absence of ATP would extend the time between the different conformational changes allowing suboptimal substrates to be recognized and spliced (reviewed in Konarska et al., 2006).

Other factors with no such ATPase activity, such as the oncogene DEK, have been described to proofread early U2AF recognition of 3'ss with inadequate acceptor sites (Soares et al., 2006).

Different splicing machineries and splicing evolution

Class II introns: is the spliceosome a rybozyme?

Among 150-300 snRNP and non-snRNP proteins take part at some stage during the spliceosome cycle (Zhou et al., 2002; Rappsilber et al., 2002). More recently, the comparative proteomic analysis of purified spliceosomal complexes stalled at different stages has made even more evident that intron recognition and excision is an extremely dynamic process with a continuous exchange of proteins (reviewed in Wahl et al., 2009). Many of these proteins have known active roles in splicing and some others function to couple the splicing machinery to other RNA processing events such as transcription or 3' end processing (reviewed in Jurica and Moore, 2003; Maniatis and Reed, 2002). The RNA components of the spliceosome, however, have been strongly suggested to drive both steps of splicing catalysis. Both U2 and U6 snRNAs are absolutely needed for the splicing reaction, although it is very likely that

catalysis involves critical residues in U6 snRNA, while U2 snRNA is required for proper positioning of the branch point adenosine during the first transesterification step (reviewed in Valadkhan, 2007).

Currently, two pieces of evidence support this hypothesis. First, protein-free U6 and U2 snRNAs are able to sequence-specifically bind to an RNA substrate and perform splicing-related catalytic reactions in vitro (Valadkhan and Manley, 2001; Valadkhan and Manley, 2003). Second, there are striking sequence and structural similarities between the class II of self-splicing introns and RNA-RNA networks formed by U2, U6 and U5 snRNAs with the pre-mRNA, which have led to the hypothesis that pre-mRNA splicing has evolved from this class of introns (reviewed in Valadkhan, 2005). Sequence analogies include a conserved ACAGAG box in U6 snRNA and an AGC triad placed in the catalytic core, shown to be crucial for splicing catalysis by mutational analysis. Structurally, an intramolecular stem-loop in U6 snRNA clearly resembles the catalytically essential domain V of group II introns. Then, it is conceivable that fragmentation of the ancestor of type II self splicing introns evolved to U2 introns and trans-acting snRNAs (reviewed in Patel and Steitz, 2003).

However, a role for proteins in splicing catalysis has not been ruled out. Prp8p, a large, highly conserved component of the U5 snRNP has been proposed to have an active role in splicing catalysis(reviewed in Grainger and Beggs, 2005). Prp8p contacts the 5'ss, the branch point and the 3'ss and U5 and U6 snRNAs along the splicing catalysis occupying a central position in the catalytic core of the spliceosome (Dix et al., 1998; Teigelkamp et al., 1995; Umen and Guthrie, 1995; Vidal et al., 1999). Moreover, it is essential for U2 and U12-type splicing and trans-splicing reactions and is the most highly conserved component of the spliceosome, underscoring its importance for splicing catalysis (reviewed in Grainger and Beggs, 2005).

U12-class introns: the minor spliceosome

The discovery of a rare class of introns, U12-type, -representing less than 1% of all human introns- more than fifteen years ago (Hall and Padgett, 1994) added another piece in the puzzle of understanding splicing and splicing evolution. Contrarily to the canonical U2 type introns, this class of introns is characterized by relatively long, highly conserved sequences at both the 5' ss and branch point. The latter is relatively close to the acceptor site (10-16 nucleotides), without a pY tract (reviewed in Patel and Steitz, 2003; Sharp and Burge, 1997; Will and Luhrmann, 2005). Both the donor and acceptor sites were initially thought to be AT and AC dinucleotides, respectively, although later reports showed that other dinucleotides, especially GU and AG as in the U2-type introns, are used (Dietrich et al., 1997). This class of introns is spliced out by a spliceosome composed of U11, U12, U4atac, U6atac, functionally analogous to U1, U2 U4 and U6 snRNP, respectively, and the common U5 snRNP (Tarn and Guthrie, 1996a and b). None of the U12 spliceosome-specific snRNAs show high sequence

similarity to its functionally homologous snRNAs of the major spliceosome, although its secondary structure and base pairing interactions seem to be similar between both pathways. U11 and U12 snRNPs are found in the cell as an 18S di-snRNP. Although affinity purified U11 does not contain the U1 snRNP specific proteins U1A, U1C and U1 70K, other proteins associated seem to have structural similarities and, probably, functions. On the other hand, SF3b, but not Sf3a, complex can be copurified with U11/U12 di-snRNP, and the U4atac/U6atac.U5 tri-snRNP associated proteins are shared with the major U4/U6.U5 tri-snRNP. The absence of U1 snRNP-specific proteins and a pY tract, as well as and the observation of U11 and U12 snRNPs pre-assembly, suggest that initial recognition of U12 introns differs greatly from the early steps of assembly of the canonical spliceosome (reviewed in Patel and Steitz, 2003; Sharp and Burge, 1997; Will and Luhrmann, 2005).

U12 introns have been found in both plants and animals suggesting that both spliceosomes have coexisted in the eukaryotic lineage for at least one billion years (Burge et al., 1998). The fact that U2 consensus sequences are much less constrained than those of U12 points to progressive conversion of introns from U12-type to U2-type, therefore explaining why the vast majority of intervening sequences are U2-dependent. U12-class introns and the minor spliceosome, although absent in some eukaryotic lineages such as yeasts and nematodes(Burge et al., 1998; Mewes et al., 1997), might have been retained during evolution to excise few remaining introns present in genes with essential cellular functions. Interestingly, chimeric spliceosomes composed of U2 and U12 components have not been detected, although contacts across the exon between both spliceosomes have been reported (reviewed in Patel and Steitz, 2003; Sharp and Burge, 1997; Will and Luhrmann, 2005).

There are three proposed hypothesis for the origin of the coexistence of both U2 and U12 spliceosomes. The fission-fusion model suggests the existence of a common ancestor, and then both spliceosomes diverged in two different lineages that merged later on. A second model suggests a divergence of both classes in the same lineage from a unique ancestor. Finally, a parasitic invasion of a eukaryotic progenitor would introduce a new class of introns (Burge et al., 1998; Lynch and Richardson, 2002).

Trans-splicing: intermolecular splicing reaction

While canonical U2- and U12-introns splicing occurs within the same pre-mRNA molecule and non-continuous transcript (i. e. cleaved by intronic ribozymes) splicing is still controversial (Dye et al., 2006; Fong et al., 2009), some splicing-like reactions can occur between two different transcripts. More than 20 years ago it was observed that multiple mRNAs of trypanosomatids (Murphy et al., 1986; Sutton and Boothroyd, 1986) and nematodes (Krause and Hirsh, 1987) contained an identical short leader sequence, the splice leader, at their 5' ends. This sequence, which contains a consensus 5'ss, is present in the

form of an snRNP, with an Sm core, a TMG cap and a complex secondary structure. Mechanistically, trans-splicing is very similar to cis-splicing, through two transesterification reactions with the participation of U2, U4, U5 and U6 snRNPs and the presence of 3'ss, contained in the trans-spliced mRNA (reviewed in Blumenthal, 2005). Notably, U1 snRNP is dispensable for this process, although how the 5'ss is initially recognized is still unclear (Bruzik et al., 1988; Lasda et al., 2010). While trans- and cis-splicing coexist in nematodes, only trans-splicing is present in trypanosomids. Its main functions are thought to be the regulation of translation and the conversion of polycistronic to monocystronic mRNAs (Kuersten et al., 1997; Maroney et al., 1995).

Trans-splicing-like reactions, in which two different mRNA molecules are joined together, have been also identified in *Drosophila* (Mongelard et al., 2002), plants (Zhang et al., 2010) and mammalian cells (Akiva et al., 2006), and in this latter case, it has been proposed to be relevant in cancer (Li et al., 2008).

RNA binding domains: RRMs

RNA can be recognized by a variety of different motifs such as RNA recognition motifs (RRMs), hnRNP K homology (KH) domains, zinc fingers and serine/arginine-rich (RS) domains.

RRM is one of the most abundant protein domains. It is present in approximately 2% of human proteins with a wide range of cellular functions, some containing more than one RRM copy (Bateman, 2002). RRMs are composed by approximately 90 aminoacids that can accommodate four unpaired bases (reviewed in Maris et al., 2005). RRMs are characterized by the presence of two RNPs of 8 and 6 mainly aromatic and positively charged aminoacids (Adam et al., 1986; Swanson et al., 1987). Both RNPs are necessary and sufficient for binding RNA, although they can show very different affinities and specificities (Allers and Shamoo, 2001). The canonical RRM fold consists in an αβ sandwich, with a four-stranded antiparallel β-sheet (β4β1β3β2) and two α-helices packed against the β-sheet. The RNPs are located in the central strands of the β-sheet (Nagai et al., 1990). On the other hand, the loops between the secondary structures are of variable length and are usually disordered. RRM binding affinity varies depending on the contribution of both the N- and C-terminal regions of the RRM, the location of the cognate sequence in loops generated by RNA secondary structures or the combination of two or more RRMs (reviewed in Maris et al., 2005). RRMs are not only involved in RNA binding but also they can participate in protein-protein interactions by different means. RRMs can interact with other RRMs, mainly through their α-helices, enlarging the RNA-binding interface (Oberstrass et al., 2005), or with other protein domains which can act as cofactors to allow RRMs to bind to RNAs (Rideau et al., 2006). A particular case of protein interacting RRM is the U2AF-homology motif (UHM), a non-canonical RRM

with poorly conserved RNPs and an extended α1 helix, which mainly mediates the protein-protein interactions (reviewed in Kielkopf et al., 2004). For example, the UHM domain of the splicing factor SPF45 interacts with SF3b155, SF1 and U2AF65, although with different affinities, and is required for SPF45-mediated Fas exon 6 alternative splicing regulation (Corsini et al., 2007). All these features make RRMs a very diverse protein domain in terms of both function and structure (reviewed in Maris et al., 2005; Clery et al., 2008).
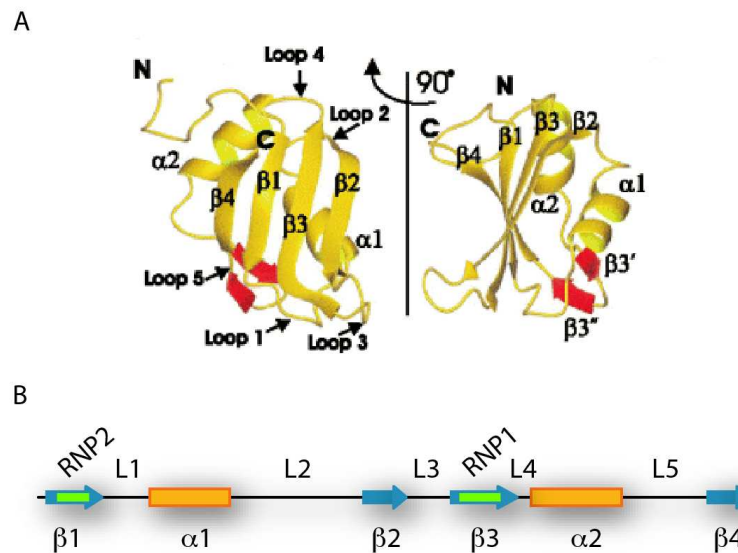
Fig. 6. Secondary estructure of RRMs. (A) hnRNP A1 RRM2. The RRM fold consist on an αβ sandwich, with two α-helices packed against a β-sheet composed by four-stranded antiparallel β-sheet (β4β1β3β2) (extracted from Maris et al., 2005; Xu et al., 1997). **(B)** Schematic representation of an RRM. The α-helices are represented by yellow boxes, the β-strands by blue arrows and the RNP1 and 2 by green boxes.

Regulation of Alternative Splicing

Alternative splicing complexity

What could be the evolutionary reason to maintain high degeneracy in splice site sequences and to build recognition upon multiple, individually weak interactions? One possible (and plausible) reason would be to allow regulation. The development of high throughput technologies, especially RNA-seq, has a provided a detailed view of the transcriptome of different cell types, leading to the surprising observation that the majority of human multiexon genes are alternatively spliced (Wang et al., 2008b; Wilhelm et al., 2008). This fact highlights the well-known relevance of alternative splicing in many cellular processes such as sex determination, cell differentiation, cell transformation or apoptosis (reviewed in Black, 2003; Nilsen and Graveley, 2010; Smith and Valcarcel, 2000; Stamm et al., 2005).

The alternative inclusion or exclusion of different sequences from a specific pre-mRNA represents a versatile mechanism to regulate gene expression and protein function, expanding enormously the cell proteome (Blencowe and Khanna, 2007). Alternative splicing occurs in a variety of modes, for example, alternative 3' and 5'ss use, whole exon or intron skipping or inclusion, alternative promoters or cleavage and polyadenylation sites. Also, contiguous exons can be mutually exclusive, where just one or the other exon is included in the final transcript (Fig. 7) (Black, 2003). Alternative inclusion or skipping of cassette exons are the most common events in U2-type introns, while alternative 5' and 3' ss choices are the most observed alternative splicing events in U12-introns, due to the incompatibility of chimeric U12-U2 spliceosomes (reviewed in Chang et al., 2008).
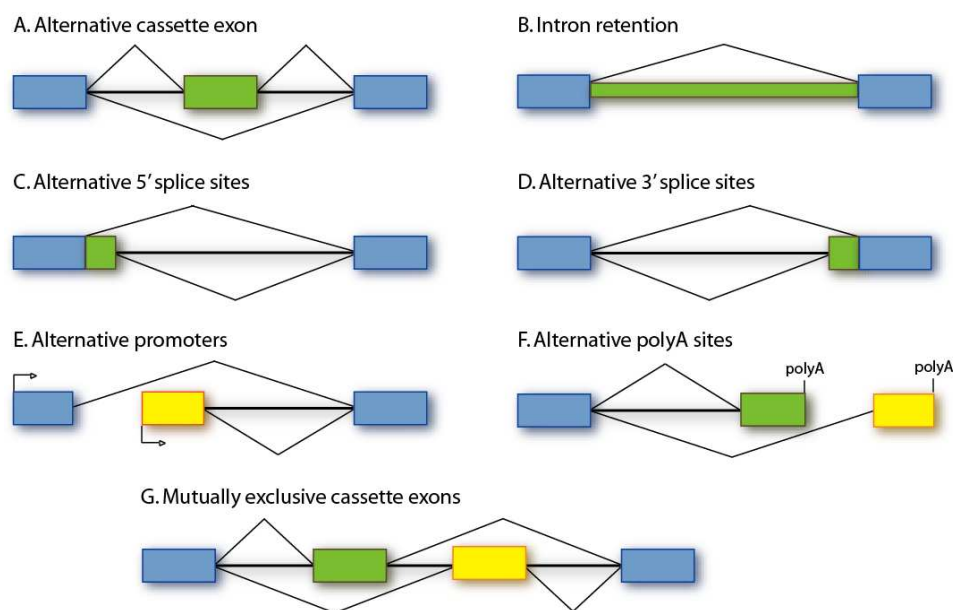
Fig. 7. Alternative splicing patters. Schematic representation of the different alternative splicing events. Constitutive and alternatively included sequences are represented by blue and green/yellow boxes, respectively. Thick black lines indicate introns and thin lines splicing patterns. **(A)** Alternative cassette exons. **(B)** Intron retention. **(C)** and **(D)** Alternative 5' and 3' splice sites, respectively. **(E)** and **(D)** Alternative promoters and polyA sites. **(G)** Mutually exclusive exons.

As mentioned previously, splicing can be found in all kingdoms of life, pointing to an ancestral origin. However, alternative splicing seems to have appeared more recently in the eukaryotic tree. This alternative processing is much more common in higher than in lower eukaryotes: more than 90% of human multiexon genes show more than one isoform (Wang et al., 2008b; Wilhelm et al., 2008), while in the yeast *Saccharomyces cerevisiae* just 3% of the annotated genes (≈250) undergo splicing and only six genes contain more than one intron (Spingola et al., 1999). In lower organisms, as yeast, plants and nematodes, intron retention is the most common form of alternative splicing (Kim et al., 2008). Contrarily, in higher metazoans, exon skipping is the most prevalent event (reviewed in Black, 2003). Despite the

obvious correlation between alternative splicing frequency and organism complexity, the contribution of the former to the latter is rather controversial (reviewed in Keren et al., 2010).

A particularly well-studied event is the evolution of alternatively spliced exons. These seem to have evolved via three different evolutionary mechanisms: exon shuffling, exonization of intronic sequences and transition of a constitutive to an alternative exon (reviewed in Keren et al., 2010). A classical example of exonization, which has been proposed to be important for the primate evolution, is that of Alu elements, a primate lineage-specific retrotransposable family (Lev-Maor et al., 2003).

A paradigm of alternative splicing-driven complexity is the *Drosophila*Down syndrome cell adhesion molecule (Dscam) gene, which can potentially give rise to more than 38.000 isoforms via alternative splicing of 95 cassette exons. This hipervariability provided by alternative processing of Dscam pre-mRNA is essential for neuronal wiring and for the generation of antibody-like factors with very different, though specific, affinities (Schmucker et al., 2000; Watson et al., 2005).

Although the inclusion of certain sequences in the final transcripts can cause the insertion of whole domains in the translated proteins, more subtle but still functional insertions/deletions are a common mode of alternative splicing. The presence of tandem 5'ss and 3'ss preserving the reading frame (e. g. GUNGUN or NAGNAG where N is any nucleotide) have been described as a major form of alternative splicing (Akerman and Mandel-Gutfreund, 2006) and postulated to fine tune the activity of the resulting proteins (Hiller et al., 2008). In the case of the NAGNAG motifs, usually the first AG is selected by a scanning-like process carried out by the spliceosome, which would recognize the first AG located downstream of the branch point during the second step of splicing (Smith and Nadal-Ginard, 1989; Smith et al., 1993). However, the nucleotide preceding each of the AG dinucleotides and the distance between both acceptor sites can determine which of them will be selected (Lev-Maor et al., 2003; Hiller et al., 2004).

Alternative splicing does not only alter the coding sequence of a gene, and consequently its function, but can also shut down its expression by introducing premature stop codons, targeting mRNAs to non-sense mediated decay (reviewed in McGlincy and Smith, 2008) and contributing to attenuate splicing defects due to spliceosome component mutations (Kawashima et al., 2009). However, non-productive splicing has been observed in many splicing regulators associated to highly or ultraconserved sequences, suggesting regulation of translation via unproductive splicing (RUST) (Lareau et al., 2007).

Combinatorial control

Many factors, not just the core components of the spliceosome, contribute to a characteristic splicing signature different depending on the cell type, developmental stage and cell environment. Splice site strength, exon-intron organization, splicing enhancers and

silencers, RNA secondary structure and other RNA processing events contribute importantly in the recognition of the highly degenerated splicing signals (reviewed in Hertel, 2008) (Fig. 8). A splicing code, which would allow the prediction of mRNA isoforms in different conditions and cell types, has been pursued for many years, but the huge complexity involving splice site selection has greatly complicated reaching this goal. However, an important step forward to decipher the splicing code has been recently made (Barash et al., 2010).

a. Splice site strength

A factor clearly contributing to the differential inclusion of coding sequences in the final mRNA is the splice site strength. This is mainly predicted by the potential base pairing of U1 snRNA with the 5'ss, the ability of U2 snRNA to base pair with the branch site and bulge out the branch adenosine, the length and pyrimidine content of the pY tract and the nucleotides surrounding the AG dinucleotide acceptor site(Horowitz and Krainer, 1994; Ruskin and Green, 1985; Zhang, 1998).

However, a non-perfect complementarity of the 5'ss not always leads to reduced splice site recognition. Recent data has shown that shifting the U1 snRNA-5'ss base pairing by just one nucleotide can help to recognize non-canonical 5'ss (Roca and Krainer, 2009). Moreover, U1 snRNP is not required for splicing of certain pre-mRNAs, at least in the presence of an excess of SR proteins (Crispino et al., 1994; Crispino et al., 1996).

The pY tract and the AG acceptor site are bound by the 65 KDa and 35 KDa subunits of the U2 snRNP auxiliary factor, respectively (Merendino et al., 1999; Wu et al., 1999; Zamore et al., 1992; Zorio and Blumenthal, 1999). The recognition of the AG dinucleotide by U2AF35 is required for U2AF65 binding in short pY tracts with poor pyrimidine content (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999). This fact allows to conceptually classifying introns as AG-independent, if the acceptor site is not required for the first transesterification reaction, or AG-dependent if the AG is necessary for the first step (Reed, 1989). PUF60, a factor functionally related and homologous to U2AF65, can also bind the pY tract as well as interact with SF1, U2AF35 and SF3b155 (Corsini et al., 2009) and promote splicing of introns with weak 3'ss (Hastings et al., 2007). Which elements dictate the specificity for U2AF65 and PUF60 binding to the pre-mRNA and how these two factors interplay remain to be determined.

Other factors, like the distance between acceptor site and branch point and the presence of competing acceptor sites can also alter the recognition of the 3'ss (Ruskin et al., 1984; Reed and Maniatis, 1985; Chua and Reed, 2001).

b. Exon-intron organization

In humans, the mean length of U2-type introns is more than 3.4 Kb (Deutsch and Long, 1999), while exons are much smaller, with an average size of 137 bp (Hawkins, 1988). On the other hand, in lower organisms such as yeast and the nematode *C. elegans*, introns are

significantly smaller while exons are slightly larger. This raises the question of how exons are recognized amid the huge introns in higher eukaryotes and if this mechanism applies also to lower organisms harboring different exon-intron architecture.

In organisms with rather short introns, the spliceosome defines the intron as a unit, i.e. assembles solely across the intron recognizing the 5'ss and the 3'ss of the same intron (Fox-Walsh et al., 2005). On the other hand, to facilitate the recognition of the small exons among vast flanking intronic sequences, the basal splicing machinery recognizes the splice sites across the exon, rather than across the intron, thus defining the coding sequences at early steps of the splicing reaction (Berget, 1995). Since SR proteins (see below) can specifically interact with U1 snRNP and U2AF35, they have been postulated to mediate the cross-exon communication (Ellis et al., 2008; Wu and Maniatis, 1993). Special cases are exons containing just one splice site (e. g. first and terminal exons). In these cases the exon definition process is mediated by the crosstalk between the cap binding complex and the first 5'ss (Izaurralde et al., 1994; Lewis et al., 1996), and the cleavage and polyadenylation machinery and the last 3'ss of the transcript (Vagner et al., 2000).

The efficiency of exon inclusion in the final mRNA largely depends on both intron and exon lengths. An increase in exon size correlates with decreased inclusion, probably due to problems in crosstalk along the exon, while shortening of introns promote recognition of the splice sites. Thus, exons flanked by long introns show a higher probability to be alternatively spliced (Fox-Walsh et al., 2005).

c. Splicing silencers and enhancers

Recognition by the splicing machinery of the short and degenerated sequences that are the splice sites is highly dependent on the recognition of other regulatory cis-acting factors or splicing enhancers. Conversely, splicing silencers are sequences that can inhibit this recognition, leading often to exon skipping or to the use of alternative splice sites. These sequences, which are present both in introns and exons, are bound by non-snRNP trans-splicing factors. The presence of intronic or exonic splicing enhancers (ISE and ESE)/silencers (ISS and ESS) as well as the relative concentration of the trans-acting factors, particularly at transcription site, have a strong effect on the mRNA outcome (reviewed in Smith and Valcarcel, 2000; Black, 2003). There are two main non-snRNP, RNA-binding families of proteins involved in regulation through auxiliary sequences: SR and hnRNP proteins.
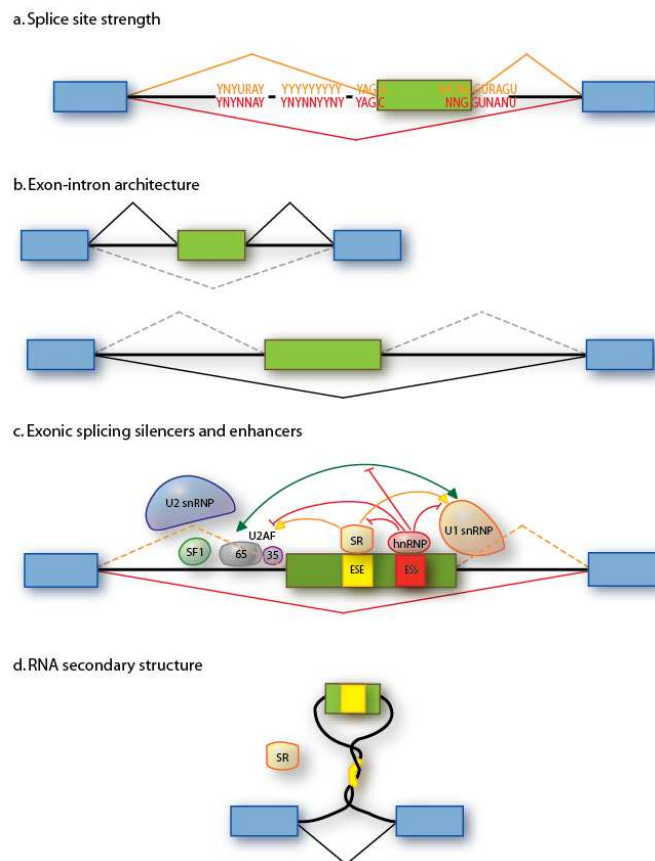
Fig. 8. Combinatorial control of alternative splicing. Alternative splicing is highly regulated by many different factors. Blue and green boxes represent constitutive and alternative exons, respectively. Thick lines correspond to introns and thin lines to splicing patterns. Dashed lines represent minor splicing events. **(A)** Splice site strength. Splice site similarity to the consensus promotes exon inclusion. **(B)** Exon-intron architecture. The longer introns and exons are, the more exon skipping is usually observed. **(C)** Splicing silencers and enhancers. Splicing enhancers, mainly bound by SR proteins, interact with components of the spliceosome to stabilize their binding to adjacent splice sites. SR proteins bind to exonic splicing enhancer elements (ESE, yellow box) to stimulate U2AF binding to the upstream 3'ss, or U1 snRNP binding to the downstream 5'ss. hnRNP proteins bind exonic splicing silencers (ESS, red box) having opposite effects, inhibiting the recognition of the splice sites by U2AF and U1snRNP and of ESE by SR proteins. **(D)** RNA secondary structure. Regulatory sequences (yellow boxes) or exons in stems or looped out by secondary structures cannot be bound by trans-acting factors.

The recent use of in vivo crosslinking and immunoprecipitation (CLIP) assays followed by high-throughput RNA sequencing has allowed drawing functional maps of several splicing factors, showing the correlation between splicing regulators binding upstream or downstream of an alternatively spliced exon and their activity (Llorian et al., 2010; Ule et al., 2006; Yeo et al., 2009; Xue et al., 2009). Nova1 and Fox1 promote inclusion when bound downstream of a cassette exon and skipping when bound upstream (Ule et al., 2006; Yeo et al., 2009). However, contradictory reports about the positional effect of PTB on alternatively spliced

exons have been published (Llorian et al., 2010; Xue et al., 2009). If PTB indeed acts as Nova1 and Fox1, these observations raise the intriguing possibility that a general mechanism exist for the regulation of alternative splicing in a position-dependent manner by very different trans-acting factors.

SR proteins

One of the most studied families of RNA binding proteins are the phylogenetically conserved Serine/Arginine-rich (SR) proteins. The members of this family have a modular organization containing one or two N-terminal RRMs and a C-terminal arginine/serine-rich (RS) domain of at least 50 amino acids with >40% content in RS or SR dipeptides, (Manley and Krainer, 2010) (Fig. 9). This family includes twelve members, although many other proteins, such as both U2AF subunits and U1 70K, contain serine/arginine repeats and belong to the so-called SR-related family of proteins (Boucher et al., 2001).
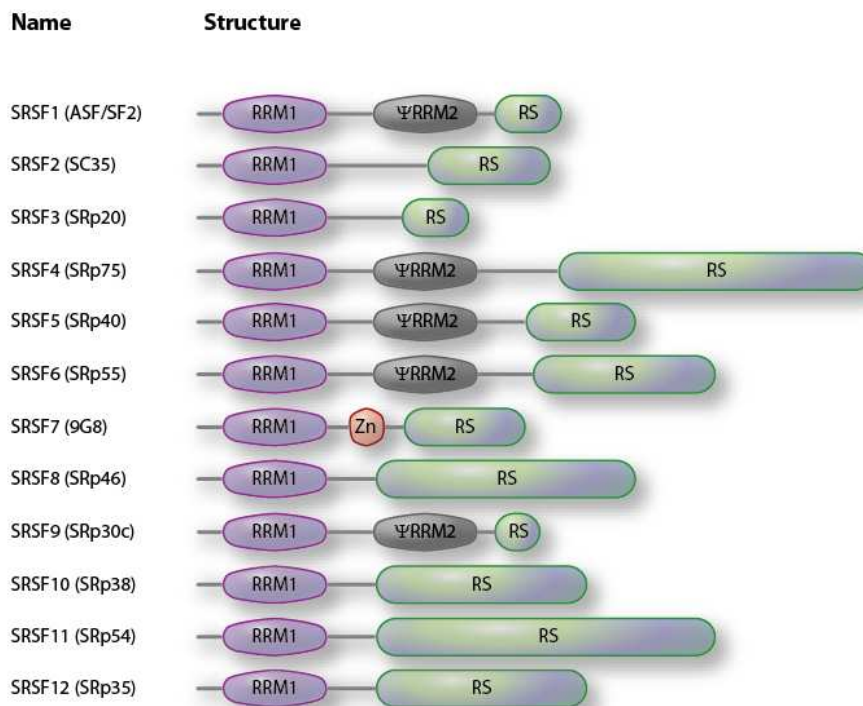
Fig. 9. SR protein family. Schematic representation of human SR proteins and their domain composition. RRM, RNA recognition motif. ΨRRM: pseudo-RRM; Zn: zinc finger; RS: arginine/serine-rich domain.

Although SR proteins are localized in the nucleus, concentrating in speckles (Caceres et al., 1997), some are able to shuttle between the nucleus and the cytoplasm (e.g. SF2 and SRp20 and 9G8) while others (e.g. SC35 and) are confined to the nucleus (Caceres et al., 1998). In fact, SR protein function and subcellular localization is highly regulated via post-translational modificacions, especially by cycles of phosphorylation and dephosphorylation of the RS domain adding another layer of splicing regulation (Colwill et al., 1996; Rossi et al., 1996; Gui et al., 1994; Wang et al., 1998b). Several types of protein kinases have been

identified to phosphorylate the RS domain either in vitro or in vivo: cdc2 (Okamoto et al., 1998), DNA topoisomerase I (Rossi et al., 1996; Labourier et al., 1998), Clk/Sty (Nayler et al., 1997; Colwill et al., 1996) and SR protein kinase (SRPK) (Koizumi et al., 1999; Gui et al., 1994)families of proteins. Dephosphorylation of the RS domain, possibly by protein phosphatases (PP) 1 and 2A, is also required for nuclear export and release of the mRNA once in the cytoplasm (Lin et al., 2005). Moreover, SF2 methylation at residues R93, R97 and R109 has been suggested to target SF2 to the nucleus, thus modulating its ability to regulate alternative splicing, translation and NMD (see below) (Sinha et al., 2010).

SR proteins have been implicated putatively in all gene expression regulation steps. Although initially observed to interact with transcriptionally active chromatin in both the lampbrush chromosome loops of Xenopus oocytes (Roth et al., 1990) and polytene chromosomes in Drosophila (Champlin et al., 1991), they emerged as essential factors for both constitutive and alternative splicing.

In the early 90's several groups identified the SR protein SF2, as a factor required to reconstitute splicing in cytoplasmic S100 HeLa cell extracts (Krainer et al., 1990a) and to promote usage of proximal 5'ss (Krainer et al., 1990b), although these activities seemed to require different molecular interactions (Caceres and Krainer, 1993). Later on, the use of the monoclonal antibody m104 produced against proteins from the germinal vesicle of the frog *Xenopus laevis*, which recognizes the phophorylated RS domain of SR proteins (Roth et al., 1990), led to the identification of more members of the family (Roth et al., 1991; Zahler et al., 1992). Many papers in the past 20 years have documented an essential role of SR proteins in promoting exon inclusion by binding to both intronic and exonic enhancers (reviewed in Bourgeois et al., 2004). However, some of them, such as dephosphorylated SRp38 (SRSF10) can also act as a general repressor of splicing during mitosis or upon cellular stress (Shin et al., 2004).

Although splicing regulation has been the main process in which the role of SR proteins have been well documented, other essential steps in RNA processing such as transcription, mRNA export, NMD and mRNA translation have been shown to be regulated by SR proteins.

As mentioned above, SR proteins were first identified as proteins localizing in transcriptionally active chromatin, but it has not been until few years ago that it has been obvious a link between the transcription machinery and the SR proteins. In 1999, Misteli and Spector showed that SR proteins, in an RS domain-dependent manner, are able to interact either directly or indirectly with the phosphorylated RNA pol II CTD in vivo (Misteli and Spector, 1999). Data coming from purified pol II complexes also showed that several SR proteins associate with pol II (Das et al., 2007) and ChIP against SR proteins demonstrated that these factors can interact predominantly with the elongating polymerase (i. e. Ser2-phosphorylated CTD) (Sapra et al., 2009).

SR protein recruitment by the transcription machinery has been also described to improve elongation efficiency: SC35 knockdown was reported to decrease the interaction of pol II with the positive transcription elongation factor b (pTEFb), a kinase responsible for CTD Ser2 phosphorylation which is essential for transcription elongation (Lin et al., 2008). Moreover, both SRp20 and SF2 are able to interact with the tail of histone H3, suggesting that direct binding of SR proteins to chromatin can play a role in both transcription and splicing (Loomis et al., 2009).

There is growing evidence that SF2 plays a role in enhancing translation. Hypophosphorylated SF2 has been described to associate with polyribosomes and to enhance translation of ESE-containing minigenes when SF2 was overexpressed (Sanford et al., 2004). Moreover, the same group showed that SF2-mediated translation initiation enhancement was dependent on the mammalian target of rapamycin (mTOR) pathway (Michlewski et al., 2008). The mTOR complex 1 (TORC1) promotes translation initiation upon a variety of cell signals such as insulin, growth factors and mitogens by regulating the phosphorylation state of eIF4E binding protein 1 (4E-BP1). When 4E-BP1 is hyperphosphorylated by TORC1, the cap-binding protein eIF4E is released, increasing its effective concentration and leading to binding of the eIF4F complex -composed by eIF4E, eIF4G and eIF4A- to the 5' end of the mRNA, resulting in initiation of translation (Ma and Blenis, 2009). SF2 was shown to regulate this event by interacting with both mTOR and protein phosphatase 2A, which also regulates 4E-BP1 phosphorylation state, specifically for capped mRNAs containing SF2 binding sites. Moreover, the second RRM of SF2 was required for binding to proteins of the mTOR signaling pathway and for efficient translation stimulation (Michlewski et al., 2008).

Finally, SR proteins have been shown to be implicated in mRNA export (Huang and Steitz, 2001; Huang et al., 2004), NMD (Sato et al., 2008; Zhang and Krainer, 2004) and mRNA stability by binding to 3' UTR (Lemaire et al., 2002).

Therefore, SR proteins are important regulators of mRNA metabolism and translation. Generation of null mutant mice or conditional knock-out mice of different SR proteins underscored the importance of this family of proteins on heart and retinal development (Feng et al., 2009; Xu et al., 2005; Kanadia et al., 2008).

hnRNP proteins

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are, together with histones, the most abundant class of proteins in the nucleus, although some can shuttle between the nucleus and the cytoplasm (Piñol-Roma and Dreyfuss, 1992). They contain different types of RNA binding domains (e.g. RRMs, KH domains) as well as domains rich in glycine and other aminoacids, which can mediate RNA-protein and protein-protein interactions, but unlike SR proteins, hnRNP family members do not present serine/arginine repeats (reviewed in Dreyfuss et al., 1993; Matlin et al., 2005; Singh and Valcarcel, 2005).

hnRNP proteins have roles in many steps of RNA metabolism like splicing, 3' end processing (Hector et al., 2002), translational regulation (Kalifa et al., 2006) or mRNA transport (Dreyfuss et al., 2002). Due to their ability to bind DNA, hnRNP proteins have also been described to contribute to DNA replication and repair and telomere length maintenance (reviewed in He and Smith, 2009).

These proteins bind the nascent transcripts as soon as they leave the pol II RNA exit channel and pack the mRNA into regular 40S particles (McAfee et al., 1996). The composition of this complex, which also contains SR proteins, is modified through the hnRNP particle life, probably regulating nucleocytoplasmic export and other processes in the cytoplasm.When the mRNA is exported to the cytoplasm the exlusively nuclear hnRNP proteins (e. g. C and U) dissociate from the complex, while others like A1, E and K, are exported with the mRNP to the cytoplasm (reviewed in Gorlich and Kutay, 1999;Mili et al., 2001). There, hnRNP proteins can regulate several aspects of the mRNP metabolism. For example, hnRNP A1 can promote stress-induced mRNA localization to stress granules (Guil et al., 2006) or initiation of IRES-mediated translation (Cammas et al., 2007).

The role of hnRNPs in splicing is usually associated with the recognition of splicing silencers antagonizing directly or indirectly SR proteins. A classical example is hnRNP A1, which was first identified to counteract SF2 during 5' splice site selection of the E1A pre-mRNA (Mayeda and Krainer, 1992). Many other cases of hnRNPs antagonizing SR proteins and promoting exon skipping have been reported over the last years (Marchand et al., 2002; Rooke et al., 2003; Singh et al.,2004; Sun et al., 1993). Alternative mechanisms by which hnRNP proteins promote exon skipping consist in inhibiting intron or exon definition crosstalks (Izquierdo et al., 2005; Sharma et al., 2005; Sharma et al., 2008), direct competition with U2AF65 for the binding to the pY tract (Lin and Patton, 1995; Matlin et al., 2007; Singh et al., 1995) or looping of exons through contacts between hnRNPs flanking these exons (Oberstrass et al., 2005).

d. RNA secondary structure and alternative splicing

RNA has an intrinsic tendency to form highly stable and complex secondary and tertiary structures. However, in vivo, as soon as the pre-mRNA leaves the polymerase pore, it is bound by many sequence-specific and non-specific RNA binding proteins generating mRNP complexes. Consequently, pre-mRNA has a very short timespan to form intramolecular interactions. These RNA secondary structures can affect splicing by two simple mechanisms. First, hindering the accessibility of splicing factors to functional sequences (either splice sites or other regulatory sequences) present in the pre-mRNA by sequestering them in stems or looping them out (Hiller et al., 2007). Second, different secondary structures vary the relative distance between these elements, consequently altering their ability to modify splice site recognition (reviewed in Buratti and Baralle, 2004).

Related to the latter, a recently reported, unique example of the contribution of secondary structures in splice site choice is the selection of just one out of 48 mutually exclusive exons of *Dscam* pre-mRNA. A docking site, located in the intron downstream of a constitutive exon, is partially complementary to a sequence, the selector site, situated upstream of each 3'ss, thus ensuring that just one exon can be included in the final transcript (Graveley, 2005).

Moreover, the presence of secondary structures on the pre-mRNA can also promote the deamination of adenosine to inosines (I) by the RNA editing process, mediated by the ADAR family of proteins (Bass and Weintraub, 1988; Kim et al., 1994; Melcher et al., 1996). These enzymes bind to the RNA via double-stranded RNA recognition motifs (Patterson and Samuel, 1995) and can decrease the stability of the structure by editing a A-U base pair into a less stable, non-canonical I-U base pair (Bass and Weintraub, 1988). The destabilization of the helix can affect the recognition of splicing functional sequences by the mechanisms previously mentioned. Furthermore, the presence of helixes can lead to mRNA hyperediting and, consequently, mRNA retention in the nucleus (reviewed in DeCerbo and Carmichael, 2005).

e. mRNA processing, transcription and chromatin

It has been known for several years that splicing is coupled to other pre-mRNA processing events and, due to the co-transcriptional nature of splicing, to pol II and transcription elongation rate. As mentioned above, the recognition of the pre-mRNA first exon 5'ss and last exon 3'ss are highly dependent on capping (Lewis et al., 1996) and polyadenylation (Vagner et al., 2000), respectively, which are in turn dependent on pol II phosphorylation state (Ahn et al., 2004; Bird et al., 2004; Cho et al., 1997; McCracken et al., 1997).

Several mechanisms of alternative splicing regulation by the transcription machinery have been suggested. The first, and maybe the more intuitive, mechanism relates to pol II rate of transcription: the slower the RNA is transcribed, the more time an alternative exon has to be included before the downstream, competing exon is synthesized (de la Mata et al., 2003). A reduced elongation rate promoted by UV radiation-induced DNA damage (Munoz et al., 2009), chromatin condensation triggered by siRNA transfections (Allo et al., 2009) or directly by chromatin remodeling by the Brm subunit of the complex SWI/SNF (Batsche et al., 2006) can also affect splice site choice. A second, non-exclusive possibility, coming from both in vitro and in vivo data, is that the C-terminal domain of pol II acts as a landing pad for RNA processing factors, increasing their effective concentration in the proximity of the nascent transcripts and facilitating the recognition of their cognate sequences (Cramer et al., 1999; Das et al., 2007; de la Mata and Kornblihtt, 2006).

Recently, the relationship between nucleosomes, covalent histone modifications (epigenetic marks) and splicing regulation has emerged. Already in 1991, a possible link between exon-intron architecture and nucleosomes was suggested, on the basis of the periodicity of around 200 nucleotides between 5' and 3' ss, similar to the DNA length wrapped in a nucleosome (Beckmann and Trifonov, 1991). This observation led to the hypothesis that nucleosomes may be preferentially positioned on the chromatin sequence corresponding to exons, therefore defining where the position of coding sequences already at the DNA level. This idea was confirmed by a number of laboratories, combining genome-wide nucleosome-positioning data sets from several organisms (including humans) with expressed sequence tag (EST) databases. Nucleosomes seem to preferentially position on genomic regions that will be transcribed as exons, especially if their surrounding splice sites are weak. This positioning is not dependent on gene transcription by pol II and it is absent in pseudoexons (e. g. sequences that contain splicing signals but are not recognized as such). Moreover, a histone mark, histone H3 trimethylation in lysine 36, was preferentially found in exons compared to introns (Kolasinska-Zwierz et al., 2009; Nahkuri et al., 2009; Schwartz et al., 2009; Spies et al., 2009; Tilgner et al., 2009).

Recently, a compelling care towards defining a complete splicing code was reported combining microarray data from different tissues, binding sites for regulatory factors, motifs enriched around alternatively splice exons, exon/intron organization, evolutionary conservation and RNA secondary structure (Barash et al., 2010). Although many factors were taken into account, other splicing regulatory processes, such as transcription rates and epigenetic marks, were not considered. Despite this progress and recent advances in understanding the importance of different cellular processes that influence alternative splicing, there is still much to be learnt to generate a fully predictive splicing code (Tejedor and Valcarcel, 2010).

Fas alternative splicing

Fas receptor (Fas), also known as CD95 or APO1, is a member of the tumor necrosis factor receptor (TNF-R) family (Itoh et al., 1991). Fas is a type I transmembrane receptor composed by a single transmembrane domain, an N-terminal extracellular domain, which will bind to the ligand (FasL), and an intracellular C-terminal domain, which contains the death domain (DD) that transmits the death signal from the cell's surface into the cell via protein-protein interactions (Itoh et al., 1991; Nagata, 1997).

Fas signaling plays a critical role in the control of the immune system. This was demonstrated after discovering that certain spontaneous mutant mouse strains developing lymphadenopathy and systemic lupus erythematosus and a large fraction of patients suffering from autoimmune lymphoproliferative syndrome carry defects in Fas and FasL genes (Watanabe-Fukunaga et al., 1992; Takahashi et al., 1994; Fisher et al., 1995; Rieux-Laucat et

al., 1995). Moreover, these patients and mouse models show an increased predisposition to lymphoma development, indicating that Fas signaling is critical for tumor suppression at least in the lymphoid compartment (Straus et al., 2001; Watanabe-Fukunaga et al., 1992; Takahashi et al., 1994). However, there is evidence that in addition to inducing cell death, Fas can mediate other activation signals, including increased cell proliferation and activation of the transcription factor NF-κB (reviewed in Peter et al., 2007). In agreement with these data, FasL-FasR signalling has been reported to be important for the generation and maintenance of some cancers (Chen et al., 2010).

The FasR is encoded by the tumor necrosis factor receptor superfamily member 6 (TNFRSF6), located in chromosome 10, giving rise to eight splice variants (Fig. 10A). Of especial interest is the alternative inclusion or skipping of the exon 6, which encodes the transmembrane domain of the receptor regulating its ability to induce apoptosis: inclusion of exon 6 leads to the membrane-bound isoform (mFas) of the receptor, which is able to induce apoptotic signaling upon binding of the Fas Ligand, while skipping of exon 6 leads to a soluble, anti-apoptotic form of Fas (sFas) (Cheng, 1994). The switch from skipping to inclusion of Fas exon 6 is essential for programmed cell death of the expanded T cell population, once the immune response has eliminated the antigen (Russell, 1995b) (Fig. 10B).
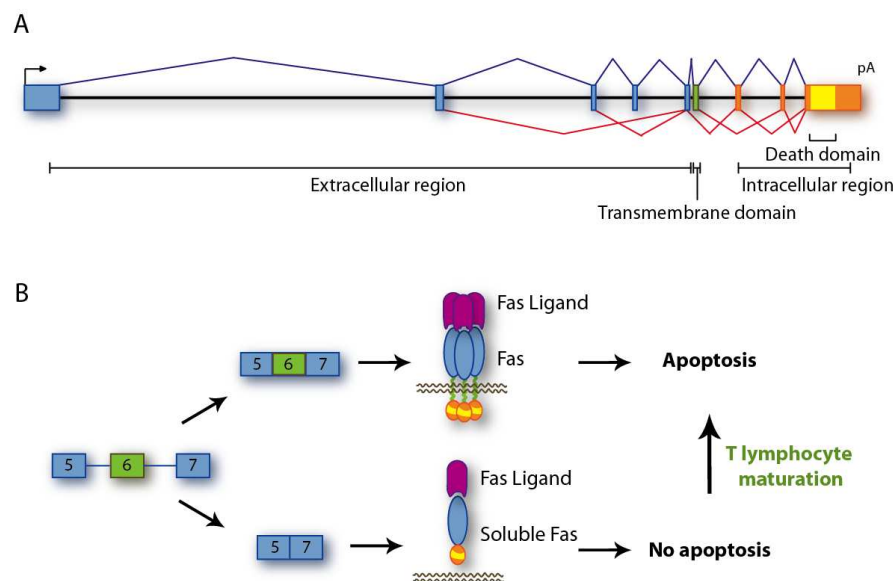
Fig. 10. Fas receptor is alternatively spliced. (A) Map of Fas locus. Exons 1 through 5 encode the extracelluar region (blue), exon 6 the transmembrane domain (green) and exons 7-9 encode the intracellular region (orange). In yellow the death domain is depicted. Blue lines show the reference transcript splicing pattern. Red lines indicate described alternative splicing events. (Modified from Fast DB: http://www.fast-db.com/fastdb2/frame.html). **(B)**Schematic representation of part of the human Fas receptor gene. Inclusion of Fas exon 6 (green) generates a membrane bound form of the receptor leading to apoptosis upon Fas ligand binding (purple). Skipping of exon 6 produces a soluble form that inhibits apoptosis by binding to the ligand.

Our laboratory and others have characterized several Fas exon 6 alternative splicing trans-acting regulators:

- The polypyrimidine-tract binding protein (PTB) and the antiapoptotic regulator Hu antigen R (HuR, ELAVL1) bind to a uridine-rich region in exon 6 (URE6) and block the exon definition process, by which U1 snRNP binding to the downstream 5'ss promotes U2AF65 interaction with intron 5 pY tract. This leads to inhibition of exn 6 definition and to exon 6 skipping (Izquierdo et al., 2005; Izquierdo, 2008b).

- The apoptosis-inducing protein TIA-1 promotes U1 snRNP interaction with exon 6 5'ss by binding to a uridine-rich region downstream of this site (URI6). Better binding of U1 snRNP results in enhanced U2AF65 interaction with the upstream 3'ss and consequently exon 6 inclusion (Izquierdo et al., 2005).

- The tumor suppressor RBM5 promotes exon 6 skipping by inhibiting the incorporation of U4/U6/U5 tri-snRNP on either intron 5 or 6, while promoting pairing between the distal 5' and 3'ss (Bonnal et al., 2008).

- Other splicing factors such as SPF45 (Corsini et al., 2007) and SF2 (Goncalves et al., 2009) have been described to promote exon 6 skipping.

Splicing and disease

Splicing defects: key contributors to human disease

The capacity that alternative splicing provides the cell with to regulate efficiently and accurately its transcriptome and, consequently, its proteome in different situations also makes the cell vulnerable to mutations that affect this tight regulation of gene expression. Alterations in both cis- and trans-acting sequences have been shown to be an important source of human disease. For example, single point mutations directly affecting splice sites account for 9-10% of the genetic diseases (Wang and Cooper, 2007). On the other hand, many other nonsense, missense and translationally silent mutations, which would not in principle be considered as a cause of pathology, occurring in intronic or exonic enhancers/silencers can lead to unpredicted aberrant, disease-causing alternative splicing patterns. An additional limitation is that, as mentioned above, these regulatory sequences are short and highly dependent on their location with respect to other functional elements and probably act in essentially all splicing events (Roca et al., 2008). These factors, together with the interesting observation that genes involved in disease contain a higher intron content than those genes not involved in disease, led to the prediction that more than 50% of all disease-causing mutations would be affecting splicing (Lopez-Bigas et al., 2005). It is not yet possible to predict from genomic

sequence alone if a disease-causing mutation exerts its effects by disrupting splicing. The combination of bioinformatic tools and high-throughput technologies for RNA sequencing will eventually lead to a splicing code able to predict isoform changes associated with disease-causing or disease-modifying mutations. Relevant to these considerations, recent data have shown that individual-specific polymorphisms affect between 21% and 30% of all the alternative splicing events (Nembaware et al., 2004; Wang et al., 2008a). This heterogeneity can potentially influence multiple aspects of disease, such as severity, susceptibility and response to therapeutic compounds.

Many splicing-disrupting mutations causing pathologies have been described, but the mechanisms underlying these effects have been deciphered in only very few cases. Interestingly, a number of these mutations lead to devastating neurodegenerative and neuromuscular diseases and cancer (reviewed in Cooper, 2009). For example, Duchenne muscular dystrophy (DMD) is caused by mutations or insertions/deletions in the 78 intron-containing dystrophin gene and has an incidence of 1:3000 of male births. A mild form of DMD is caused by a mutation generating a stop codon and at the same time a splicing silencer recognized by hnRNP A1 in exon 31 (Disset et al., 2006). Increased exon 31 skipping preserves the reading frame and avoids the inclusion of the premature stop codon, leading to the translation of a partially functional protein and a less severe form of the disease (Disset et al., 2006), suggesting a therapeutic approach based upon splicing modulation.

Other mutations affecting splicing may not directly cause disease but importantly affect its severity. Such is the case of mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene causing cystic fibrosis. This disease can be caused by different mutations but the observed phenotypes vary greatly depending on the level of exon 9 inclusion or skipping. This event is regulated by polymorphic (UG)m and (U)n tracts on the intron 8 3'ss (Chu et al., 1991). Longer (UG)m tracts correlate with more exon 9 skipping, due to an increased binding of TAR DNA-binding protein-43 (TDP-43), and therefore modulate the severity of the disease (Buratti et al., 2004).

A different mechanism involves mutations disrupting functions of RNA binding proteins. Microsatellite expansions (e. g. triplet repeats) not only can cause a loss or gain of protein function if present in the coding sequence, but can also affect RNA metabolism by sequestering RNA binding proteins in expanded non-coding RNA regions. The best-studied example is the expansion of CUG repeats at the 3' UTR of the dystrophia myotonica-protein kinase (DMPK) gene that leads to myotonic dystrophy (DM) (Brook et al., 1992; Fu et al.). CUG repeats form a stable hairpin that is specifically recognized by muscleblind-like 1 (MBNL1) protein (Miller et al., 2000), which is sequestered and unable to carry out its activity in different processes, such as splicing, mRNA localization or nuclear export (Pascual et al., 2006). Moreover, the presence of these repeats leads to activation of protein kinase C which hyperphosphorylates and stabilizes CUG binding protein 1 (CUGBP1) (Kuyumcu-Martinez et

al., 2007), altering CUGBP1-regulated alternative splicing events (Philips et al., 1998). The unbalanced ratio between MBNL1 and CUGBP1 induces embryonic rather than adult splicing pattern in adult tissues, causing the multisystemic symptoms of the disease (reviewed in Lee and Cooper, 2009).

Even though mutation of core components of the spliceosome could be considered incompatible with life or cell viability, a few diseases have been associated with mutations in these factors. Retinitis pigmentosa, characterized by retinal degeneration and, ultimately, blindness, is caused by autosomal dominant mutations in several core spliceosome factors (prp3, prp8, prp31 and rp9). Interestingly, only photoreceptor neurons are affected by the haploinsufficiency of these essential components of the spliceosome (Mordes et al., 2006; Mordes et al., 2007). Systemic lupus erythematosus is a chronic autoimmune disease, the activity of which can be correlated with the presence of anti-U1 snRNP and anti-Sm auto-antibodies (Agris et al., 1984; Hoet et al., 1992; van Venrooij et al., 1990). Also, spinal muscular atrophy (SMA), a motor neuron degenerative disease, is another prevalent disease that involves snRNP processing. SMA is caused by the inactivation of the SMN1 gene (Lefebvre et al.), component of the SMN complex which is important for the assembly of Sm proteins on most U snRNAs, an essential step for U snRNAs biogenesis (reviewed in Will and Luhrmann, 2001).

Complex diseases, such as cancer, show also a high prevalence in splicing defects. Alternative splicing regulates many pro- and anti-apoptotic factors. For example, members of the B-cell lymphoma family of proteins are often alterntatively spliced (Akgul et al., 2004). Presence of all four BH (Bcl-2 homology) domains allows the protein to antagonize apoptosis, which in turn is promoted by isoforms lacking one or more BH motifs (reviewed in Taylor et al., 2008). Many anti-apoptotic isoforms found in cancers are expressed as well in undifferentiated cells. For example, the stem cells/foetal variants of caspase 8, neurofibromatosis gene NF1, insulin receptor and fibronectin are also found in leukemias, brain and thyroid tumors and other kind of cancers (Ebbinghaus et al., 2004; Mochizuki et al., 1992; Mohr et al., 2005; Vella et al., 2002). Many other aspects of tumor development and progression such as invasion and angiogesis are also modulated by alternative splicing (reviewed in Venables, 2006). Moreover, misbalance in the expression of certain trans-regulatory factors, such as SF2, can promote tumorigenesis. Overexpression of this factor is sufficient to transform immortal mouse fibroblasts, forming sarcomas in nude mice, and has been observed in various human tumors (Karni et al., 2007). This proto-oncogenic effect is probably caused both by altering the splicing pattern of pro-apoptotic/proto-ocogenic genes (Karni et al., 2007; Moore et al., 2010) and by direct and/or indirect activation of the mTOR pathway (Karni et al., 2007; (Karni et al., 2008).

Splicing modulation and spliceosome targeting as novel therapies forhuman diseases

Because, splicing is a prevalent cause of disease, significant efforts have been made to modulate splicing in a tissue-specific, transcript-specific manner. For that purpose, several strategies have been tested in order to reverse disease-causing splicing defects and alleviate its symptoms. Antisense oligonucleotides (ASOs) and small molecules are two of the most promising approaches (reviewed in Cooper et al., 2009).

ASOs are designed to target and block pre-mRNA sequences that are functionally important for a disease-causing alternative splicing event. The use of ASO as therapeutic molecules is providing encouraging results for the treatment of spinal muscular atrophy (SMA), a motor neuron degenerative disease. SMA is mainly caused by inactivation of the SMN1 gene, component of the SMN complex (Lefebvre et al., 1995). A second, almost identical copy of this gene, SMN2, could generate a functional protein but exon 7 is skipped due to a single nucleotide difference in the position 6 of exon 7, giving rise to an inactive, unstable protein (Lorson et al., 1999). Therefore, promoting inclusion of SMN2 exon 7 would compensate for the lost of SMN1 and potentially cure the disease. ASOs designed to target an ISS in intron 7 have been shown to significantly delay and reduce the phenotypic effects on a mild SMA mouse model (Hua et al., 2008; Hua et al., 2010).

ASOs show high specificity, reducing the probability of causing side-effects, but delivery presents significant hurdles. In contrast, the use of small molecules, even broadly used drugs, has emerged as another tool to modify alternative splicing. Several small molecules have been described to stall spliceosome assembly at different stages of the pathway (Kuhn et al., 2009), to act as inhibitors of SR proteins (Bakkour et al., 2007; Keriel et al., 2009; Soret et al., 2005) and SR protein kinases (SRPKs and Clks) (Fukuhara et al., 2006; Muraki et al., 2004), or to bind core components of the spliceosome (Kaida et al., 2007; Kotake et al., 2007). The use of these molecules as therapeutic tools depends entirely on the side effects that they might cause. However, Spliceostatin A (Kaida et al., 2007) and pladienolide (Kotake et al., 2007) have been shown to bind components of the SF3b complex and to have a potent antitumor activity, and inhibitors of SR proteins and SR protein kinases can prevent early splicing events required for HIV and MLV replication (Bakkour et al., 2007; Fukuhara et al., 2006; Keriel et al., 2009; Soret et al., 2005).

Preview

In vivo UV crosslinking-immunoprecipitation (CLIP) followed by high-throughput sequencing technologies and combined with functional information (i. e. splicing-sensitive microarrays) has allowed determining the position-dependent effect of several proteins on alternative splicing (Llorian et al., 2010; Ule et al., 2006; Yeo et al., 2009; (Xue et al., 2009). In this Preview (Corrionero and Valcarcel, 2009), we analyze the importance of these maps even for well-studied factors such as PTB (Xue et al., 2009).

RNA Processing: Redrawing the Map of Charted Territory

Using genome-wide RNA-binding data, Xue et al. (2009) draw a regulatory map in this issue of *Molecular Cell* for the much-studied polypyrimidine tract-binding protein (PTB) that reveals a unique paradigm in posttranscriptional gene regulation.

When great explorers like Chen Ho, Columbus, or Shackleton set sail to uncharted land, they were prepared to endure hardship and danger but were also confident that, if successful, their expeditions would bring back goods, stories, and detailed maps of unheard-of wonders. In contrast, few explorers would be thrilled to chart well-mapped terrain. In this issue of *Molecular Cell*, Xue et al. (2009) reinvestigate one of the most studied posttranscriptional regulators of gene expression, the polypyrimidine tract-binding protein (PTB), by analyzing its binding sites in the transcriptome of human HeLa cells. Their results show that novel principles remain to be discovered even in well-charted territory.

PTB (also known as hnRNP I) is an RNA-binding protein expressed in multiple tissues, with binding preference for CU-rich sequences (Singh and Valcarcel, 2005;Oberstrass et al., 2005). Like other members of the hnRNP family, PTB plays a variety of roles in RNA metabolism, including regulation of alternative pre-mRNA splicing, polyadenylation, mRNA stability, and translation. For example, PTB has emerged as a repressor of neuron-specific alternative splicing in nonneural cells. Indeed, downregulation of PTB and upregulation of a functionally distinct neuron-specific paralog (nPTB) may be involved in about 25% of the splicing changes observed during neuron differentiation (Boutz et al., 2007). A similar switch has been reported between PTB and its hematopoietic cells paralog ROD1 (Spellman et al., 2007). Given the variety of functions and targets associated with PTB, genome-wide identification of its binding sites is of obvious general interest.

For this, Xue et al. (2009) employed an in vivo UV crosslinking and immunoprecipitation (CLIP) technique developed by Darnell and coworkers (Ule et al., 2003). This method captures close interactions between proteins and RNA molecules in living cells. After UV irradiation, cells are lysed and the crosslinked RNA-protein complexes immunoprecipitated.

Treatment with RNase digests away the RNA molecules except for those fragments protected by their interaction with the protein. Sequencing of these RNA tags not only identifies target transcripts, but also informs about the location of protein binding within the transcript. When combined with high-throughput sequencing technologies, the method allows one to exhaustively characterize binding sites in whole transcriptomes, thus providing a panoramic view of targets and possible functions of the RNA-binding protein (Licatalosi et al., 2008;Yeo et al., 2009).

The results of Xue et al. (2009) significantly expand the catalog of PTB-binding sites, confirm known gene targets of PTB-mediated regulation, refine their possible mechanisms of regulation, identify novel regulated alternative splicing events, and open intriguing new possibilities for PTB function. PTB-binding sites are found in nearly 50% of human protein-coding transcripts and in 20% of all annotated alternatively spliced regions, confirming the potential genome-wide effects of the protein. Consistent with the variety of functions that PTB plays in gene regulation, only 30% of PTB sites are associated with alternative splicing events, and additional roles in constitutive splicing and in repression of cryptic sites can be envisioned. Nearly 30% of PTB sites map to intergenic regions, perhaps indicative of a function in the metabolism of noncoding transcripts. Of relevance, most, but not all, binding sites could be predicted bioinformatically by previous knowledge of PTB RNA binding preferences, and only a small fraction of the predicted sites were detected experimentally. Sequence context effects associated with cooperative or antagonistic interactions with other factors are likely to explain these differences, arguing for the added value of experimental data.

When the binding landscape provided by high-throughput methods like CLIP is combined with functional information, for example obtained using splicing-sensitive microarrays, correlations between the location of binding sites and regulatory outputs emerge. Such representations are known as RNA maps (Fig. 11). RNA maps have been reported for the tissue-specific splicing regulators Nova and Fox2 (Ule et al., 2006;Licatalosi et al., 2008;Yeo et al., 2009). These maps share interesting features: binding of Nova/Fox2 upstream of the regulated exon correlates with Nova/Fox2-mediated exon skipping, whereas binding to the intronic region downstream of the exon correlates with exon inclusion (Fig. 11 A and B). Although the mechanisms for these position-dependent effects remain to be determined, these RNA maps suggest that Nova and Fox2 use common molecular principles to modulate splice site choice.

In contrast, the RNA map drawn for PTB by Xue et al. (2009) is remarkably different. PTB-induced exon inclusion or skipping correlates with the location of PTB-binding sites relative to the constitutive or alternative splice sites involved in the regulated event (Fig. 11C). As expected from previous knowledge of PTB-mediated regulation, PTB binding near an alternative exon correlates with skipping. Unexpectedly, PTB binding near the constitutive

distal splice sites correlates with enhanced exon inclusion. Although not without exceptions, these results greatly populate the catalog of genes in which PTB promotes exon inclusion and reveal an internal logic for the dual activity of this protein in alternative splicing regulation.
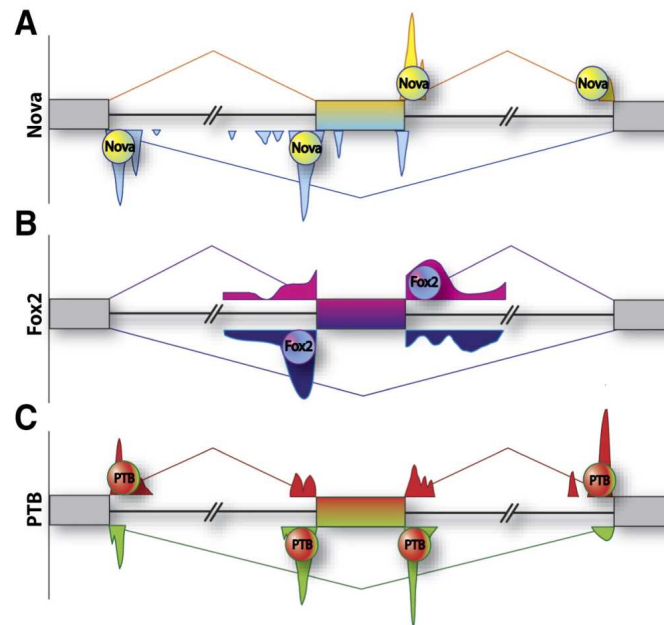
Fig. 11. RNA Maps of Three Splicing Regulators. RNA maps depict the correlation between sites of regulatory protein binding (determined by CLIP) relative to a model alternative exon and the outcome of splicing regulation (exon inclusion or skipping). Peaks represent regions with higher density of CLIP reads. Binding of Nova or Fox2 upstream of the exon correlates with exon skipping, whereas binding downstream of the exon correlates with inclusion. For PTB, skipping correlates with binding of the protein near the alternative exon, whereas inclusion correlates with binding near the distal, constitutive sites. **(A)** Nova RNA map. **(B)** Fox2 RNA map. **(C)** PTB RNA map.

A variety of mechanisms have been proposed for PTB-mediated effects on splice site selection, including steric hindrance of splice sites, interference with intron or exon definition, looping out RNA regions, or cooperative spreading of PTB from high-affinity sites (Spellman and Smith, 2006). None of these models can easily explain PTB-mediated exon inclusion. As argued by Xue et al., PTB binding may make neighboring sites less competitive by interfering with thermodynamic or kinetic parameters of their recognition by basal splicing factors. PTB could therefore serve as a valve that diverts the flow of splicing activity from a region of the pre-mRNA and thus modulates splice site competition.

The results of Xue et al. eloquently argue that high-throughput-enabling technologies provide panoramic views, illuminating mechanistic concepts that escaped the attention of researchers doing detailed work on more limited sets of genes. They also bring hope that RNA maps will one day allow us to navigate through genome sequences to predict their posttranscriptional outputs in different tissues or physiological situations.

OBJECTIVES

I. **Uncover the mechanism underlying the splicing inhibitory effect of the antitumor drug Spliceostatin A.**

II. **Determine the molecular mechanism by which the duplication of Fas intron 5 acceptor site leads to complete exon 6 skipping and Autoimmune Lymphoproliferative Syndrome (ALPS).**

III. **Identify the molecular basis for how the proto-oncogene SF2 regulates Fas alternative splicing to promote the anti-apoptotic form.**

RESULTS

Chapter I: Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug Spliceostatin A

Anna Corrionero, Belén Miñana and Juan Valcárcel

The data included in this chapter have been submitted for publication

Abstract

Spliceostatin A (SSA) is a potent anti-tumor drug that inhibits pre-mRNA splicing and binds SF3b, a protein subcomplex of the pre-mRNA branch point-recognizing U2 small nuclear ribonucleoprotein (snRNP). We report that SSA prevents interaction of SF3b155 kDa subunit with the pre-mRNA and induces non-productive recruitment of U2 snRNP to sequences 5' of the branch point. Differences in base pairing potential with U2 snRNA in this region lead to different sensitivity of 3' splice sites to SSA and to SSA-induced changes in alternative splicing. These lead to down-regulation of genes important for cell proliferation, including cyclin A2 and Aurora A kinase. Our results reveal a SF3b-mediated proofreading mechanism that prevents non-productive base pairing interactions in the spliceosome and highlight the regulatory and cancer therapeutic potential of perturbing the fidelity of splice site recognition.

Results and Discussion

Spliceostatin A (SSA) is a stabilized methyl ketal derivative of FR901464, a fermentation product of *Pseudomonas* bacteria that displays potent cytotoxic activity against various solid tumor cell lines at nanomolar IC50 and prolongs the life of tumor-bearing mice (Nakajima et al., 1996; Kaida et al., 2007). Both SSA and pladienolide, a structurally related anticancer compound from *Streptomyces*, bind to protein components of SF3b (Kaida et al., 2007; Kotake et al., 2007), a subcomplex of U2 snRNP essential for 3' splice site recognition (Brosi et al., 1993a; Brosi et al., 1993b), and inhibit pre-mRNA splicing *in vitro* and in cultured cells (Kaida et al., 2007; Kotake et al., 2007; Roybal and Jurica, 2010). To study the mechanism of splicing inhibition by SSA, a model pre-mRNA substrate from adenovirus (AdML) was incubated with HeLa nuclear extracts under splicing conditions in the absence or presence of increasing concentrations of SSA. Accumulation of splicing intermediates and products was inhibited at concentrations starting at 45 nM, with essentially complete inhibition at 100 nM (Fig. I. 1A).

Fully assembled splicing complexes contain five snRNP particles (U1, U2 and the tri-snRNP U4/5/6), and distinct sub-complexes can be resolved by native gel electrophoresis (reviewed in Wahl et al., 2009): complex A contains U2 snRNP bound to the 3' splice site region of the pre-mRNA and complexes B and catalytically active C contain in addition U4/5/6 snRNPs. SSA inhibits U2 snRNP-containing A and subsequent complexes (Fig. I. 1B), in contrast with a previous report (Roybal and Jurica, 2010). Complex A assembly on an RNA containing the 3' half of the intron and downstream exon was also inhibited, indicating that a 5' splice site is not required for SSA-mediated inhibition (Fig. I. 1C). When heparin (5 µg/µl) was omitted from the mix prior to electrophoresis, however, addition of SSA altered the electrophoretic mobility but did not inhibit the complex formed (Figs. I. 1D and E), indicating

that SSA does not prevent U2 snRNP assembly but rather weakens its interactions with the pre-mRNA, which can be displaced by the highly negatively charged RNA mimetic properties of heparin (Fig. I. 1E).

U2 snRNP binding involves base pairing between sequences flanking the pre-mRNA branch point adenosine (which forms a 2'-5' phosphodiester bond with the 5' end of the intron after the first catalytic step of the splicing reaction) and the branch point recognition sequence (bprs) in U2 snRNA (Wahl et al., 2009). While RNase H-mediated degradation of the bprs is commonly used to inactivate U2 snRNP (ΔbprsU2), a complex was unexpectedly detected between ΔbprsU2 snRNP and the 3' splice site region in the absence of heparin (Figs. I. 1D, E), indicating that –contrary to the established model of complex A formation- base pairing interactions between the pre-mRNA branch point and U2 snRNA bprs are not essential for U2 snRNP recruitment. As was the case for the complex formed in the presence of SSA, the complex was heparin-sensitive (Fig. I. 1E) and showed hallmarks of functional U2 snRNP recruitment, including the presence of U2 snRNP components, dependence on ATP, U2AF and the presence of a branch point, and sensitivity to exon definition effects (Figs. I. S1-3 and see below). We conclude that, although of lower stability, complexes between U2 snRNP and the pre-mRNA can form in the presence of SSA or in the absence of U2 bprs.

To further characterize these complexes, the interactions of protein and RNA components of U2 with the pre-mRNA were analyzed. Ultraviolet light (UV) irradiation and immunoprecipitation experiments showed that the interactions between the U2 snRNP Auxiliary Factor U2AF65 or SF3a60 with the pre-mRNA were not affected by SSA, while crosslinking of SF3b155 was significantly decreased (Fig. I. 1F, right panel). The presence of an additional crosslinked species immunoprecipitated by anti-SF3a 60 antibodies –which may correspond to the 66 Kda subunit of the SF3a complex (Brosi et al., 1993a)- further points to SSA-induced conformational rearrangements of RNA-protein contacts between U2 snRNP components and the pre-mRNA. Equivalent results were obtained when the complex formed with ΔbprsU2 was analyzed (Fig. I. 1F, left panel).

Base pairing interactions between U2 snRNA and the pre-mRNA were monitored using psoralen-mediated, long wave UV light-induced crosslinking (Fig. I. 2A). Crosslinking between U2 snRNA and the pre-mRNA was largely unaffected by SSA treatment (Fig. I. 2B, right panel). Surprisingly, crosslinking could be also detected between Δbprs U2 snRNA and the pre-mRNA, which was not affected by SSA either (Figs. I. 2A-B, left panels). RNase H-mediated digestion and primer extension assays were used to map the crosslinked residues in U2 snRNA and the pre-mRNA (Figs. I. 2A, C and D). In the presence or absence of SSA, crosslinking involved the same nucleotides in the bprs of U2 snRNA, while it involved nucleotides 3' from the bprs in Δbprs U2 (Figs. I. 2A and D). Crosslinking in the presence of SSA was detected in a region of the pre-mRNA 5' of the nucleotides crosslinked in the absence of the drug (Fig. I. 2C lanes 13 vs 15 and Fig. 2E), suggesting that SSA induces

altered base pairing between U2 snRNA and the pre-mRNA. Crosslinking to nucleotides significantly 5' distant from the branch point was also observed between Δbprs U2 snRNA and the pre-mRNA (Figiffer. I. 2C lanes 6 vs 7; 10 and 12 vs 13; and Fig. 2E). Taken together, the results indicate that SSA treatment or U2 snRNA bprs inactivation alter the pattern of U2 snRNA base pairing with the branch point, concomitant with loss of SF3b155 interaction with the pre-mRNA. To assess whether the potential of sequences 5' of the branch point to form alternative base pairing interactions with U2 snRNA can influence the effects of SSA, mutant RNAs were generated that either increase the potential of base pairing interactions of the natural branch point region or offer an alternative site of perfect base pairing (Fig. I. 2F, top). While enhancing base pairing with the proper branch site region reduced the sensitivity of the 3' splice site region to SSA, stronger inhibition was observed in a mutant that improves the base pairing potential of the region 5' of the branch point (Fig. I. 2F). These results are consistent with the possibility that the splicing inhibitory effects of SSA involve the establishment of alternative, non-productive base pairing between U2 snRNA and the pre-mRNA.

Differential sensitivity of 3' splice sites to SSA could lead to altered splice site selection in alternatively spliced transcripts. Consistent with this possibility, splicing-sensitive microarrays revealed that treatment of HeLa cells with SSA affected 11% of the alternative splicing events monitored by the array in 21% of the genes, with a validation rate by quantitative RT-PCR of 72% (Figs. I. S4-7). Gene expression changes were also observed in 10% of the genes, 72% of which were concomitant with changes in alternative splicing, often coupled to the generation of premature translation termination codons (PTC) and consequently subject to nonsense-mediated mRNA degradation (AS-NMD) (Lareau et al., 2007). Consistent with the inactivating effect of SSA on SF3b155 function, a sample of validated SSA-induced alternative splicing events showed similar changes upon SF3b155 knock down (Figs. I. 3A and 3B) and 30% of the SSA-induced alternative splicing events were also observed in microarray analyses of RNAs isolated after knock down of SF3b155 using two different siRNAs (Fig. I. S6), which represents a highly significant statistical overlap (p<0.001 Fischer exact test). Both SSA treatment and SF3b155 knock down reduced cell proliferation (Kaida et al., 2007; and data not shown) and significant down-regulation by AS-NMD of genes that encode important regulators of cell proliferation was observed in both situations (Fig. I. 3C). For example, in addition to the previously reported inhibition of p27 splicing (Kaida et al., 2007), AS-NMD was observed for CCNA2 (cyclin A2) and STK6 (Aurora A kinase) (Fig. I. 3D), which are important for cell cycle transitions similar to those blocked by SSA (Mitra and Enders, 2004; Saeki et al., 2009). Taken together, the results reveal that SSA treatment and SF3b155 inactivation induce specific changes in alternative splicing and gene expression, including changes in key regulators of cell division that can contribute to the anti-proliferative properties of the drug.

To address the mechanism of alternative splicing changes induced by SSA, we focused on an alternative splicing event in RBM5, a putative tumor suppressor gene encoding a splicing regulator implicated in the control of cell division and apoptosis (Oh et al., 2006; Fushimi et al., 2008; Bonnal et al., 2008). Both SSA treatment and SF3b155 knock down induce skipping of exon 6 (Fig. I. 3B and Fig. I. S4D) and the competing 3' splice site regions are amenable for biochemical studies (Fig. I. 4). Spliceosome assembly assays showed that recruitment of U2 snRNP to the competing 3' splice sites display differential sensitivity to SSA (Fig. I. 4A). Contrary to expectation, however, the inhibited 3' splice site of intron 5 displays higher apparent resistance to SSA than the 3' splice site of intron 6, which remains active in the presence of the drug. This counter-intuitive pattern of differential sensitivity was shared by another tested set of alternative 3' splice sites involved in similar SSA-mediated alternative splicing effects (Fig. I. S8). Consistent with the concept that differential sensitivity to SSA is modulated by sequences 5' of the branch point, a 10 nucleotide fragment located upstream of intron 5 branch point (but not a neighboring sequence of 7 nucleotides) was sufficient to confer SSA resistance to intron 6 3' splice site region (mutant uAS5-dAS6), as observed in complex formation assays (Fig. I. 4B) and as enhanced psoralen-mediated crosslinking with U2 snRNA compared to intron 6 sequences (Fig. I. 4C). A general function for the region 5' of the branch point (known as anchoring site-AS) in U2 snRNP recruitment was reported previously (Gozani et al., 1996), although its proposed role was to mediate sequence-independent interactions with U2 proteins in the context of normal U2 snRNP recruitment. As observed for AdML RNA, mapping of U2 snRNA/pre-mRNA interactions revealed that SSA induces interactions of U2 snRNA with regions 5' of the branch point (Fig. I. 4D compare lanes 5 and 7, 12 and 14, and Fig. E), concomitant with reduced SF3b155 (but not SF3a60) crosslinking (Fig. I. S9). Significantly, the novel sites of U2 interaction included a relatively distant crosslinking site located in the region that confers SSA resistance in the uAS5-dAS6 mutant (Fig. I. 4D). This region has substantial base pairing potential with U2 snRNA (Fig. I. 4E), and mutation of residues involved in this base-pairing scheme strongly increased sensitivity to SSA, while mutation of another nucleotide not involved in base pairing did not (Figs. I. 4E, F). We conclude that SSA induces base pairing of U2 snRNA to sequences substantially 5' of the branch point and that the base pairing potential of these regions with U2 snRNA influences the differential effects of the drug on alternative 3' splice site regions.

To test whether such sequences can indeed influence splice site choice, we transiently transfected HeLa cells with minigenes that include the genomic region of RBM5 between exons 5 and 7. Co-transfection with siRNAs against SF3b155 resulted in increased levels of exon 6 skipping (Figs. I. 4G and 4H), as observed for the endogenous gene upon SSA treatment or upon SF3b155 knock down (Fig. I. 3B). Based on the U2 snRNP recruitment and mapping results above, we predicted that the competitive disadvantage of intron 5 3' splice site under conditions of SSA treatment or SF3b155 knock down would be attenuated if

the sequence acting as a decoy site for U2 snRNA binding would be replaced by the corresponding sequence of intron 6 (mutant uAS6-intron 5). Consistent with this prediction, SF3b155 knock down did not induce exon 6 skipping (Figs. I. 4G and H). We conclude that the presence of sequences harboring strong base pairing potential with U2 snRNA 5' of the branch point can modulate the competitiveness of 3' splice sites in the presence of SSA and thus explain alternative splicing changes induced by the drug or by knock down of SF3b155.

While snRNA/snRNA and snRNA/pre-mRNA base pairing interactions play key roles in the dynamic architecture of the spliceosome (reviewed in Wahl et al., 2009), base-paired regions are usually short and complementarity is often imperfect, raising the question of how specific recognition of particular sequences (e.g. splice sites vs decoy neighboring sequences) can be achieved. Proofreading activities have been identified that monitor spliceosomal transitions (Staley and Woolford, 2009) and even early events in 3' splice site recognition (Soares et al., 2006). Collectively, our results reveal that U2 snRNP recruitment to 3' splice site regions is subject to a proofreading mechanism that prevents non-productive base pairing interactions between U2 snRNA and complementary sequences 5' of the branch point region, and that SSA relaxes this proofreading step. SF3b155 can play a key role in this proofreading activity, as the protein crosslinks to both sides of the branch point (Gozani et al., 1998), becomes phosphorylated coupled to catalytic activation of the spliceosome (Wang et al., 1998a) and its interaction with the pre-mRNA is specifically decreased in the presence of SSA concomitant with the establishment of alternative U2/pre-mRNA base pairing interactions (Figs. I. 1, 2 and 4). The significant overlap between changes in alternative splicing induced by SSA and by knock down of SF3b155 in tissue culture cells is also consistent with this idea. Prominent among these changes are RNA processing events leading to down-regulation of key cell cycle regulatory factors, potentially contributing to the antitumor effect of the drug. Our results raise the intriguing possibility that physiological variations in the levels or activity of SF3b155 modulate alternative splicing programs, possibly linked to the control of cell proliferation. They also suggest that modulating the accuracy of 3' splice site recognition can help in the design of novel anti-cancer therapies, a general principle with analogies to the function of a class of antibiotics that promote inaccurate base pairing interactions in the function of the ribosome (Ogle and Ramakrishnan, 2005), another complex ribonucleoprotein machinery.

Figures and Figure legends
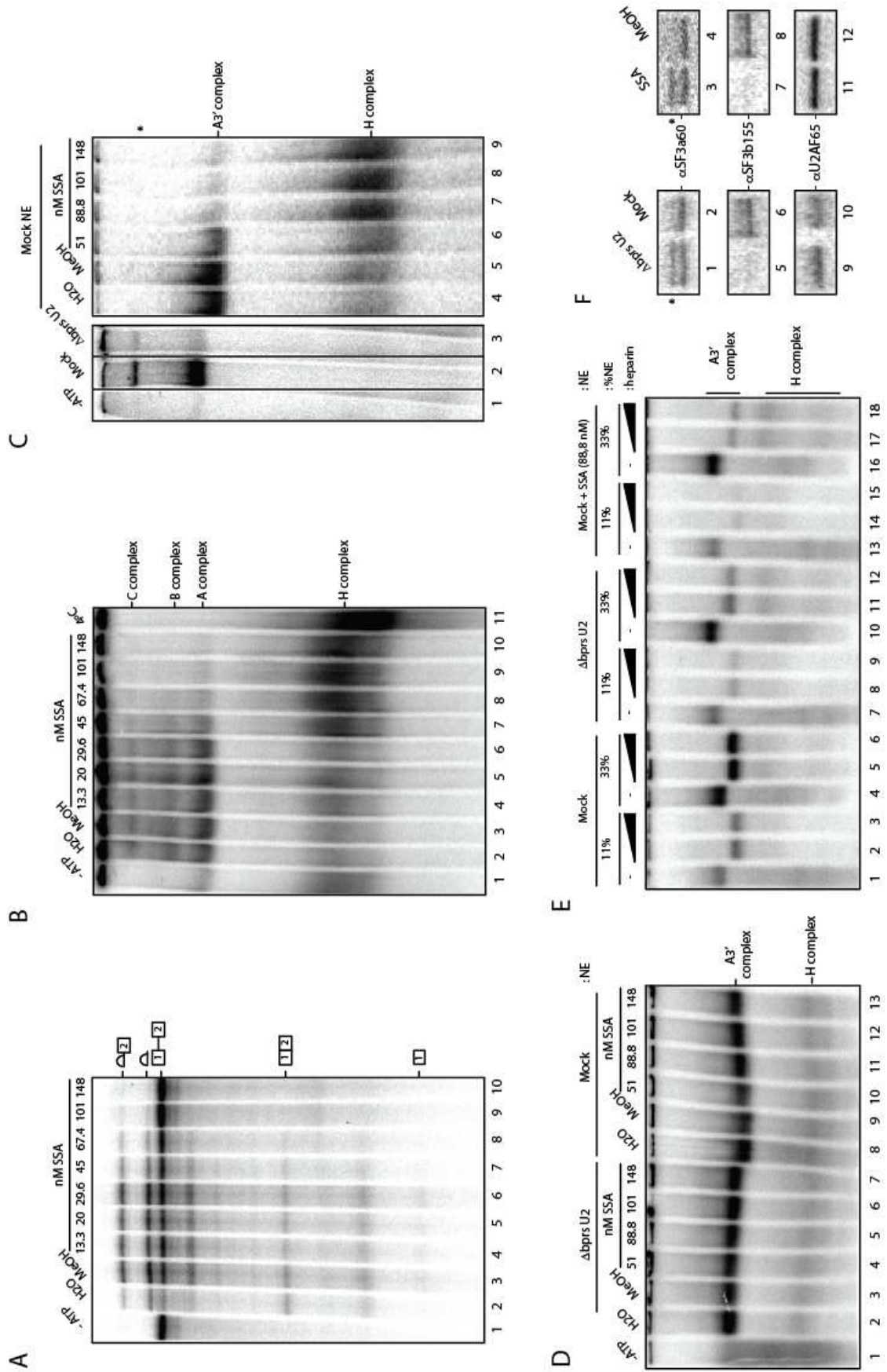
Fig. I. 1

Fig. I. 1. SSA inhibits stable U2 snRNP recruitment. (A) In vitro splicing assays using adenovirus major late (AdML) RNA in the presence of the indicated concentrations of SSA or its carrier (methanol), analyzed by denaturing polyacrylamide gel electrophoresis. Pre-mRNA, splicing intermediates and products are indicated. **(B)** Spliceosome assembly assays corresponding to in vitro splicing mixes as in (A) in the presence of heparin (5 mg/ml) analyzed by native agarose-polyacrylamide gel electrophoresis. The positions of heterogeneous nuclear RNP (H) and spliceosomal A, B and C complexes are shown. **(C)** Spliceosome assembly assays as in (B) using AdML 3' RNA, containing the 3' 40 nucleotides of AdML intron 1 and exon 2. The positions of the A3' and H complexes are indicated. **(D)** Spliceosome assembly assays as in (C) in the absence of heparin, analyzed by low melting-agarose gel electrophoresis, using either mock-treated or nuclear extracts in which the bprs of U2 snRNA was inactivated by RNase H-mediated degradation (ΔbprsU2). **(E)** Spliceosome assembly assays as in (D) in the presence of different concentrations of the indicated nuclear extracts with or without SSA, and in the absence or presence (1.67 and 5 mg/ml) of heparin. **(F)** Ultraviolet light-induced crosslinking followed by immunoprecipitation of SF3b155, SF3a60 and U2AF65 in the presence of SSA (88 nM) or methanol, incubated with mock- or ΔbprsU2 nuclear extracts. Asterisk indicates a crosslinked protein co-immunoprecipitated with the SF3a60 antibody, possibly SF3a66 (Brosi et al., 1993a).

Fig. I. 2
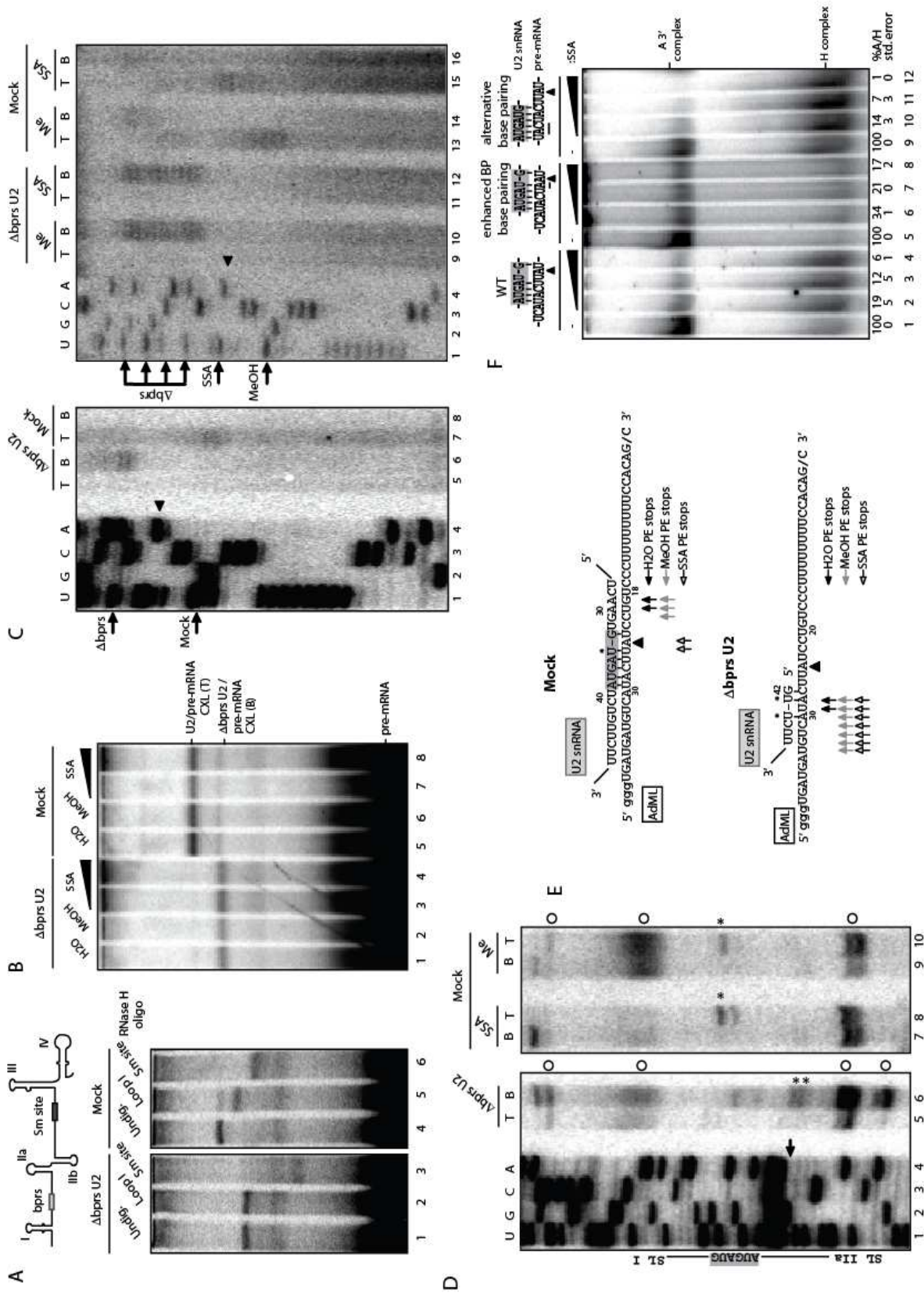
Fig. I. 2.SSA and U2 bprs inactivation lead to altered U2 snRNA/pre-mRNA base pairing interactions.(A) Schematic representation of U2 snRNA secondary structure domains (top) and mapping of U2 snRNA/pre-mRNA interactions by RNase H-mediated degradation of psoralen-induced crosslinks between radioactively labeled AdML 3' RNA and either U2 snRNA (Mock) or ΔbprsU2 snRNA, analyzed by denaturing gel electrophoresis. As expected, the position of the crosslinked RNA is affected by RNAse H-mediated degradation of U2 snRNA at sequences 5' (Loop I) or 3' (Sm site) of the bprs in Mock extracts, while crosslinking of the ΔbprsU2 snRNA occurs 3' of the bprs, as indicated by altered electrophoretic mobility of the RNA-RNA crosslink upon degradation of the 3' Sm site but not of the 5' Loop I. **(B)** SSA does not inhibit U2 snRNA base pairing. Psoralen-induced crosslinking was analyzed as in (A) using either mock- or Δbprs U2 snRNA nuclear extracts, in the presence or absence of SSA (53, 267 nM) or methanol. The positions of crosslinked adducts (CXL) formed with wild type and ΔbprsU2 are labeled as T (top) and B (bottom), respectively. **(C)** Mapping of U2 snRNA-AdML crosslinks by primer extension. U2 snRNA-AdML adducts from either mock or ΔbprsU2 nuclear extracts and from methanol- or SSA (88 nM)-treated nuclear extracts were gel extracted and purified. To assess the presence of background levels of U2 snRNA in the lanes, gel slices corresponding to top (T) and bottom (B) positions defined as in (B) were analyzed in parallel for each sample. Primer extension was carried out using a 32P-labeled oligonucleotide complementary to AdML exon 2. Lanes 1 to 4 correspond to dideoxy sequencing reactions of AdML DNA using the same primer. Arrows indicate primer extension stops. Triangles indicate the position of the branch point adenosine. Circles indicate non-specific stops observed in primer extension of U2 snRNA in the absence of psoralen crosslinking. **(D)** Mapping of U2 snRNA crosslinked residues by primer extension as in (C) using a 32P-labeled oligonucleotide primer complementary to the Sm site sequence in U2 snRNA. Lanes 1 to 4 correspond to dideoxy sequencing of U2 snRNA using the same primer. The bprs (5'GUAGUA3') and boundaries of the flanking stem-loops I and IIa are indicated. Asterisks indicate specific primer extension stops. **(E)** Schematic representation of base pairing interactions between U2 snRNA and the AdML branch point region in mock- or ΔbprsU2 nuclear extracts, in the presence of SSA, methanol or water. Asterisks indicate U2 snRNA crosslinked nucleotides as mapped in (D). Triangles indicate the branch point adenosine and residues in the grey box indicate the bprs. Arrows summarize primer extension (PE) stops observed in (C) using different conditions as indicated. **(F)** Spliceosome assembly assays (as in Fig. I. 1C) using AdML wild type and mutants harboring either a consensus branch point (enhanced BP base pairing) or a sequence located 5' of the branch point region which is fully complementary to U2 snRNA bprs (alternative base pairing). Mutated residues are underlined. Quantification of the A/H complex ratio for a minimum of three independent experiments (with standard deviation) is shown. The ratio in methanol-treated samples was considered as 100% in each case.

Fig. I. 3
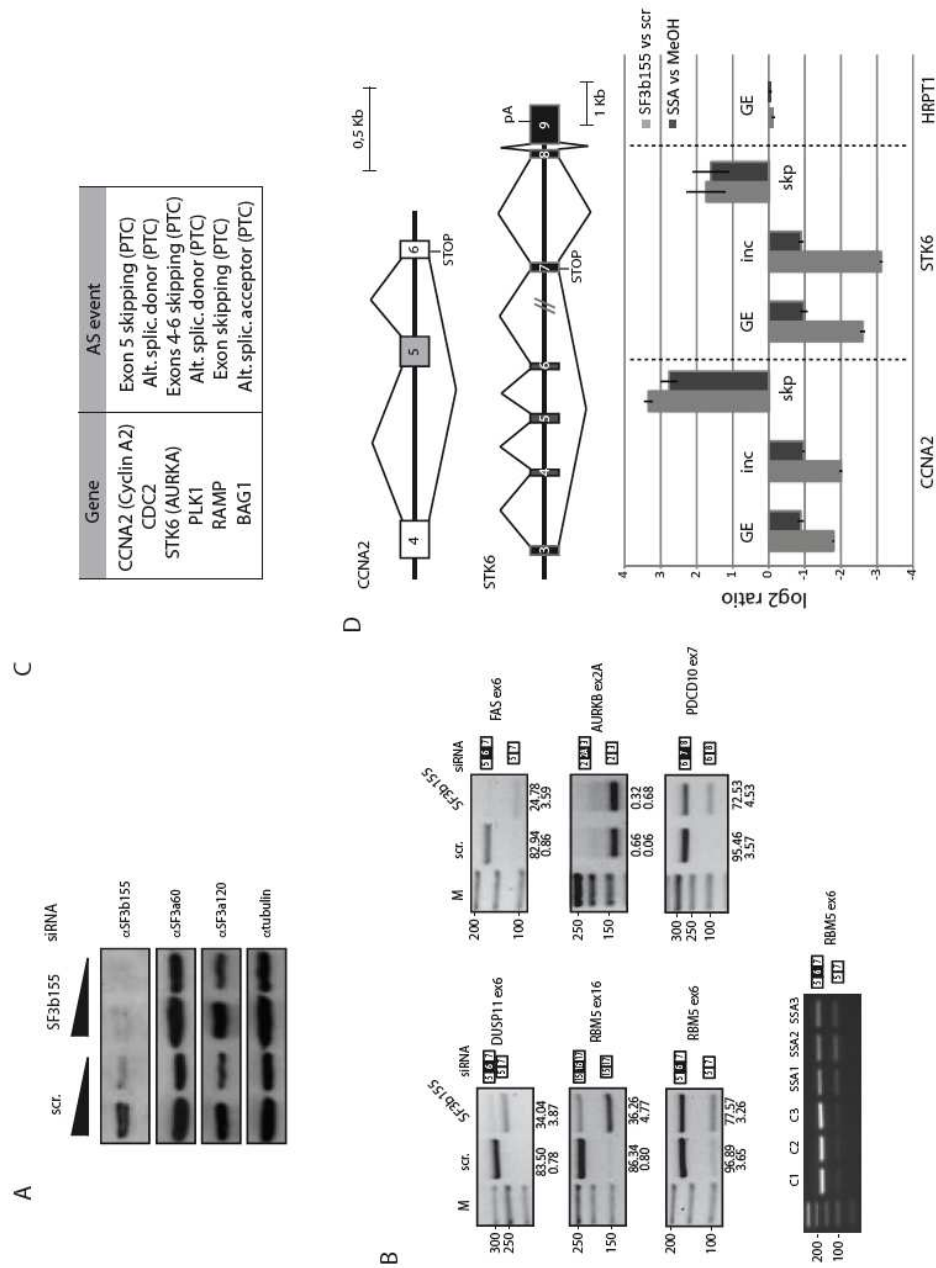
Fig. I. 3. **SSA induces alternative splicing and expression changes in genes important for cell cycle control.(A)** Western blot analysis of the effects of SF3b155 RNAi on the levels of SF3b155 and other protein components of U2 snRNP (as well as α-tubulin as loading control) demonstrate specific depletion of SF3b155. **(B)** Semi-quantitative RT-PCR analysis of alternative splicing changes upon SF3b155 knock down from a selection of alternative splicing events induced by SSA. Genes and alternative splicing events are indicated, as well as the position of the alternatively spliced products. The numbers below each lane indicate percentage of exon inclusion (average and standard deviation for a minimum of three indpendent experiments). Alternative splicing changes common to those induced by SSA are observed for all except AURKB, which displays essentially quantitative exon skipping. The lower panel shows changes in RBM5 exon 6 induced by SSA treatment: C1-3 show results from three control samples, while SSA1-3 correspond to three independent RNA samples from

cells treated with SSA. **(C)** Genes encoding important cell cycle regulators whose expression is downregulated by SSA and by knock down of SF3b155 using two independent siRNAs. Downregulation is coupled to the generation of alternatively spliced isoforms containing premature stop codons (PTC) and consequently nonsense-mediated-decay (AS-NMD). **(D)** (top) Schematic representation of alternative splicing events leading to NMD in the cyclin A2 (CCNA2) and Aurora Kinase A (STK6) genes, activated both by SSA treatment and by SF3b155 knock down; indicated are PTCs associated with the alternative splicing patterns leading to NMD; (bottom) quantitative RT-PCR analysis using primers that monitor general gene expression in a constitutive exon (GE), or primers specific for splice junctions corresponding to exon inclusion (inc) or skipping (skp). Also shown is the absence of effects of the treatments on expression of a housekeeping gene (HRPT1).
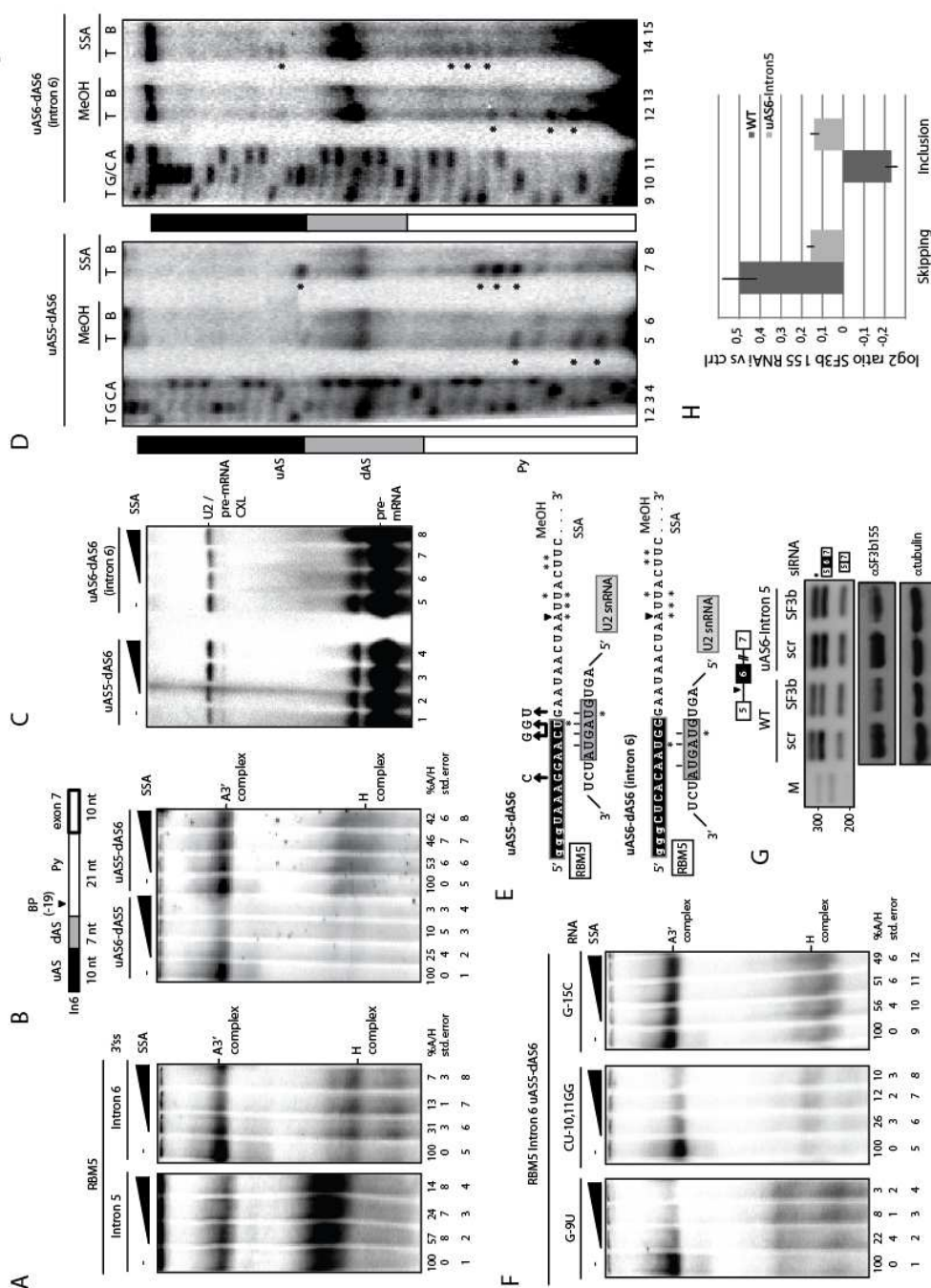
Fig. 4. Base pairing potential with U2 snRNA of sequences 5' of the branch point can modulate SSA-induced alternative splicing.(A) Spliceosome assembly assays (as in Fig. 2F) were carried out using RNAs corresponding to RBM5 introns 5 and 6 3' splice site regions. **(B)** Spliceosome assembly assays (as in A) using chimeric RBM5 intron 6 RNAs in which the indicated sequences have been replaced by the corresponding fragments of intron 5. uAS and dAS indicate upstream and downstream regions of the Anchorage Site (AS), BP and Py indicate branch point and polypyrimidine tract, respectively. The length of the different regions of the RNAs utilized is also indicated. **(C)** Psoralen crosslinking assays (as in Fig. I. 2B) using the indicated RNAs. **(D)** Mapping of U2 snRNA/RBM5 intron 6 psoralen-induced crosslinks by primer extension in the absence or presence of SSA, as in Fig. I. 2C. Lanes 1 to 4 and 9 to 11 correspond to dideoxy sequencing reactions of the corresponding DNA products using the same primer. Asterisks indicate primer extension stops. Black and white boxes designate the different segments of the 3' splice site region as in B. **(E)** Schematic representation of potential U2 snRNA/RBM5 pre-mRNA base pairing interactions induced by SSA. Asterisks mark the crosslinked nucleotides mapped in (C) in methanol- and SSA-treated samples. Nucleotides in the grey box represent the U2 snRNA bprs. **(F)** Spliceosome assembly assays (as in A) using RNAs harboring mutations (also indicated in E) in the residues proposed to base pair with U2 snRNA bprs (or control residues) of the uAS5- dAS6 chimeric RNA. **(G)** RT-PCR analysis of RBM5 RNA isoforms expressed from plasmids expressing the alternatively spliced RBM5 region (exons 5 to 7), either wild type or a mutant that replaces the upstream region of intron 5 anchoring site by the corresponding sequence from intron 6 (uAS6-intron 5) in control cells or in cells in which expression of SF3b155 has been knocked down. The asterisk indicates an aberrant amplification product caused by primer hybridization in intron 5, as determined by sequencing. The lower panel shows western blot analysis of the levels of SF3b155 and α-tubulin in control and SF3b155 knock down cells. **(H)** Analysis of RBM5 exon 6 skipping and inclusion of RNAs analyzed in G in SF3b155 knock down vs control cells by quantitative RT-PCR, using splice junction primers specific for exon inclusion or skipping as one of the oligonucleotides utilized in the PCR amplification reactions.

Materials and Methods

In vitro splicing assays

Splicing assays were performed in 9 µl mixtures containing 3 µl of HeLa nuclear extracts (CilBiotech, Belgium) and 10 fmol of *in vitro* transcribed 32P-labeled Adenovirus-Major-Late pre-mRNA under standard conditions (3 mM MgCl2, 1.1 mM ATP, 22 mM creatine phosphate and 1.67% polyvinyl alcohol). After incubation for 1h at 30ºC, RNA was isolated by treatment with proteinase K and ethanol precipitation. Pre-mRNA precursor, splicing intermediates and products were resolved by electrophoresis on denaturing 13% polyacrylamide gels and analyzed by Phosphorimager.

Spliceosome assembly

Splicing mixtures were assembled as above using 32P-uridine labeled pre-mRNA or fragments corresponding to the 3' splice site region and downstream exon. After 15 min incubation at 30ºC, 5 µg/ul of heparin were added to the mix, kept for 10 min at room temperature and resolved by electrophoresis on 4 % acrylamide:bisacrylamide (80:1)-0,5 % agarose gels in 50 mM Tris base – 50 mM glycine buffer. The gels were dried and exposed to a Phosphoimager screen. Alternatively, 2.2 µl of 6x DNA loading dye (20 mM Tris-HCl pH 7.5, 0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol) were added to the reaction directly after incubation at 30ºC and mixtures were then resolved by electrophoresis in 1.5% low melting agarose gels (Ecogen, Cat.# AG-420) in 50 mM Tris base – 50 mM glycine buffer. The gels were fixed in 10% methanol – 10% acetic acid for 10' at room temperature, dried and exposed to a Phosphoimager screen.

UV crosslinking and immunoprecipitation

32P-labeled RNAs were incubated under in vitro splicing conditions in a final volume of 27 µl. Mixtures were incubated for 30 min at 30ºC and irradiated with UV light (254 nm; 0.4 J). After further incubation with 1 mg/ml RNase A at 37ºC for 30 min, immunoprecipitation of U2AF65 and U2 snRNP components was carried out by addition of anti-U2AF65 MC3 monoclonal antibody hybridoma supernatant (Gama-Carvalho et al., 1997, J. Cell Biol. 137: 975-987) or anti-SF3b155 or anti-SF3a 60 polyclonal antibodies (kind gifts of Prof. Angela Krämer, University of Geneva). After incubating for 60 min at 4ºC, 30 µl of a 50% slurry of protein A/G sepharose beads (GE Healthcare, Cat.#17-5280-01 and Cat.#17-0618-01, respectively) were added and the mixture was incubated under rotation for another 60 min. Beads were sedimented by centrifugation and washed 4 times with 800 µl of high salt buffer (500 mM NaCl, 50 mM Tris-HCl pH 8.0, 1% NP-40) and once with the same buffer except that the concentration of NaCl was 150 mM. Sedimented beads were resuspended in 4xSDS

loading dye, boiled and the supernatant was resolved by electrophoresis on 8-10% SDS polyacrylamide gels. Gels were dried and exposed to a Phosphoimager screen.

Psoralen-mediated RNA-RNA crosslinking

Fifteen-microliter reactions containing 80 fmol of 32P-uridine labeled 3' half RNAs were incubated at 30ºC for 20' in the presence of 5 µl HeLa nuclear extracts and 22 µg/ml of 4'aminomethyltrioxsalen hydrochloride (Sigma, Cat.#A4330). The mixture was irradiated with 365 nm wavelength UV light for 10 min. RNA was then treated with proteinase K, ethanol precipitated and analyzed by electrophoresis on a denaturing 6% polyacrylamide gel. RNase H (Invitrogen, Cat.#18021-014)-mediated inactivation of U2 snRNP was performed by incubating the nuclear extracts with a oligodeoxynucleotide complementary to positions 28 to 42 of U2 snRNA. RNase H–mediated digestion after crosslinking was performed similarly using the specified oligodeoxynucleotides.

Mapping of psoralen-crosslinked nucleotides by primer extension

Psoralen-mediated crosslinked RNAs and splicing lariat intermediates were gel purified and primer extension was performed using the indicated 5' 32P-labeled oligodeoxynucleotide and Superscript II (Invitrogen, Cat.#18064014) following the manufacturer's protocol. Primer extension products were resolved by electrophoresis on denaturing polyacrylamide gels, in parallel with a sequencing reaction performed using the same primer and *fmol* DNA Cycle Sequencing System (Promega, Cat.#Q4100).

Spliceostatin A treatments

A. Nuclear extract. Nuclear extracts (3 µl for splicing reactions and 5 µl for psoralen crosslinking assays) were treated for 30' at 4ºC with SSA previous to their use in subsequent assays. The indicated amounts correspond to the final concentration in the reaction mixtures.

B. Tissue culture. Sub-confluentHeLa cells were treated with Spliceostatin A (at a final concentration of 100 ng/ml -260 nM-) for 3 hours. Control cells were treated with methanol, the solvent in which the SSA stock was dissolved.

RNA isolation

Total RNA was isolated using the Qiagen RNeasy mini kit (Cat. No. 74104), with DNase digestion on column before elution using the RNase-free DNase Set (Cat.#79254). RNA quality was assessed by Agilent Bioanalyzer nano assay and nanodrop spectrophotometry.

Microarray analyses

Cy3 and Cy5 labeled cRNAs were generated from 500 ng of total RNA using the Quick-Amp Labeling Kit, No Dye (Cat.#5190-0447) and Cy3 and Cy5-CTP from Perkin-Elmer (NEL580 and NEL581). 7µg of each cRNA were hybridized into our custom oligo microarray platform (Agilent) in an Agilent rotating oven at 60ºC for 18 hours. After hybridization, arrays

were washed, and scanned images analyzed. Statistical analyses were carried out with LIMMA package (Bioconductor) as previously described (Munoz et al., 2009). Microarray results have been deposited in GEO database (GSE22614 and GSE 22952). Microarray data as well as a detailed explanation of data analysis using this platform can be anonymously downloaded from the following web site: http://davinci.crg.es/splicingarrays/ ; User: jvlab; Password: Web01.

Real time qPCR analyses

Primers were designed using Primer 3 Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) using the reference and the alternative RefSeq Accession numbers. Primers used are listed in Supplementary Table 1. The primers were designed such that their annealing temperature was 60ºC, generating single amplification products in the range of 60 - 200 base pairs long. For analysis of alternatively spliced isoforms, one of the primers was chosen as overlapping an exon junction such that only upon hybridization to the correctly spliced mRNA the primer was able to produce an amplification product. First strand cDNA synthesis was carried out using 1 mg of RNA, 50 pmoles of oligo-dT (Sigma-Aldrich), 75 ng of random hexamers (Roche 11034731001) and Reverse Transcriptase Superscript II (Invitrogen, Cat. No. 18064-014) following the manufacturer's instructions. PCR amplification was carried out with 1ml of the 1:10 diluted reverse transcription sample with 5 ml of 2x SYBR Green Master Mix (ROCHE, Cat. No. 4309155), and 4 pmol of specific gene primer pairs to a 10 ml total volume, in 384-well microtiter plates. PCR reactions were run in triplicates on a Lightcycler 480® system (ROCHE). Relative copy number (RCC) or fold-change ratio was calculated according to (Pfaffl, 2001). A list of the primers utilized for RT-PCR analyses using this platform can be anonymously downloaded from the following web site: http://davinci.crg.es/splicingarrays/ ; User: jvlab; Password: Web01.

Semi-quantitative RT-PCR

Total RNA was isolated 72 hours after SF3b155 siRNAs transfection using the GE Healthcare RNAspin Mini kit (Cat.#25-0500-72) and analyzed by RT-PCR. 1 µg of RNA was reverse-transcribed using random primers and Superscript II (Invitrogen, Cat.#18064014). One-fourth of the cDNA was amplified by PCR using appropriate primers. After 24 cycles of 1' at 93℃, 45" at 56℃ and 1' at 72℃, the products w ere analyzed on 2% agarose gels.

RNAi

To knock down SF3b155, cells were transfected with Lipofectamine 2000 following the manufacturer's protocol with three different siRNAs against SF3b155 (Invitrogen):

siRNA #3: 5'-GACAGCAGAUUUGCUGGAUACGUGA-3'

siRNA #4: 5'-GCUUGGUCAGAAGAAGCCAGGAUAU-3'

siRNA #5: 5'-CCCUGUGGCAUUGCUUAAUGAUAU-3').

siRNA #4 was used in initial knock down experiments, siRNAs #3 and 5, which achieved higher levels of knock down, were used for independent microarray analyses.

U2AF depletion and complementation with recombinant GST-U2AF65

U2AF depletion by chromatography on oligo-dT cellulose and complementation assays were carried out as previously described (Valcarcel and Gebauer, 1997; Guth et al., 2001).

Supplementary Figures and Figure Legends

Fig. I. S1
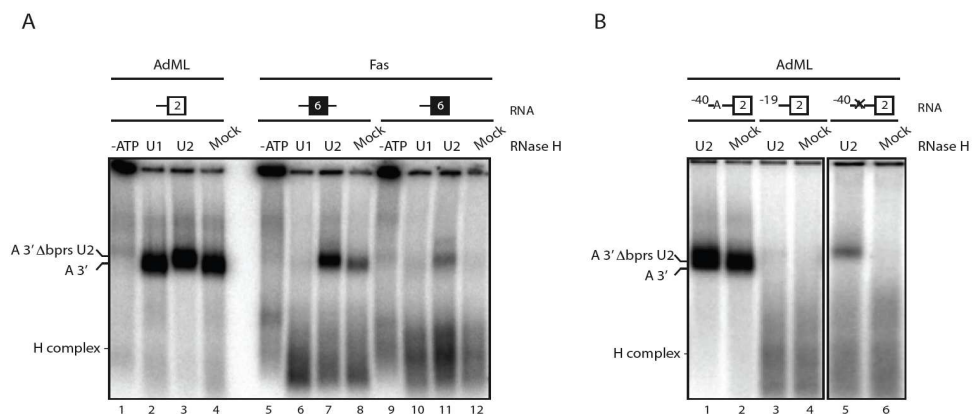
Fig. I. S1. Δbprs U2 complex is ATP- and branch point-dependent and is sensitive to exon definition effects. (A) ATP dependence and exon definition effects.Spliceosome assembly assays were carried out as in Fig. 1D using nuclear extracts in the presence or absence of ATP / ATP regenerating system or in extracts in which either U1 or U2 were inactivated by RNase H-mediated degradation of the 5' end of U1 snRNA or the bprs of U2 snRNA, respectively (Merendino et al., 1999). Tested RNAs included AdML (3' half of the intron + exon 2) and Fas intron 5 (3' 68 nucleotides + exon 6, followed or not by the 5' 25 nucleotides of intron 6). The result shows that the ATP-dependent complex formed on AdML RNA was not sensitive to U1 inactivation, while the complex formed on Fas exon 6 and flanking sequences was dependent on the integrity of U1 snRNP as well as on the presence of intron 6 5' splice site region, as expected from the strong dependence on exon definition effects of complex A formation in this transcript (Izquierdo et al., 2005). In contrast, U2 snRNA bprs inactivation altered the mobility but did not reduce (even increased) complex formation. **(B)** Spliceosome assembly assays as in (A) using wild type AdML (-40), a mutant in which the branch point region was deleted (-19) or a mutant in which the branch point region (ACUU<u>A</u>U, where the underlined adenosine is the branch site) was mutated to GTCCTC (-40 X). The result indicates that formation of both the U2 and U2 Dbprs complexes requires a functional branch point region.

Fig. I. S2
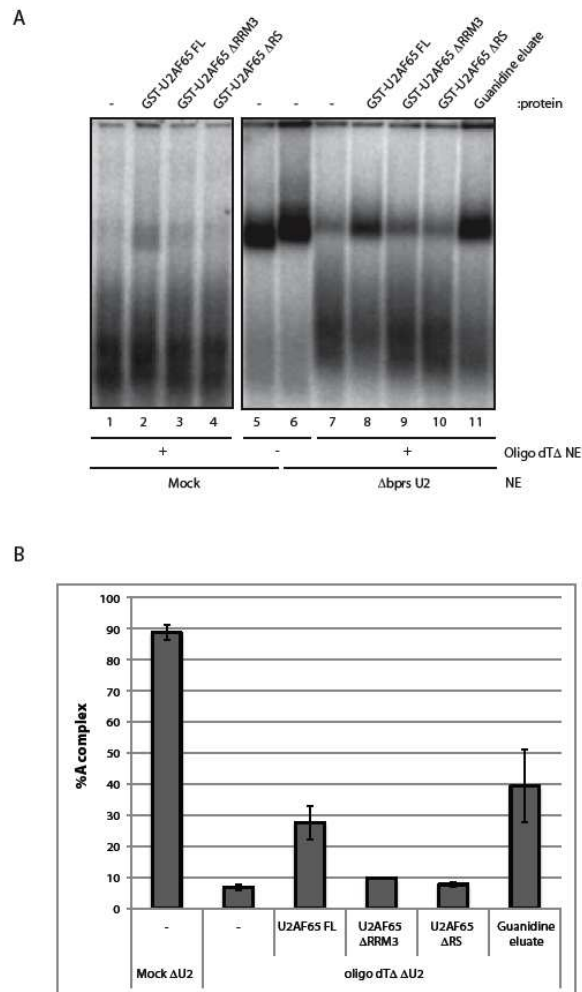
Fig. I. S2. Δbprs U2 complex formation requires full length U2AF65. (A) Spliceosome assembly assays were carried out as in Fig. 1D using Mock- or U2-inactivated (Δbprs) nuclear extracts and either mock-depleted (-) or U2AF-depleted (by chromatography in oligodT-cellulose; Oligo dTD NE, (Guth et al., 1999) nuclear extracts, as indicated. Depleted extracts were complemented using 75 ng/ml of recombinant full length GST U2AF65 full length (FL), GST-U2AF65 lacking RNA Recognition Motif 3 (ΔRRM3, lacking aminoacids 347–475), GST-U2AF65 lacking the arginine-serine rich domain (ΔRS, lacking aminoacids 1-94) or the guanidine-HCl eluate from the oligo dT column used for U2AF depletion (which contains U2AF65 and U2AF35). The concentrations of the different recombinant proteins were adjusted to use protein preparations with equivalent specific activity in RNA binding. **(B)** Quantification of A versus H complex formation in Δbprs U2 U2AF65-depleted extracts complemented with recombinant U2AF65. Error bars represent standard deviations. The results indicate that U2AF is required for formation of the U2 Δbprs complex (compare lanes 6 vs 7 in part A) as is the case for the complex formed with intact U2 snRNA (compare lanes 1 vs 5). Addition of recombinant full length U2AF65 provides some splicing activity to U2AF-depleted extracts (compare lanes 7 and 8), while U2AF65 lacking the RS or RRM3 domains do not (compare lanes 7 with 9 and 10). We conclude that recruitment of Δbprs U2 snRNP is dependent on U2AF activity, as is the case for recruitment of wild type U2 snRNP (lanes 1-4).

Fig. I. S3
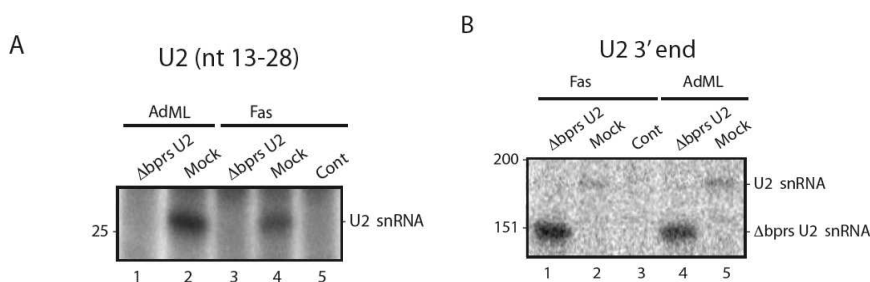
Fig. I. S3. U2 snRNA, but not the U2 snRNA 5' region, is detected in Δbprs U2 complex.
Spliceosome assembly assays were performed as in Figs. 1D and Sig. S1 on AdML or Fas RNAs. After electrophoresis, complexes were extracted from the gel, the RNA was purified and the presence of U2 snRNA in the complex analyzed by primer extension using 5'-32P-labeled oligonucleotides complementary to U2 snRNA nucleotides 13-28 **(A)** (expected length of the primer extension product = 28 nucleotides) or 171-187 **(B)** (expected lengths of primer extension products = 187 nucleotides for full-length U2 snRNA and 145 nucleotides for Δbprs U2 snRNA). The results indicate that U2 snRNA is readily detectable in the complexes using the primer complementary to the 3' end of U2 snRNA (B). Note that the different size of the primer extension product in the Δbprs and Mock-treated U2 extracts is due to RNAse H-mediated degradation of the bprs (nucleotides 28-42) in the Δbprs extracts. In contrast, no primer extension signal is detected for the complex formed in Δbprs U2 extracts using the primer complementary to the 5' end of U2 snRNA, indicating that this region of the snRNA is not present in the complex formed after degradation of the bprs.

Fig. I. S4
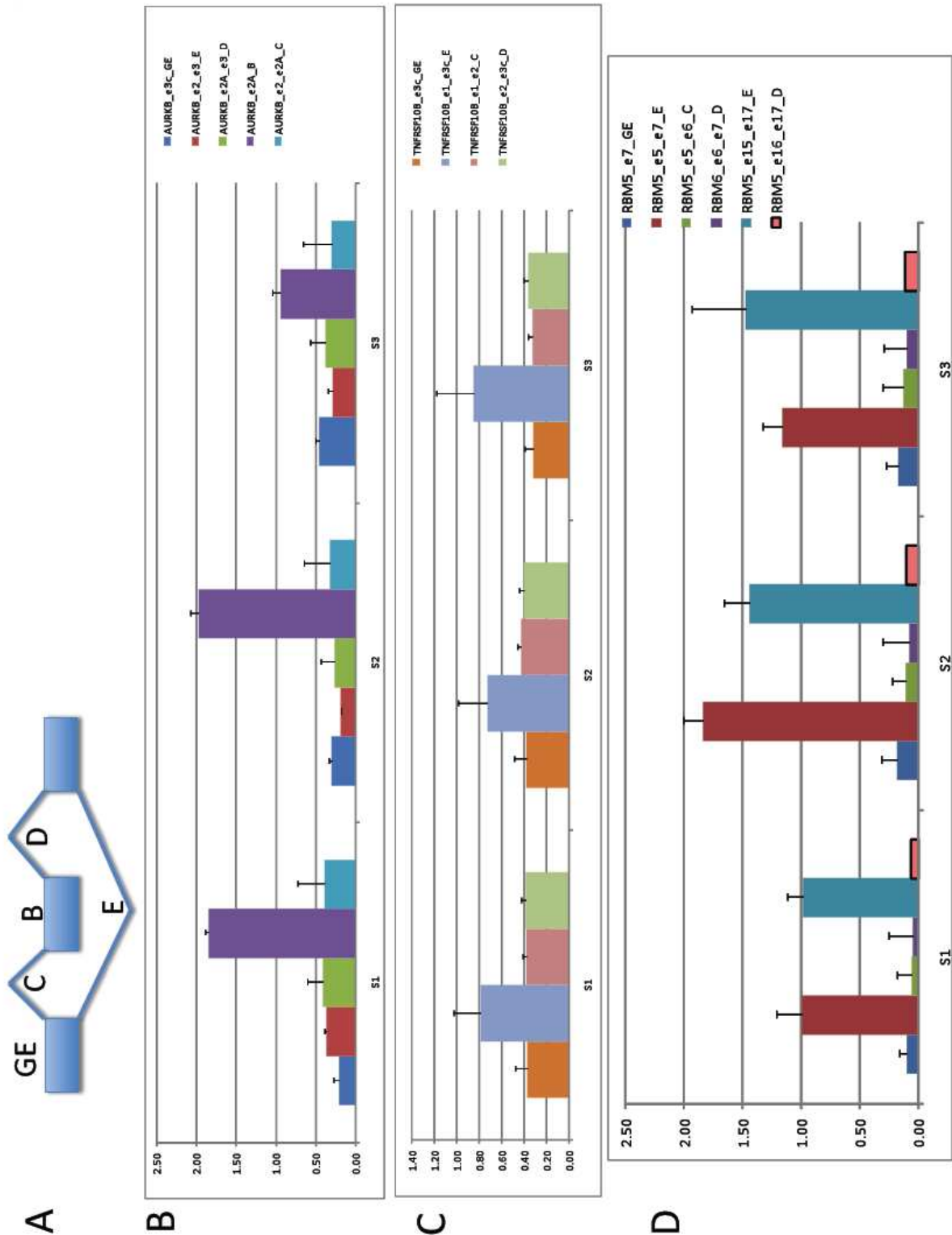
Fig. I. S4. Examples of microarray-predicted alternative splicing events validated by quantitative RT-PCR. (A) Scheme of a typical alternative splicing cassette exon event, indicating the nomenclature of the probes present in the array to detect changes in expression of alternatively spliced isoforms, which are compared with gene expression values provided by constitutive exons. Validation by quantitative PCR is accomplished using oligonucleotides corresponding to the alternatively spliced region (B) or to alternative splice junctions (C,D,E). **(B)** Detection of increased inclusion of exon 2A in

the gene AURKB, detected using exon 2A-specific oligonucleotides. The y axis indicates expression ratios between RNAs from methanol-treated and SSA-treated cells. The x axis represents the nature of the amplification products using oligonucleotides indicated in the color code on the right. For example, AURKB_e3c_GE indicates quantification of constitutibe exon e3 to determine changes in general expression; AURKB_e2_e3_E indicates quantification of the spliced product joining exons 2 and 3 (probe E in the scheme in A) measuring exon 2A skipping; AURKB_e2A-e3-D indicates quantification of the spliced product between exon 2A and exon 3 (probe D in the scheme in A) measuring exon 2A inclusion; etc. **(C)** Detection of increased skipping of exon 2 in the gene TNFRSF10B, detected using an oligonucleotide complementary to the junction between exons 1 and 3 as one of the PCR primers. **(D)** Detection of increased skipping of exons 6 and 16 in the gene RBM5, detected using oligonucleotides complementary to the junctions between exons 5 and 7, or exons 15 and 17, as one of the primers in the amplification reaction.

Fig. I. S5
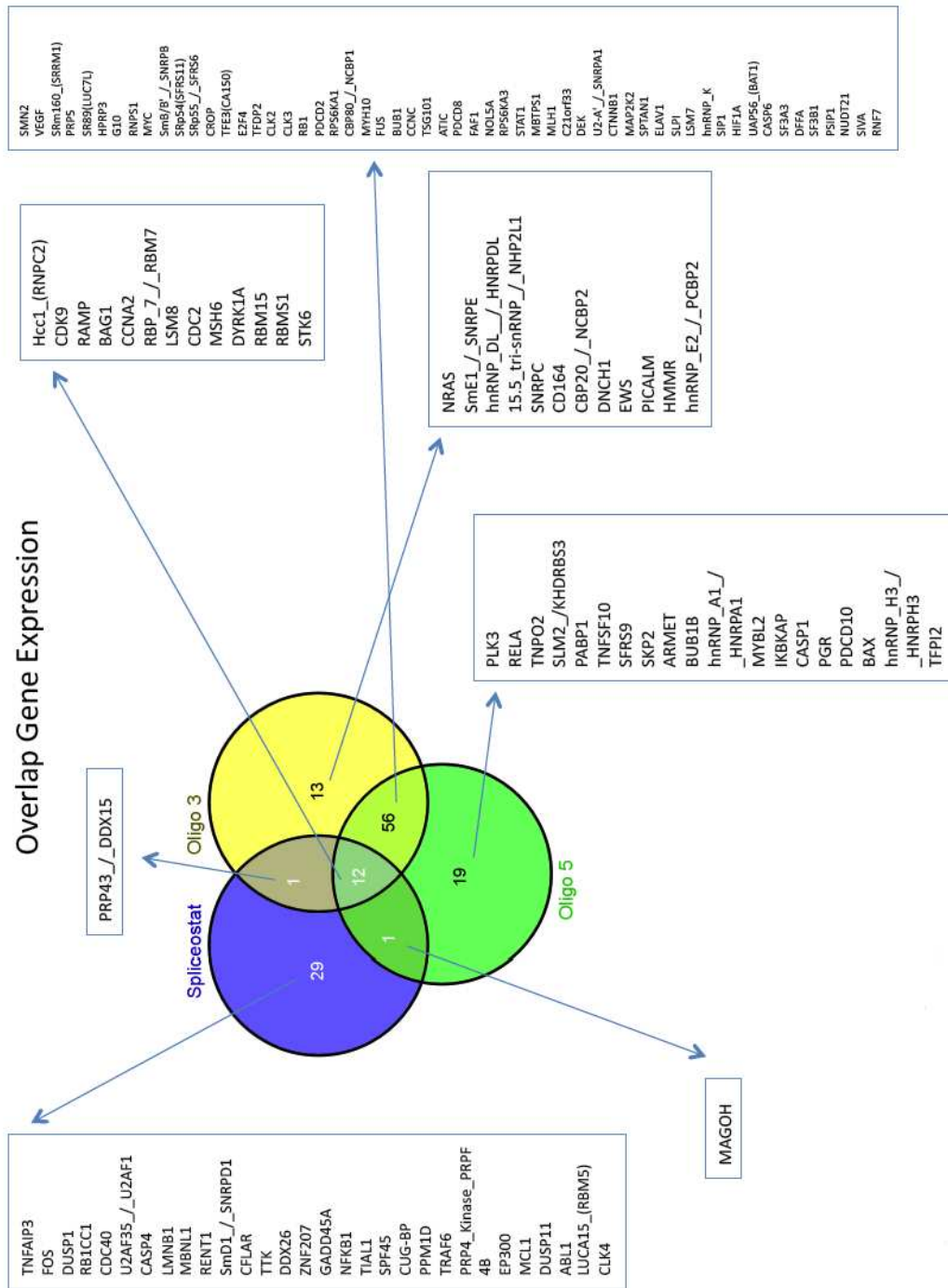
Fig. I. S5. Microarray-detected changes in gene expression upon SSA treatment and upon SF3b155 knock down using two different siRNAs. The diagram represents the overlap of genes displaying gene expression changes of at least 1.3 fold between methanol- or SSA-treated cells and between cells treated with control scrambled siRNAs or treated with two different siRNAs corresponding to SF3b155 (oligos 3 and 5). The overlap between genes affected by SF3b155 knock down using the two siRNAs is 80% (Fischer exact test p value < 0.0001). The overlap between genes affected by SSA and by either of the two siRNAs is 32% (Fischer exact test p value = 0.026), and the overlap with genes affected by both siRNAs is 28% (Fischer exact test p value = 0.008).

Fig. I. S6

Overlap Alternative Splicing Changes (in Genes without Gene Expression Changes)

Venn diagram: Oligo 3 = 53, Oligo 5 = 43, Spliceostat = 119; overlaps: Oligo 3 ∩ Oligo 5 = 190, Oligo 3 ∩ Oligo 5 ∩ Spliceostat = 31, Oligo 5 ∩ Spliceostat = 8, Oligo 3 ∩ Spliceostat = 12.

Common elements in "Oligo 3", "Oligo 5" and "Spliceostat":

ID	Gene	Splice type
96.4.1	15.5_tri-snRNP / NHP2L1	novel_exons
38.6.1	SFRS2	alt_splice_acceptor
00234.0002.1	RELA	novel_exons
00188.0060.1	NFKBIA	exon(s)skipped
123.2.1	9G8 / SFRS7	exon(s)skipped
123.3.1	9G8 / SFRS7	exon(s)skipped
123.4.1	9G8 / SFRS7	exon(s)skipped
123.5.1	9G8 / SFRS7	novel_exons
39.4.1	SRp20 / SFRS3	novel_exons
39.5.1	SRp20 / SFRS3	novel_exons
00087.0023.1	DAXX	alt_splice_acceptor
00214.0035.1	PLK1	alt_splice_donor
186.1.1	NOSIP	novel_exons
36.2.1	RENT1	alt_splice_acceptor
00016.0067.1	AURKB	exon(s)skipped
00016.0033.1	AURKB	novel_exons
00016.0058.1	AURKB	alt_splice_donor
00208.0011.1	PDCD6IP	exon(s)skipped
00457.0019.1	CDKN1A	alt_splice_donor
73.1_10:3	C1QBP	novel_exons
140.1.1	U2AF35 / U2AF1	novel_exons
140.3.1	U2AF35 / U2AF1	novel_exons
00263.0017.1	TNFRSF10B	exon(s)skipped
00123.0013.1	HRAS	exon(s)skipped
00312.0001.1	BUB1B	exon(s)skipped
00312.0011.2	BUB1B	novel_exons
00049.0003.1	CASP9	exon(s)skipped
00462.0029.1	MAP2K5	exon(s)skipped
00033.0008.2	BRCA1	exon(s)skipped
00033.0014.1	BRCA1	exon(s)skipped
232.1.1	hnRNP_E2 / PCBP2	alt_splice_acceptor

Common elements in "Oligo 5" and "Spliceostat":

ID	Gene	Splice type
00016.0038.2	AURKB	alt_splice_acceptor
00231.0003.1	RB1CC1	alt_splice_acceptor
00253.0005.1	TCL1A	alt_splice_acceptor
00253.0009.1	TCL1A	exon(s)skipped
00279.0011.3	ZNF198	exon(s)skipped
88.1.1	SRp54(SFRS11)	exon(s)skipped
88.2.1	SRp54(SFRS11)	novel_exons
116:1_4:1	CDC40	novel_exons

Common elements in "Oligo 3" and "Spliceostat":

ID	Gene	Splice type
70.10.1	CLK2	exon(s)skipped
70.4.1	CLK2	novel_exons
00116.0010.1	GRB2	exon(s)skipped
32.1.1	PPP1R8	alt_splice_donor
00249.0100.1	STAT3	exon(s)skipped
74.2.1	p68 / DDX5	novel_exons
00203.0044.1	PDCD10	exon(s)skipped
00203.0076.2	PDCD10	alt_splice_acceptor
00203.0114.1	PDCD10	exon(s)skipped
00203.0016.2	PDCD10	alt_splice_donor
00203.0009.2	PDCD10	alt_splice_donor
158.5.1	hnRNP_H3 / HNRPH3	exon(s)skipped

Fig. I. S6. Microarray-detected changes in alternative splicing in genes that do not display gene expression changes upon SSA treatment and upon SF3b155 knock down using two different siRNAs. The diagram represents the overlap of genes displaying changes in expression of an alternatively spliced isoform of at least 1.4 fold between methanol- or SSA-treated cells and between cells treated with control scrambled siRNAs or treated with two different siRNAs corresponding to SF3b155 (oligos 3 and 5). The overlap between genes affected by SF3b155 knock down using the two siRNAs is 81% (Fischer exact test p value < 0.0001). The overlap between genes affected by SSA and by either of the two siRNAs is 30% (Fischer exact test p value < 0.001), and the overlap with genes affected by both siRNAs is 18% (Fischer exact test p value = 0.005).

Fig. I. S7

Overlap Alternative Splicing
Changes (in Genes with Gene
Expression Changes)

Common elements in "Oligo 3", "Oligo 5" and "Spliceostat":		
188.3.1	RBP_7 / RBM7	exon(s)skipped
00313.0021.1	CCNA2	exon(s)skipped
410:1_2:21	LSM8	novel_exons
00015.0024.2	STK6	exon(s)skipped
648:1_3:6	RBMS1	novel_exons

Common elements in "Oligo 5" and "Spliceostat":		
00383.0016.1	BAG1	alt_splice_acceptor
00015.0001.1	STK6	novel_exons

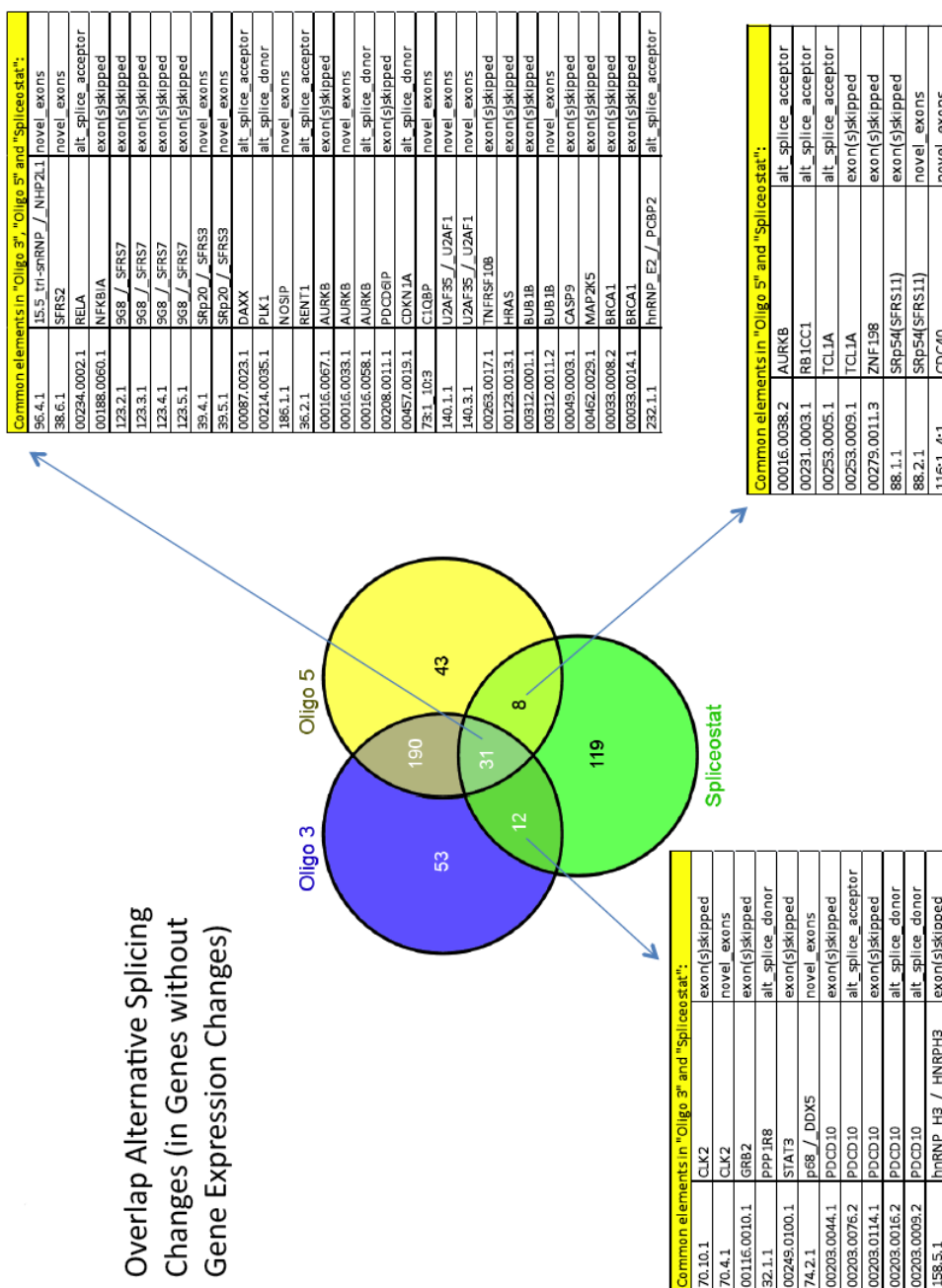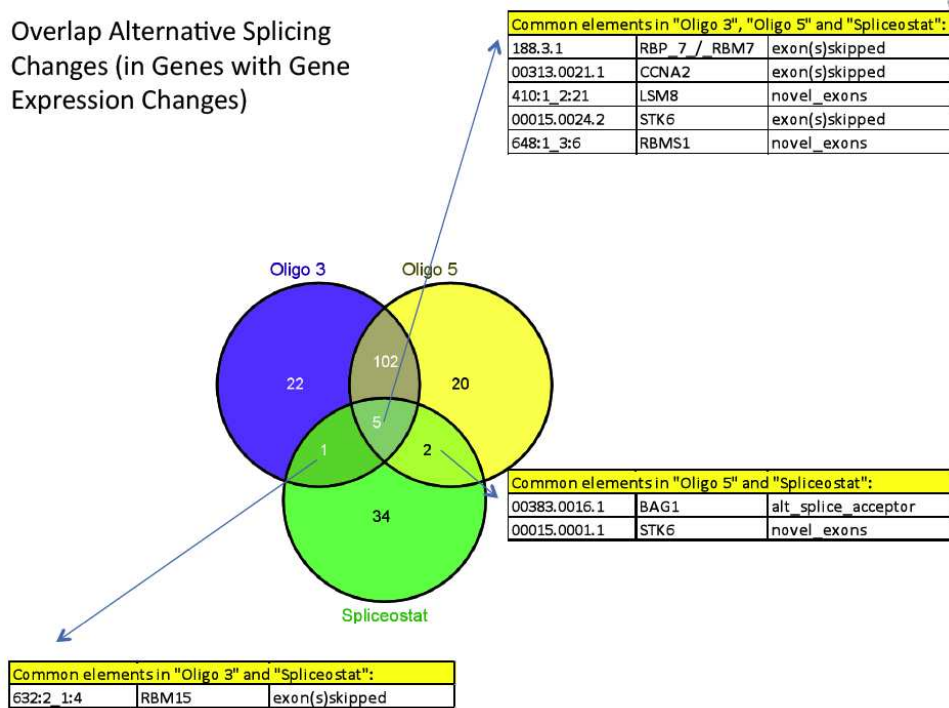| Common elements in "Oligo 3" and "Spliceostat":		
632:2_1:4	RBM15	exon(s)skipped

Fig. I. S7. Microarray-detected changes in alternative splicing in genes that also display gene expression changes upon SSA treatment and upon SF3b155 knock down using two different siRNAs. The diagram represents the overlap of genes displaying changes in overall gene expression as well as in expression of an alternatively spliced isoform of at least 1.4 fold between methanol- or SSA-treated cells and between cells treated with control scrambled siRNAs or treated with two different siRNAs corresponding to SF3b155 (oligos 3 and 5). The overlap between genes affected by SF3b155 knock down using the two siRNAs is 82% (Fischer exact test p value = 0.001). The overlap between genes affected by SSA and by either of the two siRNAs is 19% (Fischer exact test p value = 0.014), and the overlap with genes affected by both siRNAs is 12% (Fischer exact test p value = 0.064).

Fig. I. S8
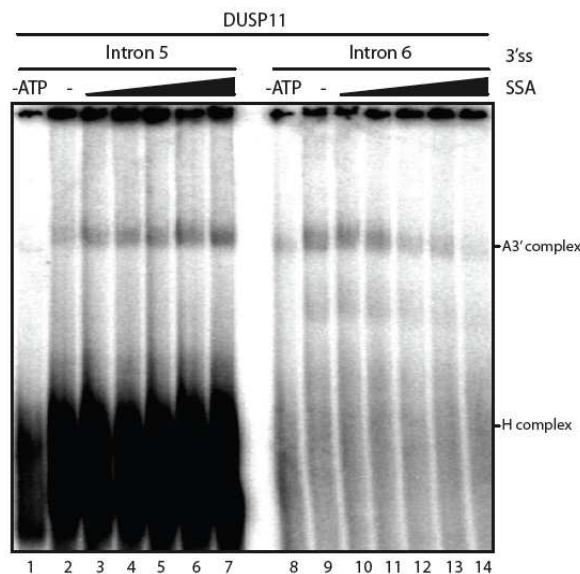
Fig. I. S8. Differential sensitivity to SSA of U2 snRNP recruitment to alternative 3' splice site regions of the gene DUSP11. Treatment with SSA induces skipping of DUSP11 exon 5. Spliceosome assembly assays were carried out as in Fig. I. 4A using RNAs corresponding to intron 5 and intron 6 3' splice site regions of DUSP11 gene, in the absence or presence of SSA (10, 20, 51, 88 and 178 nM). The complex formed on DUSP11 intron 5 3' splice site region remains largely resistant to SSA, while complexes formed on DUSP11 intron 6 3' splice site region are inhibited by SSA. The pattern of differential sensitivity is analogous to that observed for RBM5 (Fig. I. 4A).
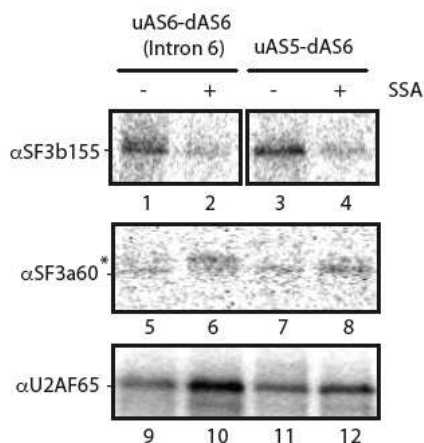
Fig. I. S9. SSA decreases crosslinking of SF3b155 to RBM5 intron 6 and uAS5-dAS6 RNAs. UV light-induced crosslinking and immunoprecipitation was performed as in Fig. 1F using radioactively labeled RNAs from RBM5 intron 6 3' splice site region or a mutant derivative (uAS5-dAS6) in which the upstream anchoring site (uAS) (Fig. I. 4B) was replaced by the equivalent sequence in RBM5 intron 5 3' splice site region, using the indicated antibodies against U2 snRNP components and U2AF65. As observed for the AdML 3' splice site region, SSA reduced crosslinking of SF3b155, while it did not reduce crosslinking of SF3a 60 or U2AF65. The asterisk indicates a crosslinked protein co-immunoprecipitated with the anti-SF3a60 antibody, possibly SF3a 66.

Chapter II: Strict 3' splice site sequence requirements for U2 snRNP recruitment after U2AF binding underlie a genetic defect leading to autoimmune disease

Anna Corrionero, Veronica Raker, José María Izquierdo and Juan Valcárcel

The data included in this chapter have been submitted for publication

Abstract

Pre-mRNA splice site sequences provide unequivocal information for intron excision, yet consensus splice sites are highly degenerate in higher eukaryotes, suggesting that multiple sequence arrangements can accurately signal exon/intron boundaries. Here we report that the 3' splice site associated with the alternatively spliced exon 6 of the Fas receptor CD95 displays strict sequence requirements and that a mutation that disrupts this particular sequence arrangement leads to constitutive exon 6 skipping in a patient suffering from Autoimmune Lymphoproliferative Syndrome (ALPS). Specifically, we find an absolute requirement for RCAG/G at the 3' splice site (where R represents purine and / indicates the intron/exon boundary) and that the balance between exon inclusion and skipping is exquisitely sensitive to single nucleotide variations in the uridine content of the upstream polypyrimidine (Py)-tract. Biochemical experiments revealed that the ALPS patient mutation reduces U2 snRNP recruitment to the 3' splice site region and that this effect cannot be explained by decreased interaction with the U2 snRNP Auxiliary Factor U2AF, whose 65 and 35 KDa subunits recognize the Py-tract and 3' splice site AG, respectively. The effect of the mutation, which generates a tandem of two consecutive AG dinucleotides at the 3' splice site, can be suppressed by increasing the distance between AGs, mutating the natural 3' splice site AG or by increasing the uridine content of the Py-tract. Unexpectedly, the latter requires that uridine enrichment occurs at a position distal from the 3' splice site. The suppressive effects of these additional mutations correlate with increased recruitment of U2 snRNP but not with U2AF binding, once again suggesting that the strict architecture of Fas intron 5 3' splice site region is tuned to regulate alternative exon inclusion through modulation of U2 snRNP assembly after U2AF binding.

Introduction

Removal of introns from pre-mRNA precursors (pre-mRNA splicing) is a necessary step for expression of the majority of higher eukaryotic genes (Wahl et al., 2009) and alternative patterns of intron removal greatly expand the information content of complex genomes (Nilsen and Graveley, 2010). Three sequence elements help to specify the 3' end of introns in higher eukaryotes: the branch point, the 3' splice site AG and the intervening polypyrimidine (Py)-tract. The branch point, a virtually invariant adenosine that forms a 2'-5' phosphodiester bond with the 5' end of the intron after the first catalytic step of the splicing reaction, is flanked by the sequence UACUAAC in budding yeast (A denotes the branch point adenosine) and by the much more variable consensus YNYURAY in human genes (Y indicates pyrimidine, R purine, N any nucleotide) (Zhang, 1998). The branch point region is first recognized by the Branch Point Binding Protein (BBP/SF1) (Berglund et al., 1997; Liu et al., 2001a) and subsequently through base pairing interactions with U2 snRNA, the RNA component of the U2 snRNP ribonucleoprotein (Zhuang and Weiner, 1989; Nelson and Green, 1989; Wu and Manley,

1989). The invariant 3' terminal AG dinucleotide and the preceding Py-tract are recognized, respectively, by the 35 and 65 KDa subunits of the U2 snRNP Auxiliary Factor (U2AF), which assists in the recruitment of U2 snRNP to the branch point region (Zamore et al., 1992; Wu et al., 1999; Zorio and Blumenthal, 1999; Merendino et al., 1999).

A large variety of sequences fit into this general organization of the intron 3' end, and cooperative interactions help to functionally define 3' splice sites. For example, interaction between U2AF and BBP/SF1 (Selenko et al., 2003) mutually stabilizes the binding of these factors to their loosely defined adjacent cognate sites (Berglund et al., 1998). Another example are stabilizing interactions mediated by arginine-serine-rich (RS) domains, which support protein-protein interactions as well as interactions between proteins harboring these domains and double stranded RNA regions (e.g. those resulting from base pairing between U2 snRNA and the branch site region) (Wu and Maniatis, 1993; Shen et al., 2004; Shen and Green, 2004; Shen and Green, 2006). Furthermore, the arrangement of protein and RNA contacts in the initial complex has been proposed to enforce a unique RNA structure in the 3' splice site that may be critical to trigger subsequent events in the splicing process (Kent et al., 2003). In addition, proofreading activities exist that ensure proper assembly of splicing factors in bona fide splice site signals, e.g. allowing the assembly of U2AF on uridine-rich sequences only if followed by a 3' splice site AG dinucleotide (Soares et al., 2006).

A third mechanism contributing to splice site identification is the sequential recognition of the same signal by different factors at different stages of spliceosome assembly or during the processes leading to catalysis. For example, the 3' splice site AG dinucleotide is recognized initially by U2AF35, an interaction that becomes essential for spliceosome assembly on AG-dependent introns, which are characterized by relatively uridine-poor Py-tracts that can only weakly recruit U2AF65 in the absence of the U2AF35/AG interaction (Wu et al., 1999). The splicing factor hSlu7 re-structures the spliceosome after the first catalytic step of the splicing reaction and positions the 3' splice site AG for engaging in the second catalytic step over neighboring AG dinucleotides, thus helping to generate properly spliced mRNA products (Chua and Reed, 1999a, b). Neighboring AGs can however strongly influence splice site choice, facilitating early interactions with the spliceosome but allowing other sites to be used during catalysis (Chua and Reed, 2001; Lallena et al., 2002; Lev-Maor et al., 2003). 3' splice sites are further recognized by proofreading mechanisms implemented in budding yeast by the DExD/H-box ATPase Prp22p that prevent exon ligation in premRNAs containing mutations in splicing signals, including the 3' splice site (Mayas et al., 2006).

The Fas / CD95 gene encodes a pro-apoptotic membrane receptor (Strasser et al., 2009) that has been recently shown to be also required for efficient growth of cancer cells (Chen et al., 2010). An alternatively spliced variant that skips exon 6 lacks the trans-membrane domain and encodes a soluble, form of the receptor that can be anti-apoptotic (Cheng et al., 1994). A switch in alternative splicing of exon 6 occurs during T cell activation

such that the anti-apoptotic form of the receptor is produced during antigen-triggered clonal amplification, while the pro-apoptotic form contributes to eliminate the expanded cell population upon completion of the immune response (Liu et al., 1995; Russell, 1995a). Failure to accomplish this alternative splicing switch is linked with persistent high levels of lymphocytes in the lymph nodes, liver and spleen, which lead to Autoimmune Lymphoproliferative Syndrome (ALPS) (Roesler et al., 2005). Several splicing factors have been shown to modulate the levels of Fas exon 6 inclusion or skipping by various molecular mechanisms, including targeting early events in splice site recognition / exon definition and modulating splice site pairing after U2 snRNP binding (Forch et al., 2000; Izquierdo et al., 2005; Corsini et al., 2007; Bonnal et al., 2008).

In this manuscript we have characterized the molecular defect behind a mutation in the Fas/CD95 gene identified in a patient with ALPS, which leads to constitutive exon 6 skipping. The results reveal strict sequence requirements at the 3' splice site and the existence of a rate-limiting step for exon inclusion in the recruitment of U2 snRNP after U2AF binding.

Results

Strict sequence requirements in the 3' splice site region of Fas intron 5

It has been previously shown that a C to G mutation at position −3 of the 3' splice site of Fas intron 5 correlates with increased exon 6 skipping in activated T and B cells derived from a patient suffering from Autoimmune Lymphoproliferative Syndrome (ALPS) (Roesler et al., 2005). Because this mutation generates a tandem duplication of the 3' splice site AG dinucleotide, the mutant will be described hereafter as "AG". To confirm that the splicing defect can be exclusively attributed to this mutation and also to evaluate the possibility of a cell type-specific contribution, HeLa cells were transfected with wild type and mutant minigenes containing Fas genomic region from exon 5 to 7 under a CMV promoter. After RNA isolation, the pattern of Fas splicing was analyzed by semi-quantitative RT-PCR using vector-specific primers. As observed in ALPS patient's cells, the mutation led to a dramatic switch from preferential exon 6 inclusion to quantitative skipping (Fig. II. 1A).

To test the relative effect of the AG mutation on splicing of each of the introns flanking exon 6, minigenes containing the genomic regions between exons 5 and 6 (including the 5' splice site of intron 6, to allow for exon definition effects), or between exons 6 and 7 (including also the 3' 68 nucleotides of intron 5 and a deletion of intron 6 (50-1102) known not to affect Fas splicing or regulation (Izquierdo et al., 2005) were transfected in HeLa cells and RNA was analyzed as above. The mutation led to detection of only unspliced RNA from intron 5-containing transcripts, while the efficient splicing of intron 6 was unaffected by the mutation (Fig. II.1B). We conclude that the AG mutation leads to exon 6 skipping primarily by inhibiting splicing of intron 5, without major effects (e.g. through exon definition) on splicing of intron 6.

To test whether other mutations in the region can also influence exon 6 inclusion, single point mutations were tested. Mutation of the C at position -3 to U (C-3U) also led to exon skipping (Fig. II. 1D, lane 3), indicating that the effect of the ALPS mutation is not necessarily due to the generation of a tandem of potential 3' splice sites AG dinucleotides. Mutation of the A at position –4 to C or U also caused exon skipping (Fig. II. 1D, lane 4 and data not shown), while transition to G had only a minor effect (Fig. II. 1D, lane 5), suggesting that a purine at this position is necessary for exon inclusion. Mutation of position –5 (U) to G had only minor effects and, in fact, mutation to C increased exon 6 inclusion (Fig. II. 1D, lanes 6 and 7). Mutation of position –6 (C) to another pyrimidine maintained the pattern of preferential exon 6 inclusion, while mutation of this or the C at position 7 to purine residues led to exon 6 skipping (Fig. II. 1D, lanes 8-10). These results suggest that the pyrimidine content of this region is important for exon 6 inclusion. Consistent with this, mutation of A to U at position –9 increased exon 6 inclusion, while mutation of pyrimidines to purines at positions -10 to -12 significantly increased exon 6 skipping (Fig. II. 1D, lanes 11-15). Finally, mutation of the G at exon 6 position +1 also led to quantitative exon skipping, while mutation of position +2 did not (Fig. II. 1E). We conclude that Fas exon 6 inclusion depends on the consensus RCAG/G at the 3' splice site (R designates a purine; / indicates the intron/exon boundary) and that the balance between exon inclusion and skipping is finely tuned by single nucleotide differences in the pyrimidine content of the upstream sequence.

Consistent with this notion, regeneration of precisely the RCAG/G consensus around the original AG mutation led to complete reversion of the effect of the mutant and to the use of the newly generated AG as the 3' splice site (Fig. II. 1F, lanes 1-5 and data not shown).

The AG mutant inhibits intron 5 splicing in vitro and prevents pre-spliceosomal A complex formation

To investigate the molecular mechanism by which the AG mutant causes inhibition of intron 5 splicing, in vitro splicing assays were carried out using HeLa nuclear extracts and in vitro transcribed RNAs corresponding to wild type and AG mutant Fas pre-mRNA sequences from exon 5 to nucleotide +25 of intron 6. Although the difference was not as dramatic as in minigene-transfected cells, the splicing efficiency of the AG mutant was significantly lower than that of the wild type. This was the case for the accumulation of products of the first catalytic step (lariat intermediate) at early time points as well as for the accumulation of products of the second catalytic step (lariat and spliced products) at later time points (Fig. II. 2A).

Next we analyzed the step of the spliceosome assembly pathway at which the AG mutant RNA was inhibited. Lower levels of the ATP-dependent pre-spliceosome A and spliceosome B/C complexes were assembled on the AG mutant compared to the wild type RNA, starting at early incubation time points (Fig. II. 2B). Similar results were obtained using RNAs spanning the 3' 68 nucleotides of intron 5, exon 6 and the 5' 25 nucleotides of intron 6,

which assemble an ATP-dependent A3' complex (Fig. II. 2C), indicating that the effect of the AG mutant does not require the presence of an upstream 5' splice site. The same RNAs were used to monitor the interaction of U1 and U2 snRNAs with the 5' and 3' splice site sequences flanking exon 6 using psoralen-mediated UV crosslinking (Wassarman and Steitz, 1992), as previously described (Izquierdo et al., 2005). The results of Figure II 2D show that the AG mutation did not cause significant changes in U1 crosslinking, while crosslinking of U2 was substantially decreased. We conclude that the ALPS-causing mutation reduces the interaction of U2 snRNP with Fas intron 5 3' splice region.

Initial U2AF recruitment is not affected in the AG mutant

Next we tested whether the decrease in U2 snRNP recruitment in the AG mutant could be due to decreased binding of U2AF using short wavelength ultraviolet light-mediated crosslinking and immunoprecipitation with specific antibodies (Gama-Carvalho et al., 1997). Interestingly, at early times of incubation (5 minutes), when clear differences in A complex formation between the wild type and mutant AG are observed (Figs. II. 2B, 2C), no significant differences in U2AF65 crosslinking could be detected (Fig. II. 3A, lanes 1-2). Interestingly, U2AF65 crosslinking to the wild type RNA was higher than to the mutant AG at later incubation times (lanes 3-6). These observations suggest that the AG mutation reduces the stability (e.g. by accelerating the dissociation) of U2AF65 interaction with the pre-mRNA.

Interaction of U2AF35 with the pre-mRNA was also analyzed using 32P-guanidine-labeled transcripts, to improve the sensitivity of detection of the U2AF35 / 3' splice site AG interaction. Figure 3B shows that neither the levels of U2AF35 crosslinking nor the levels of crosslinked U2AF65 co-immunoprecipitated with U2AF35 were lower for the AG mutant than for the wild type RNA at 5 minutes of incubation, while they were slightly decreased at 15 minutes, as observed in Figure 3A for U2AF65.

We conclude that reduced levels of U2 snRNP recruitment in the AG mutant do not correlate with reduced levels of U2AF heterodimer crosslinking to the pre-mRNA. Interestingly, reduced crosslinking of U2AF65 was observed, even at early incubation times, for other two mutants (C-3U and A-4C) (Fig. II. 3C) which also led to exon 6 skipping in transfection assays, suggesting that reduced U2AF recruitment contributes to the splicing defect in these mutants.

To further explore the relationship between U2AF recruitment and the effect of the AG mutation, two additional mutants were tested. Previous work showed that increasing the complementarity between U1 snRNA and the 5' splice site site of intron 6 enhances exon 6 inclusion through exon definition effects that enhance U2AF65 binding to the Py-tract (Izquierdo et al., 2005). These bridging interactions could be prevented by the function of the splicing repressor PTB acting from a silencer sequence located in the central region of exon 6 (Izquierdo et al., 2005). Improving the strength of intron 6 5' splice site (U1 comp) or mutating the PTB binding site in the exonic silencer (m0) increased exon 6 inclusion (Fig. II. 3D,

compare lanes 1 and 3 and data not shown). Neither of these mutations, however, increased exon 6 inclusion in the context of the AG mutant (Fig. II. 3D, lanes 4 and 8). These results are once again consistent with the notion that U2AF binding is not the rate-limiting step leading to exon 6 skipping in the AG mutant.

Suppressing the effects of the AG mutation

Previous studies have shown that the relative use of neighboring 3' acceptor AGs depends upon their relative distance to the branch point and to each other (Chua and Reed, 2001; Lev-Maor et al., 2003). To determine whether these effects were relevant in the case of the AG mutant, the distance between AGs was increased up to five nucleotides. Introducing one, two or three nucleotides did not prevent complete exon 6 skipping (Fig. II. 4A, lanes 1-5). Introducing four or five nucleotides, however, led to exon inclusion levels similar to those of the wild type RNA (lanes 6-8), resulting in the use of the proximal 5' acceptor site (as determined by sequencing of the cDNA products, data not shown). These distance effects were observed with different intervening sequences (lanes 1-8). The presence of a downstream acceptor AG plays a role in inhibition, because mutation of the 3' AG partially relieved exon 6 skipping (compare lane 11 with lane 4), while it did not promote exon inclusion in a transcript where the more distal AG did not display an inhibitory effect (compare lane 10 with lane 6). These observations indicate that the presence of a distal acceptor 3' splice site less than four nucleotides from the proximal AG contributes to the inhibition of exon 6 inclusion in the AG mutant.

Next, we evaluated the contribution of the strength and configuration of the Py-tract on the effect of the AG mutation (Fig. II. 4B). Inserting 8 uridines and 3 cytidines upstream of the AGAG sequence led to a modest increase in exon 6 inclusion (Fig. II. 4C, lane 5). Insertion of 9 or 10 uridines, however, led to quantitative exon inclusion (lanes 3-4), revealing the remarkable effect of a single additional uridine for activation of the mutant AG as a 3' acceptor site.

Because insertion of 10 nucleotides at the Py-tract could alter 3' splice site recognition by increasing the distance between the branch point and the 3' splice site, substitution mutants in the Py-tract were also generated (Fig. II. 4B). Substantial levels of exon 6 inclusion were observed in constructs containing 6 (but not 4) contiguous uridines (Fig. II. 4D, lanes 1-5). Interestingly, however, the position of the 6 uridines was functionally important because a 6 uridine stretch located just upstream of the 3' splice site AG was not able to promote exon 6 inclusion, while the same length uridine stretch located further upstream induced predominant exon inclusion (Fig. II. 4D, lanes 5-7).

To investigate the molecular mechanisms behind the effects of these mutations, spliceosome assembly, psoralen-mediated RNA/RNA crosslinking and U2AF65 crosslinking/immunoprecipitation assays were carried out using radioactively labeled RNAs as in Figure 3. A good correlation was found between A 3' complex assembly, U2 snRNA

psoralen-mediated crosslinking and exon 6 inclusion (compare Figs. II. 4A, D, E and F). Specifically, wild type Fas showed similar U2 snRNP assembly and U2 snRNA crosslinking that mutants AG6 (AGs separated out by 4 nucleotides) and U6C4U (6 uridine stretch at a distal position from the upstream AG) (Fig. II. 4E, lanes 1-2, 5-8 and Fig. II. 4F, lanes 1, 3 and 5) consistent with preferential inclusion of exon 6 observed upon transient transfection of constructs harboring these mutations (Figs. II. 4A and D). Conversely, reduced levels of U2 snRNP assembly and U2 snRNA crosslinking were detected for the AG and the UGC4U6 (6 uridine stretch proximal to the upstream AG) mutants (Fig. II. 4E, lanes 3-4, 9-10 and Fig. II. 4F, lanes 2 and 4), consistent with preferential exon 6 skipping observed in transient transfection assays (Fig. II. 4D).

Finally, we monitored the interaction of U2AF65 with the Py-tract at 5 and 15 minutes by UV-mediated crosslinking followed by U2AF65 immunoprecipitation. At 5 minutes, U2AF65 crosslinked efficiently with the WT, AG, AG6 and U6C4U RNAs. The interaction decreased in the AG mutant at 15 minutes, while it was maintained or even increased at this time point with WT, AG6 and U6C4U (Fig. II. 4G, lanes 1 to 4). In contrast, the UGC4U6 mutant, which displayed exon 6 skipping and reduced complex A assembly in vitro, showed reduced U2AF65 crossliking at either time point (Fig. II. 4G, lanes 5).

We conclude that the splicing defect generated by the AG (C-3G) mutation can be relieved by increasing the distance between the newly formed AG dinucleotide and the natural 3' splice site AG, by mutating the natural 3' splice site or by increasing the strength of the Py-tract. The latter can be achieved with a Py-tract of 6 or more consecutive uridines located not immediately upstream of the 3' splice site. The results also show that while decreased U2AF65 interaction can be correlated with decreased binding of U2 snRNP to the 3' splice site region of intron 5 (e.g. mutant UGC4U6), decreased exon 6 inclusion in the ALPS mutation cannot be explained by decreased interaction with U2AF65, but rather by a defect in U2 snRNP recruitment subsequent to U2AF65 binding.

Discussion

The consensus 3' splice site in human introns is NYAG/R (N=any nucleotide, Y=pyrimidine , R=purine, /=intron-exon boundary) (Zhang, 1998). Our results show that the function of the 3' splice site of Fas intron 5, and consequently inclusion of the alternatively spliced exon 6, requires the sequence RCAG/G (Fig. II. 1), which differs from the general consensus (Zhang, 1998) in several important features. First, a purine is required at position -4, while no clear nucleotide preference can be found at this position for the bulk of 3' splice sites. Second, a cytidine is strictly required at position -3, while uridine can be found at this position in up to 37% of human 3' splice sites. Third, guanidine cannot be replaced by adenine at position +1, while adenine can be found in up to 26% of human sites. The consensus CAG/G emerged as the high affinity site selected by the U2AF 35 KDa subunit in

iterative binding selection experiments using the U2AF65-U2AF35 heterodimer (Wu et al., 1999). It is therefore conceivable that the strict sequence requirements at positions -3 and +1 reflect the need for tight U2AF35 binding in Fas intron 5 3' splice site activation. While some of the mutants breaking the RCAG/G consensus decreased U2AF65 crosslinking, no differences in U2AF35 crosslinking were observed between wild type and AG mutant RNAs in nuclear extracts (Fig. II. 3B). This could reflect a limitation of the crosslinking assay, but also that other factor(s) specifically involved in recognition of this 3' splice site display exact sequence requirements for positions -3, +1 (and perhaps also for the uncharacteristic need of a purine at position -4). Biochemical experiments using site-specifically radioactively labeled RNAs as well as RNA affinity chromatography, however, failed to reveal factors that differentially interact with the wild type and AG mutant 3' splice sites (data not shown).

While U2AF65 crosslinking was similar to wild type and AG mutant at early time points (when differences in spliceosome assembly were already obvious), crosslinking signals were lower for the AG mutant at later times (Fig. II. 3A). This may reflect higher off rates of U2AF binding in the mutant or –intriguingly– the existence of a proofreading activity that displaces more readily U2AF from the mutant RNA after the initial steps of spliceosome assembly. In this regard, the protein DEK proofreads interactions of U2AF with the Py-tract RNAs followed by a bona fide 3' splice site AG (Soares et al., 2006).

Alternatively, other aspects of U2AF35 or U2AF65 activities leading to U2 snRNP recruitment may display strict 3' splice site sequence requirements in the context of Fas intron 5. Two non-mutually exclusive mechanisms have been proposed for how U2AF promotes the recruitment of U2 snRNP. Upon recognition of the Py-tract by U2AF65, its RS domain is positioned in the vicinity of the branch point region and promotes base pairing between this region and the branch point recognition site in U2 snRNA (Valcarcel et al., 1996). An interaction between the carboxy-terminal U2AF Homology Domain (UHM) of U2AF65 and the U2 snRNP protein SF3b155 can also contribute to U2 snRNP recruitment (Gozani et al., 1998). It is therefore conceivable that even though similar U2AF binding occurs on the wild type and on the AG mutant 3' splice site region, the architecture of the complex formed (including conformational changes in the U2AF subunits induced upon RNA binding) differs in the mutant such that proper positioning of the RS domain near the branch point region, or proper interaction of the UHM domain with SF3b155, cannot occur. The observation that rescue of the effect of the AG mutant by increased uridine content in the Py-tract requires at least six uridines positioned not immediately upstream from the 3' splice site is consistent with the requirement of a special RNP structure for activation of intron 5 3' splice site. Consistent with the notion of 3' splice site architectures highly sensitive to mutation, in vivo evidence has been reported in C. elegans that single nucleotide variations can have profound effects in 3' splice site selection (Ma and Horvitz, 2009).

Although similar effects are induced by the AG mutation and by other mutants in the RCAG/G consensus, the generation of a new potential 3' splice site is relevant for the splicing defect observed in the ALPS-causing mutation. The newly generated AG dinucleotide can be engaged in splicing reactions when the distance with the natural AG is increased by 4 or more nucleotides or when the natural site is mutated (Fig. II. 4), indicating that the natural 3' splice site prevents use of the newly created 3' splice site. This is another peculiarity of intron 5 3' splice site region, because tandem AG dinucleotide arrangements do not necessarily result in exon skipping but rather in activation of the mutant AG for splicing (as is the case of the hemophilia-causing mutation transmitted in certain European royal families (Rogaev et al., 2009). In addition, NAGNAG acceptors are very common in genes from a variety or organisms from land plants to humans (Hiller et al., 2004; Schindler et al., 2008). Although the use of the alternative NAGNAG sites (known as 3' splice junction wobbling) has been argued to be stochastic (Chern et al., 2006; Sinha et al., 2009), evidence exists for functional effects of the single amino acid differences generated by the use of the alternative acceptors (Tsai et al., 2008) and for arrangements of regulatory sequences at these regions (Tsai et al., 2007; Akerman and Mandel-Gutfreund, 2006).

The splicing factor hSlu7 is thought to contribute to the reorganization of the spliceosome between the first and second catalytic steps of the splicing reaction and play a role in holding the free 2' OH of the released upstream exon near the 3' splice site AG to allow exon ligation (Chua and Reed, 1999b). In the absence of the fidelity function of hSlu7, other neighboring AGs become activated and used for splicing (Chua and Reed, 1999b). Knockdown of hSlu7 failed however to affect wobble splicing choices (Tsai et al., 2010) and also failed to relieve the inhibitory effect of the AG mutant (data not shown).

Another factor that can play a role in the selection between alternative AGs is SPF45 (Lallena et al., 2002). SPF45 regulates alternative splicing of Fas exon 6, an activity that requires interactions between SPF45 and U2AF65, SF1/BBP and SF3b155 (Corsini et al., 2007). Knockdown of SPF45 (as well as knockdown of a variety of other RNA binding proteins) failed however to promote exon inclusion in the mutant AG (unpublished results).

A variety of other factors regulate Fas alternative splicing, including TIA-1, PTB and RBM5. TIA-1 promotes exon inclusion by facilitating U1 snRNP recruitment to the 5' splice site of intron 6 and exon definition, while PTB causes exon skipping by binding to an exonic silencer motif and interfering with exon definition (Forch et al., 2000; Izquierdo et al., 2005). The target of these regulators through the exon definition process is U2AF binding to intron 5 3' splice site, therefore making them unlikely contributors to the splicing defect caused by the AG mutant, which is also consistent with the results of the mutational analysis of Figure 3D. RBM5 regulates exon 6 skipping by blocking splice site pairing after U2 snRNP binding, at the time of U4/5/6 assembly (Bonnal et al., 2008), an activity that therefore operates after the step blocked in the AG mutant.

Our mutational analyses show that regulation of U2 snRNP recruitment after U2AF binding can very significantly alter the balance between exon inclusion and skipping, opening the possibility that yet to be identified regulators of Fas exon 6 splicing target this important step. More generally, our results reveal that the particular configuration of sequences at this 3' splice site is delicately tuned, being very sensitive to nucleotide replacement and positional effects. This is in contrast with the apparent flexible sequence requirements suggested by loose consensus sequences derived from comparison of a large number of 3' splice site regions. For example, a genome-wide study in fission yeast revealed a wide array of sequence arrangements at 3' splice sites, ranging from conventional position of the Py-tract to functional Py-tracts located 5' of the branch point to 3' splice site regions lacking any dectectable Py-tract (Sridharan and Singh, 2007). Consistent with the standard function of U2AF, the same study reported variable dependence of U2AF large subunit for intron removal.

Collectively, these considerations suggest, on one hand, that the specific RNP architecture of particular 3' splice site regions can be finely adjusted and therefore be very sensitive to mutation, an observation relevant for splicing regulation as well as for ALPS and a variety of other genetic diseases caused by splicing defects (Buratti et al., 2006). On the other hand, they suggest that a variety of sequences, factors and regulatory steps can contribute to 3' splice site recognition through mechanisms that for the most part remain to be understood.

Materials and Methods

Plasmid mutagenesis

Mutants were generated by polymerase chain reaction (PCR)-based site-directed mutagenesis of the expression vector containing human Fas genomic sequences as described (Forch et al., 2000) using primers carrying the desired mutation, 2,5 units TaqPlus Precision Polymerase (Stratagene) and 26 amplification cycles (94°C for 1 min, 56°C for 1 min, and 72°C for 7 min). PCR products were purifie d (PCR purification kit, Qiagen), digested with 20 units of DpnI (New England BioLabs) and treated with 5 units of T4 DNA polymerase (Fermentas). After precipitation, ligation was carried out with 400 units of T4 DNA ligase (New England BioLabs) and transformed into heat shock-or electro-competent XL1-Blue cells. All the mutants were confirmed by sequencing.

DNA transfections, RNA purification and analysis

Exponentially growing HeLa cells grown in Dulbecco's Modified Eagle's Medium (DMEM), 10% FCS (fetal calf serum) and 1% penicillin/streptomycin, were transfected at 50% confluence (5 x 10^6 cells) with 0,2 µg of plasmid DNA using ExGen 500 (Fermentas) or Fugene (Roche) following the manufacturer's instructions for adherent cell lines. Cells were incubated 24 h before RNA purification and total RNA was isolated using RNeasy kit

(Qiagen). Reverse transcription-PCR was carried out as described (Forch et al, 2000). Briefly, reverse transcription was performed using 1 µg of DNase-treated total RNA and the PT2 minigene-specific primer (5'-aagcttgcatcgaatcagtag-3') for 1 h at 42°C. Semi-quantitative PCR was then performed using GoTaq DNA polymerase (Promega) and 24 amplification cycles (93°C for 1 min, 60°C for 45 seconds, and 72°C for 1 min) using oligonucleotides PT1 (5'-gtcgacgacacttgctcaac-3') and PT2. PCR products were analyzed by electrophoresis in 2% agarose gels.

In vitro transcription of splicing substrates

Transcription templates were generated by PCR from the different plasmids, with a T7 promoter included in the upstream primer. Radiolabeled pre-mRNAs were transcribed by using approximately 1 µg of template DNA in 25 µl transcription reactions containing final concentrations of 40 mM Tris-HCl, pH 7.9, 10 mM NaCl, 6 mM $MgCl_2$, 2 mM spermidine, 0.8 mM dithiothreitol, 0.5 mM ATP, CTP and GTP,

0.05 mM UTP, 30 µCi [α-^{32}P]UTP, and 60 units of T7 or 42 units of SP6 RNA polymerases (Promega), for 2 hour at 37ºC. The transcripts were isolated by gel purification followed by phenol/choloroform/isoamylalcohol (25:24:1) extraction and ethanol precipitation.

In vitro splicing and spliceosome assembly assays

Body radiolabeled RNA (10 fmols) was incubated with 33% of HeLa cells nuclear extract, 3,3% Polyvynil Alcohol, 1 mM ATP, 2,7 mM $MgCl_2$, 20 mM creatine phosphate, 4 units of RNasin (Promega), up to 9 µl with buffer D, 0.1 M KCl, 1 mM DTT at 30°C for the indicated time points. For in vitro splicing assays, the reaction was stopped by treating for 20 minutes at 65°C with proteinase K. The RNAs were isolated by phenol/choloroform/isoamylalcohol (25:24:1), ethanol precipitated and loaded on 13% denaturing polyacrylamide gels. Gels were analyzed by exposing directly to PhosphorImager screens. For spliceosome assembly assays, the reactions were treated with 5 µg/ul of heparin for 10 min at room temperature and resolved on 4 % acrylamide:bisacrylamide (80:1)-0,5 % agarose gels in 50 mM Tris base – 50 mM glycine buffer. The gels were dried and exposed to a PhosphorImager screen. Alternatively, 2.2 µl of 6x DNA loading dye (20 mM Tris-HCl pH 7.5, 0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol) were added to the reaction after treatment with heparin (5mg/ml) for 10 minutes at room temperature. Mixtures were resolved either in agarose-polyacrylamide gels or in 1.5% low melting agarose gels (Ecogen) in 50 mM Tris base – 50 mM glycine buffer. The gels were fixed in 10% methanol – 10% acetic acid for 10' at room temperature, dried and exposed to a PhosphorImager screen.

UV cross-linking and immunoprecipitation

32P-labeled RNAs were incubated under in vitro splicing conditions in a total volume of 27 µl. After incubation for 30 min at 30ºC, mixtures were irradiated with UV light (254 nm; 0.4

J). After incubation with 1 mg/ml RNase A at 37ºC for 30 min, U2AF65 and U2AF35 were immunoprecipitated with the MC3 monoclonal antibody hybridoma supernatant (Gama-Carvalho et al., 1997) and a polyclonal antibody (Zuo and Maniatis, 1996), respectively. After incubating for 60 min at 4ºC, 30 µl of a 50% slurry of protein A/G sepharose beads were added and the mixture was incubated under rotation for another 60 min. Beads were sedimented by centrifugation and washed 4 times with 800 µl of high salt buffer (500 mM NaCl, 50 mM Tris-HCl pH 8.0, 1% NP-40) and once with the same buffer with 100 mM NaCl. Sedimented beads were resuspended in 4xSDS loading dye, heated at 95ºC for 5 minutes, briefly centrifuged and the supernatant was loaded on 10% SDS-polyacrylamide gels. Gels were dried and exposed to a PhosphorImager screen.

Psoralen-mediated RNA-RNA crosslinking

Fifteen-microliter splicing reactions containing 80 fmol of 32P-uridine labeled RNAs were incubated at 30ºC for 20' in the presence of 22.2 µg/ml 4'aminomethyltrioxsalen hydrochloride (Sigma). The mixture was irradiated with 365 nm wavelength UV light for 10 min. RNAs were then treated with proteinase K, ethanol precipitated and fractionated on a denaturing 6% polyacrylamide gel. RNase H mediated inactivation of U2 snRNP was performed as described previously (Merendino et al, 1999) by incubating nuclear extracts with oligodeoxynucleotides complementary to U2 snRNA (5'-CAGATACTACACTTG-3'), U1 snRNA (5'-CTGCCAGGTAAGTAT-3'), U4 snRNA (5'-GGGTATTGGGAAAAGTTTTCAATTAGCAATA-3') and U5 snRNA (5'-TTAAGACTCAGAGTTGTTCC-3').

Figures and figure legends
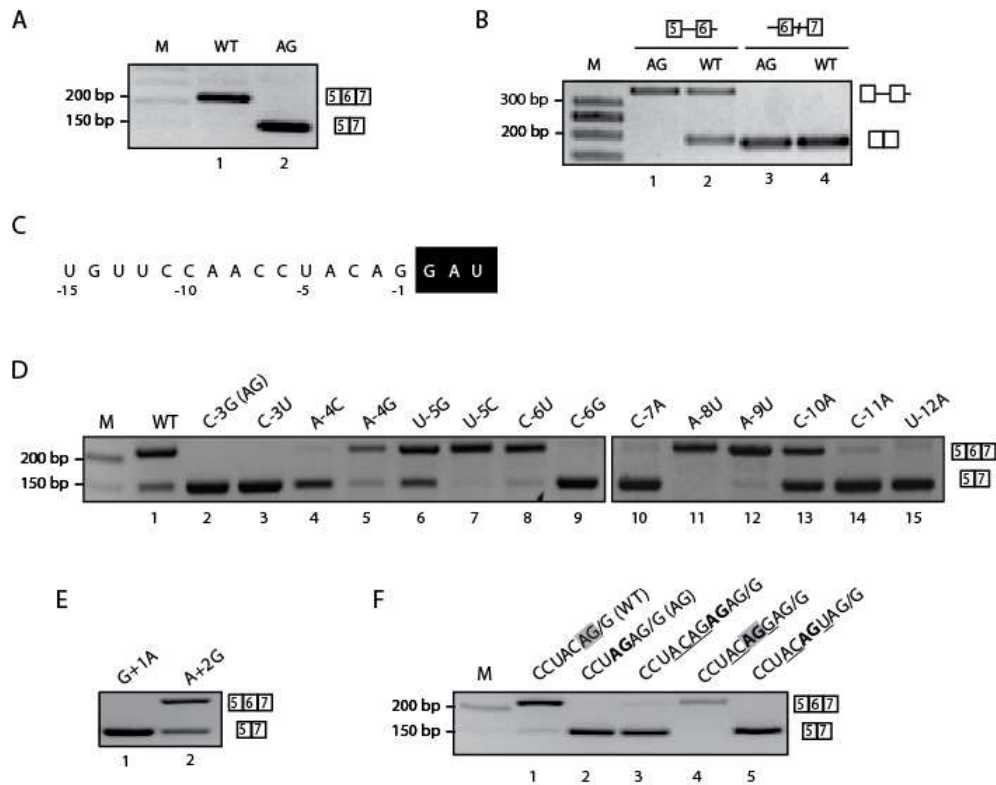
Fig. II. 1. Strict requirement of RCAG/G in Fas intron 5 3' splice site. (A) Fas minigenes encompassing exon 5 to exon 7 carrying the wild type sequence or C to G mutation at position -3 of intron 5 (referred to as AG because it generates a duplicated AG dinucleotide at the 3' splice site) found in an ALPS patient, were transfected in HeLa cells. RNA was extracted and alternative splicing pattern was monitored by reverse transcription followed by semi-quantitative RT-PCR. The alternatively spliced products were resolved by agarose gel electrophoresis. The positions of size markers and predicted alternatively spliced products are indicated. **(B)** Splicing pattern of wild type and AG minigenes containing single introns. Analyses were as in (A) using primers appropriate to detect spliced and intron retention products. The vertical line in intron 6 represents a deletion that facilitates the detection of unspliced RNA but that does not affect Fas splicing or regulation (Izquierdo et al., 2005). **(C)** Schematic representation of the Fas intron 5 3' splice site sequence. Numbers indicate the distance from exon 6, which is represented by a black box. **(D)** Mutational analysis of intron 5 3' splice site region. The indicated minigene mutants were transfected in HeLa cells and the alternative splicing pattern monitored as in (A). Lanes 1 and 2 correspond to the wild type and and AG minigenes, respectively. **(E)** Mutational analysis of exon 6 positions +1 and +2. **(F)** Reconstructing a ACAGG sequence (underlined) suppresses the effects of the AG mutant and activates the mutant AG. The indicated 3' splice site region mutants were analyzed as in (D). / indicates the original intron/exon boundary. AG in bold indicates the AG generated by the ALPS mutation.
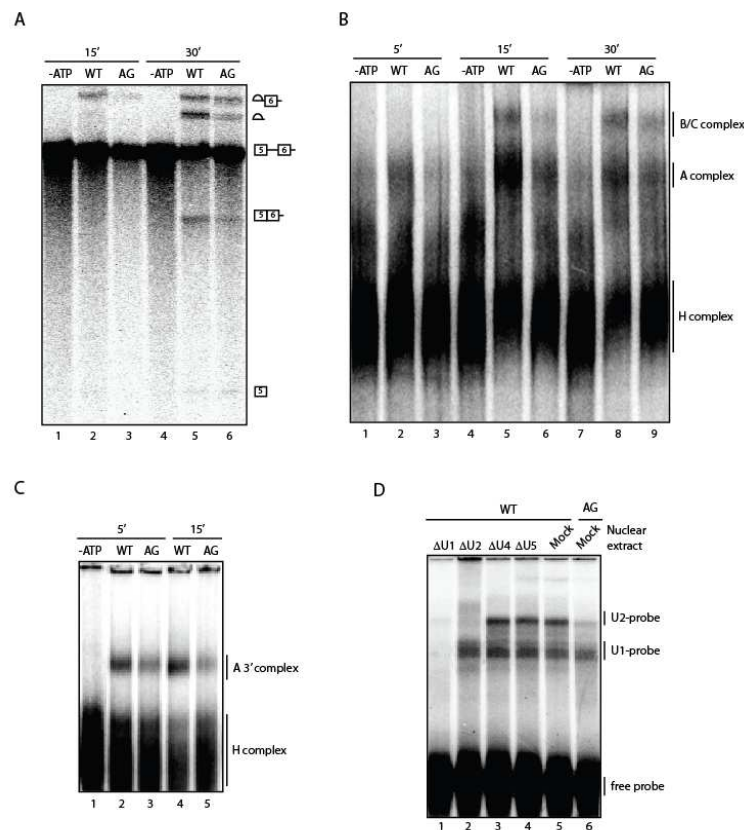
Fig. II. 2. The AG mutant inhibits U2 snRNP recruitment. (A) Inhibition of pre-mRNA splicing in vitro by the AG mutant. Uniformly α ^{32}P-UTP labeled RNAs corresponding to wild type (WT) or AG mutant Fas genomic region between exon 5 and position +25 of intron 6 were incubated with HeLa cell nuclear extracts under splicing conditions (see Materials and Methods) in the absence or presence of ATP for 15 and 30 minutes. RNA species were subsequently purified, and analyzed by electrophoresis in 13% denaturing polyacrylamide gels. The positions of pre-mRNA, splicing intermediates and products are indicated. **(B)** Spliceosome assembly assays corresponding to in vitro splicing mixes as in (A) incubated for the indicated times and treated with 5 mg/ml of heparin for 10' at room temperature prior to complex analysis by native agarose-polyacrylamide gel electrophoresis. The positions of heterogeneous nuclear RNP (H), A, B and C complexes are shown. **(C)** Spliceosome assembly analysis of WT or AG RNAs corresponding to exon 6, the 3' 68 nucleotides of intron 5 and the 5' 25 nucleotides of intron 6. Analysis were as in (B) except that electrophoresis was carried out in 1.5% low melting point agarose gels. The identities of the complexes are indicated. **(D)** Psoralen-mediated crosslinking to detect base pairing interactions between U1 and U2 snRNAs with WT or AG exon 6 splice sites. Splicing mixes as in (C) in the presence of psoralen (22.2 mg/ml), using mock-treated HeLa cell nuclear extracts or extracts in which U1, U2, U4 or U5 snRNAs were inactivated by RNase H-mediated digestion. The reactions were incubated for 20' at 30ºC and irradiated with long-wave UV light. RNA was isolated and analyzed by electrophoresis in denaturing 6% polyacrylamide gels. The positions of the free RNA and the U1/RNA and U2/RNA crosslinked species are indicated.
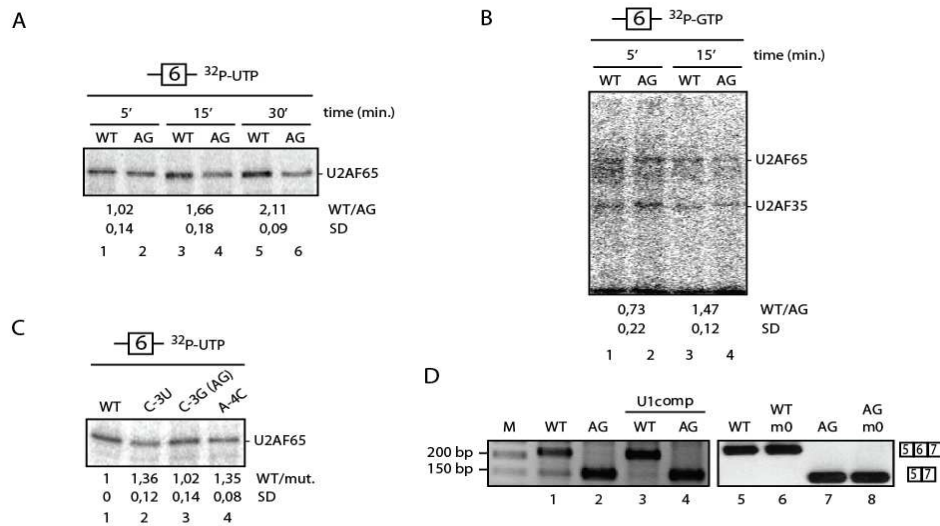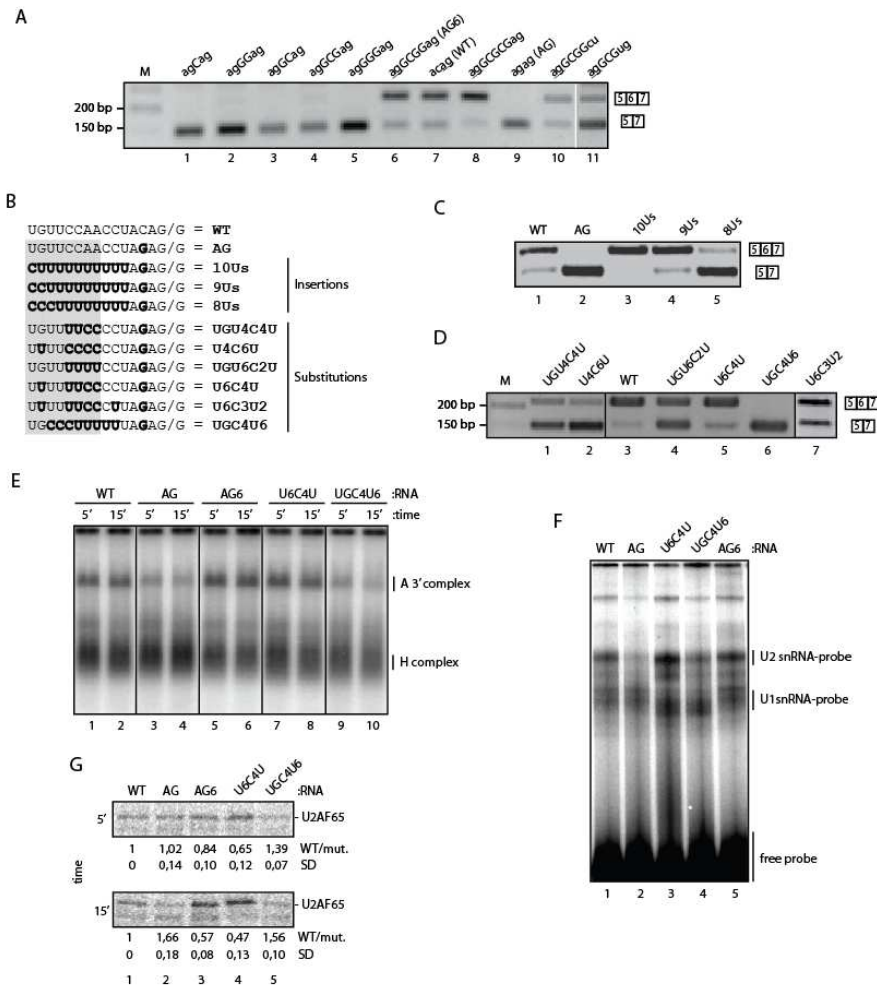
Fig. II. 3.

The AG mutation does not inhibit initial recognition of the 3'ss by the U2AF heterodimer. (A) Crosslinking-immunoprecipitation of U2AF65 to assess binding to wild type and AG mutant RNAs. 32P-uridine-labeled wild type and AG RNAs comprising the 3′ 68 nucleotides of intron 5, exon 6, and 5' 25 nucleotides of intron 6 were incubated with nuclear extracts under splicing conditions for the indicated times. After UV irradiation and RNAse treatment, U2AF65 was immunoprecipitated using the MC3 monoclonal antibody (Gama-Carvalho et al., 1997). Immunoprecipitated materials were fractionated by sodium-dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis. The position of RNA-crosslinked U2AF65 is indicated. The ratios of U2AF65 crosslinking to WT versus AG probes (WT/AG) at 5', 15' and 30' and the standard deviation (SD) corresponding to four independent experiments are indicated below the corresponding lanes. (B) U2AF35 crosslinking-immunoprecipitation to wild type and mutant AG. Assays were carried out as in (A) for the indicated times points using α-32P-guanosine-labeled RNAs. U2AF35 was immunoprecipitated using a rabbit polyclonal antibody (Zuo and Maniatis, 1996). The positions of the U2AF heterodimer components crosslinked to the RNAs are indicated. The ratio of U2AF35 crosslinking to wild type versus AG RNAs at 5 and 15 minutes and the standard deviation (SD) for three independent experiments are indicated below the corresponding lanes. (C) U2AF65 crosslinking-immunoprecipitation to various 3' splice site mutants. Assays were carried out as in (A) with the indicated mutant RNAs. Ratios and standad deviations correspond to three indpendent experiments. (D) Improving exon definition does not relieve the effect of the AG mutation. The indicated mutant minigenes were transfected and RNAs analyzed as in Figure II. 1. U1comp indicates a mutant minigene with enhanced complementarity between U1 snRNA and the intron 6 5' splice site (positions –2 and –3 from the exon 6-intron 6 boundary were mutated to A and C, and positions +7 and +8 in intron 6 were mutated to A and T). The m0 mutant replaces the uridine-rich PTB-binding silencer in exon 6 (UUUGUCUUCUUCUUUU) by a random sequence (AAUGCACACUCACCAG) (Izquierdo et al, 2005).

Fig. II. 4. Suppression of the AG mutation by increasing the distance between AGs and by stronger Py-tracts. (A) AG-AG insertion mutants. Analysis of the indicated mutants were as in Figure 1. The first ag indicates the position of the AG created by the ALPS mutation, the second corresponds to the natural intron 5 3' splice site. Underlined ags are those used when exon 6 is included (determined by sequencing of the spliced products). **(B)** The 3'ss sequence of WT, AG and Py-tract mutants. Mutated/inserted nucleotides are in bold. / indicates the intron/exon boundary. The box summarizes the conclusion from the results of parts C and D that at least six uridines are required to promote exon 6 inclusion in the AG mutant. **(C)** Analysis of the effects of uridine insertions in the AG mutant, analyzed as in A. **(D)** Analysis of mutants showing different arrangements of uridine-rich regions in the Py-tract, analyzed as in A. **(E)** Spliceosome assembly analysis of AG-AG insertion and Py-tract mutants, analyzed as in Figure 2C at the indicated times of incubation. **(F)** Psoralen-mediated crosslinking analysis for detection of U1/Fas and U2/Fas base pairing interactions, performed as in Figure II. 2D with wild type or the indicated mutant RNAs. **(G)** U2AF65 crosslinking-immunoprecipitation assays were carried out as in Figure II. 2A using wild type or the indicated mutant RNAs.

Appendix

hSlu7 does not rescue exon 6 inclusion in the AG mutant.

hSlu7 is a second step factor thought to be involved in the spliceosome reorganization in the transition between the first and the second steps of the splicing reaction (Chua and Reed, 1999b). Moreover, hSLu7 has been described to establish the AG used as a acceptor site when two neighboring AGs are present (Chua and Reed, 1999b).

To test if hSlu7 could have a role in the splicing inhibition promoted by Fas intron 5 AG duplication, we either overexpressed or knocked down this protein in HeLa cells and monitored Fas WT and AG alternative splicing (Fig. II. 5). hSlu7 knockdown promoted Fas WT exon 6 skipping. However, changes in hSlu7 levels failed to relieve the inhibitory effect of the AG mutation. This result is consistent with our in vitro data showing that the AG inhibits splicing before the first catalytic step.
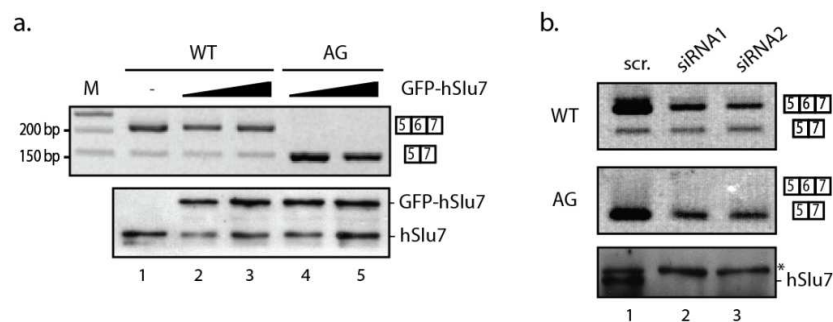
Fig. II. 5. hSlu7 does not rescue exon 6 inclusion of Fas AG minigene.(A) Fas AG and WT minigenes were cotransfected with 5 or 10 µg GFP-hSlu7 with ExGen 500 (Fermentas) into 50% confluent HeLa cells. Cytoplasmic RNA was extracted after 48 hours and analyzed by semi-quantitative RT-PCR. **(B)** Cells were transfected with Lipofectamine 2000 (Invitrogen) following the manufacturer's protocol with two different siRNAs against hSlu7 (50 pmol/mL siRNA) (siRNA1: 5'-auuguagccauuccaccgaucccu-3' and siRNA2: 5'-auaucuucucgaauccugagauuc-3') (Invitrogen). After 48 hours, a second round of transfection was carried out, followed by transfection of the WT and AG minigenes as in Fig. II. 1A. Samples were analyzed as described above.

WT and AG RNA oligonucleotides pulldowns do not show differential binding of proteins.

To determine if we could detect proteins binding differentially in the WT or mutant 3'ss, we carried out RNA-affinity chromatography. WT and AG RNA oligonucleotides encompassing the last 27 nucleotides of intron 5 and the first nucleotide of exon 6 (Fig.

II. 5A) were linked to agarose beads and incubated with HeLa nuclear extracts under splicing conditions. Beads were washed with increasing KCl concentrations with an elution step prior to every wash and eluted proteins were analyzed by SDS-PAGE followed by Coomassie-blue staining (Fig. II. 6B). We did not detect bands with striking differences in bands intensity between the WT and the AG oligonucleotides.

A

WT: UAAAAUGUCCAAUGUUCCAACCUACAG**G**

AG: UAAAAUGUCCAAUGUUCCAACCUA<u>G</u>AG**G**

B

C
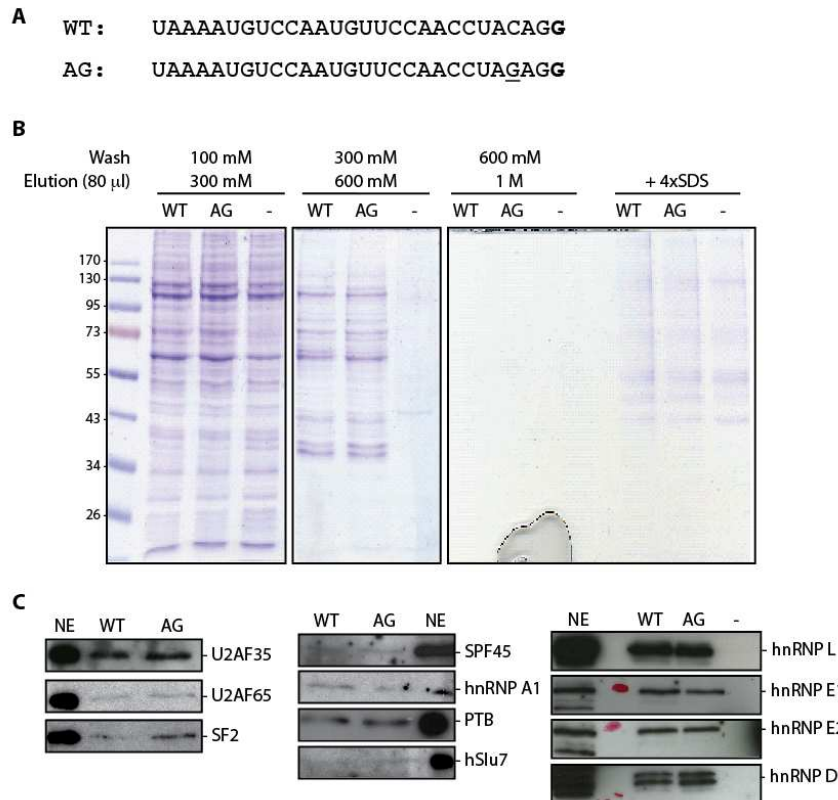
Fig. II. 6. Analysis of Proteins Bound to Fas WT and AG 3'ss by RNA-affinity chromatography.(A) WT and AG oligonucleotides corresponding to sequences -27 in intron 5 to the first nucleotide in exon 6 (in bold) were used for RNA-affinity chromatography. **(B)** The RNA oligonucleotides covalently bound to agarose beads or beads alone (-) were incubated with HeLa cell nuclear extract under splicing conditions. Beads were washed three times with the indicated KCl concentrations and eluted with 80 μl of buffer with the same salt concentration used for the next round of washes, as described in Materials and Methods. Eluted proteins ere analyzed by SDS-PAGE and Coomassie-Blue staining. **(C)** Western blotting analysis to detect specific proteins in the 600 mM KCl elution fractions (WT, AG and -). Nuclear extracts (NE) were used as positive control for the Western blotting.

However we cut several bands showing slight differences and analyzed them by mass spectrometry. As suspected, the results did not show any protein with specific affinity for any of the RNA fragments (data not shown). We performed western blotting analysis of the eluted fractions at 600 mM KCl using antibodies specific for several splicing factors and hnRNP proteins (Fig. II. 6C). Again, we did not observe preferential binding of these proteins for any RNA oligonucleotide. Although we did detect interaction of splicing factors, such as U2AF heterodimer, with these RNA fragments, it

is very likely that the absence of a downstream 5'ss does not allow a proper recognition of these oligonucleotides as functional 3'ss not allowing the observation of differences between WT and mutant RNAs.

# Chapter III: The proto-oncogene ASF/SF2 promotes the anti-apoptotic form of Fas

Results

The proto-oncogene SF2 regulates Fas alternative splicing to promote the anti-apoptotic isoform

Fas exon 6 codes for the transmembrane domain of the death receptor. Inclusion of this exon leads to the expression of the membrane-bound form, which can induce apoptosis upon binding to the Fas ligand (FasL) (Ruberti et al., 1996). On the other hand, skipping of exon 6 generates a mRNA encoding a soluble form of Fas which is still able to interact with the ligand but is incapable to transduce the apoptotic signal into the cell (Cheng et al., 1994; Cascino et al., 1995) (Fig. 10).

Recently, the SR protein SF2 has been classified as a proto-oncogene (Karni et al., 2007). High levels of SF2 have been observed in various human tumors, and mouse fibroblasts slightly overexpressing this factor were able to form sarcomas in nude mice, revealing the tumorigenic activity of this SR protein (Karni et al., 2007).

To determine if SF2 promotes the generation of the anti-apoptotic form of Fas by causing exon 6 skipping which could contribute to SF2 proto-oncogenic activity, we performed siRNA-mediated knockdown of SF2 and monitored changes in alternative splicing of a Fas minigene, encompassing Fas genomic sequence from exon 5 to 7, with a single point mutation at intron 5 pY tract (Fas C-11A), which shows a high degree of exon 6 skipping, thus allowing the detection of a positive effect on exon 6 inclusion (Fig. III.1A). Interestingly, SF2 knockdown led to increased levels of exon 6 inclusion, indicating that this SR protein does promote the generation of the anti-apoptotic form of Fas.
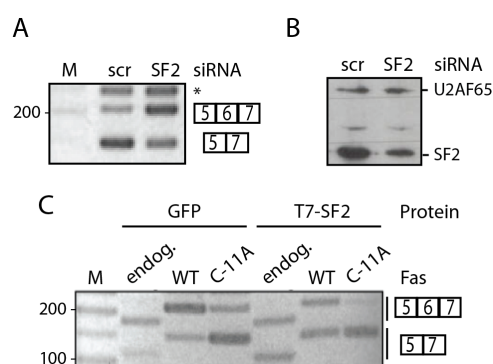
Fig. III. 1. SR and hnRNP proteins affect differentially Fas exon 6 inclusion.(A) HeLa cells were treated with 50 pmol/mL of SF2 siRNA or control siRNAs (scr) for 72 h and cytoplasmic RNA analyzed by semi-quantitative RT-PCR. * indicates an unspecific amplification product. **(B)** Protein levels of SF2 knockdown were measured by western blotting using antibodies against SF2 and U2AF65. **(C)** Effect of SF2 overexpression on alternative splicing of endogenous Fas and Fas WT and C-11A minigenes, analyzed as in (A).

To further study the effect of SF2 in Fas exon 6 inclusion, we monitored the effect of overexpressing SF2 on this alternative splicing event in either the endogenous gene

or cotransfecting with WT or C-11A minigenes. Consistent with the knockdown data, SF2 overexpression led to Fas exon 6 skipping in all the cases (Fig. III.1C).

The fact that SF2 induces the anti-apoptotic form of the Fas receptor might be one of the mechanisms by which SF2-overexpressing tumor cells escape apoptosis.

SF2 RRM2 is necessary and sufficient to promote Fas exon 6 skipping

SR proteins are characterized by a modular structure containing one or two RNA recognition motifs (RRM) and a C-terminal serine/arginine-rich (RS) domain. SF2 contains a N-terminal canonical RRM, while the second RRM lacks crucial residues in the RNP1 and RNP2 residues required for RNA binding. Moreover, SF2 RRM2 binds RNA through a tryptophan within a conserved SWQLKD motif in helix α1 and several amino acids present in strand β2 and a histidine on loop 5 (Tintaru et al., 2007), while canonical RRMs bind RNA through the central strands of the β-sheet (e. g. β1 and β3).
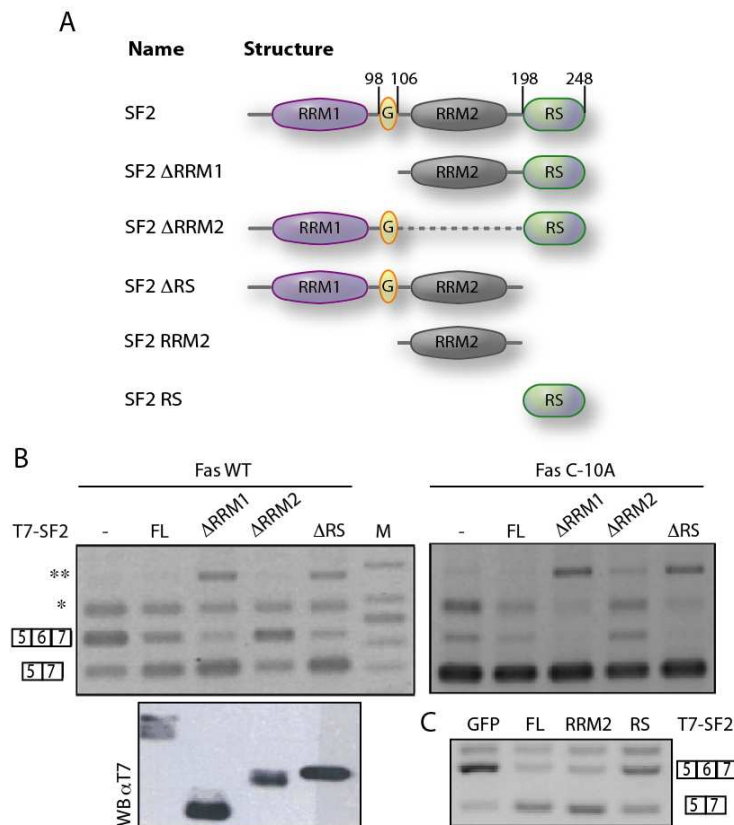
Fig. III. 2. Role of SF2 structural domains in Fas exon 6 alternative splicing regulation.(A) Structures of the SF2 domain-deletion mutants (Caceres and Krainer, 1993). RRM, RNA recognition motif; ΨRRM, pseudo-RRM (RRM2); RS, RS domain; G, Glycine patch **(B)** HeLa cells were co-transfected with T7-tagged SF2 domain-deletion mutants and Fas WT minigene. Cytoplasmic RNA was analyzed after 48 h by semi-quantitative RT-PCR. Levels of SF2 mutants overexpression were monitored by western blotting against the T7-tag. * and ** indicate an unspecific band and an unidentified amplified product **(C)** SF2 single domains (RRM2 or RS domain) were co-transfected with Fas WT minigene and the RNA analyzed as in (B).

To determine which are the SF2 structural requirements to cause exon 6 skipping, we co-transfected either Fas WT or Fas C-10A -which shows a high degree of skipping- minigenes with T7-tagged SF2 deletion mutants (Caceres et al.,

1997) (Fig. III. 2). As full length (FL) SF2, SF2 deletion mutants missing either RRM1 or the RS domain not only were still able to promote exon 6 skipping, but their activity appeared to be enhanced. On the other hand, a construct lacking the non-canonical RRM2 did not affect Fas splicing pattern (Fig. III. 2A).

To further explore the role of SF2 domains, we co-transfected Fas WT with either the RRM2 or the RS domains. Consistently, RRM2 alone was still able to cause exon 6 skipping, while the RS domain alone did not quantitatively change the inclusion/skipping ratio (Fig. III. 2B), indicating that the RRM2 of SF2 is necessary and sufficient to promote exon 6 skipping in these co-transfection assays.
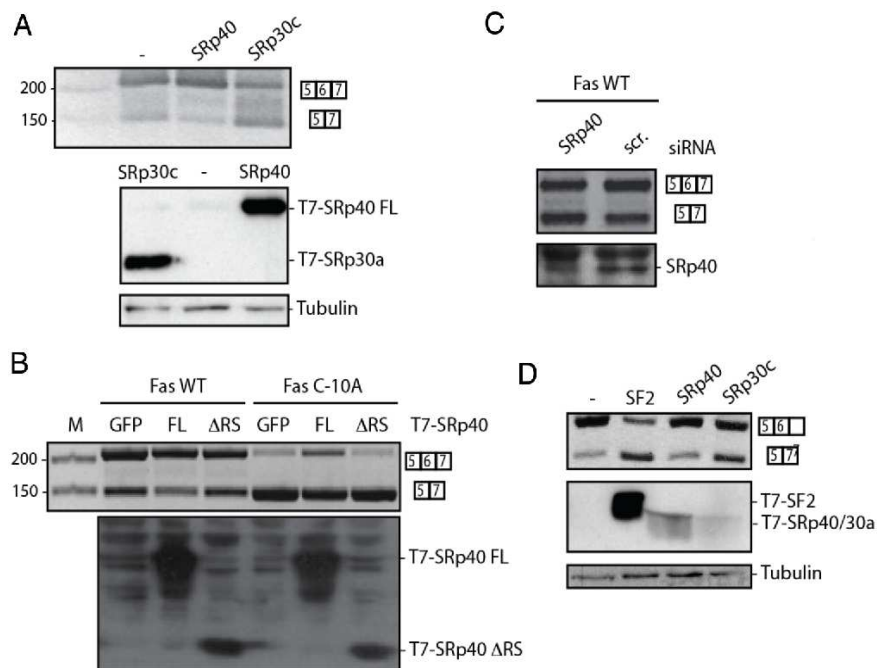
Fig. III. 3. SR proteins have antagonistic effects on Fas exon 6 alternative splicing.(A) SRp40 and SRp30c cloned in the pCG T7 vector were overexpressed and Fas WT minigene splicing pattern analyzed as in Fig. III. 2B. Transfected SR proteins expression levels were monitored by western blotting against the T7-tag. Tubulin western blotting as a loading control is shown. **(B)** SRp40 full length (FL) and DRS domain were transfected with Fas WT and Fas C-10A minigenes. Cytoplasmic RNA and SRp40 overexpression levels were analyzed as in (A). **(C)** Knockdown of endogenous SRp40 was carried out by transfecting HeLa cells with a SRp40-specific siRNA (5′- CCCGUUCCCGUAGUCGCAAAUCUUA-3′). 48 h later Fas WT minigene was transfected and total RNA extracted after 24 h. Fas alternative splicing was monitored by RT-PCR. The levels of SRp40 knockdown were analyzed by western blotting using an antibody against SRp40 (kind gift of D. R. Cooper) (Patel et al., 2005). **(D)** T7-tagged second RRMs of SF2, SRp30c and SRp40 were co-transfected with Fas WT minigene. Fas alternative splicing pattern was monitored as in (A). Overexpressed protein levels were analyzed by western blotting against the T7 tag. Tubulin western blotting is used as a loading control.

To test if other SR proteins containing two RRMs can induce exon 6 skipping, we co-transfected full length SF2, SRp40 or SRp30c together with Fas WT minigene. The latter showed similar effects on Fas alternative splicing than SF2, while SRp40 promoted exon 6 inclusion, indicating that SR proteins have antagonistic effects on this alternative splicing event (Fig. III. 3A). SRp40 RS domain was necessary to observe this effect in either Fas WT or C-10A minigenes (Fig. III. 3B) while knockdown of this protein led to exon 6 skipping (Fig. III. 3C).

To further study the role of non-canonical RRMs on exon 6 inclusion, we cloned both SRp40 and SRp30c second RRMs on the pCG T7 vector, as SF2 RRM2, and analyzed their effects on Fas WT minigene. Similarly to SF2 RRM2, SRp30c RRM2, even at very low levels of expression, was sufficient to promote exon 6 skipping while SRp40 pseudo-RRM had no effect on the inclusion of this exon (Fig. III. 3D).

The similar effects of both SF2 and SRp30c second RRMs are in agreement with the high level of similarity between these domains, evaluated by multiple sequence alignment of different pseudo-RRMs using the CLUSTALW software (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (Fig. III. 4A). In contrast, SRp40 RRM2 was more divergent, consistent with the lack of activity of this domain by itself.

Differences in the primary sequence of a protein can lead to subtle changes in the polypeptide folding and, consequently, to changes in RRM-protein and/or RRM-RNA interactions. To determine structural differences between SRp40 and SF2/SRp30c second RRMs, we modeled both SRp40 and SRp30c second RRMs using the SF2 RRM2 structure (pdb: 2o3d, (Tintaru et al., 2007) as the template and the SWISS-MODEL tool (http://swissmodel.expasy.org) and aligned them using MultiProt (http://bioinfo3d.cs.tau.ac.il/MultiProt/) (Shatsky et al., 2004) (Fig. III. 4B). As expected, the main differences between SRp40 and SF2/SRp30c were located in the unstructured regions (e. g. loops) of the RRM fold. For example, the loops between β2-β3 and β3'- β3" had different length and charge. We then generated chimeras of SF2 and SRp40 pseudo-RRMs to find out if the loops were playing a role in SF2-mediated Fas exon 6 skipping (Fig. III. 4C). While replacement of SF2 loops by SRp40 loops led to a loss of function of SF2 RRM2, the converse mutants did not produce a SRp40 RRM2 variant that could induce exon 6 skipping (Fig. III. 4D). Moreover, loop I mutants (LI) accumulated at low amounts, possibly due to reduced stability of the chimeric protein. These data suggest that all the loops in SF2 RRM2 are required to cause exon 6 skipping but individually they are not able to confer this activity to another pseudo-RRM.
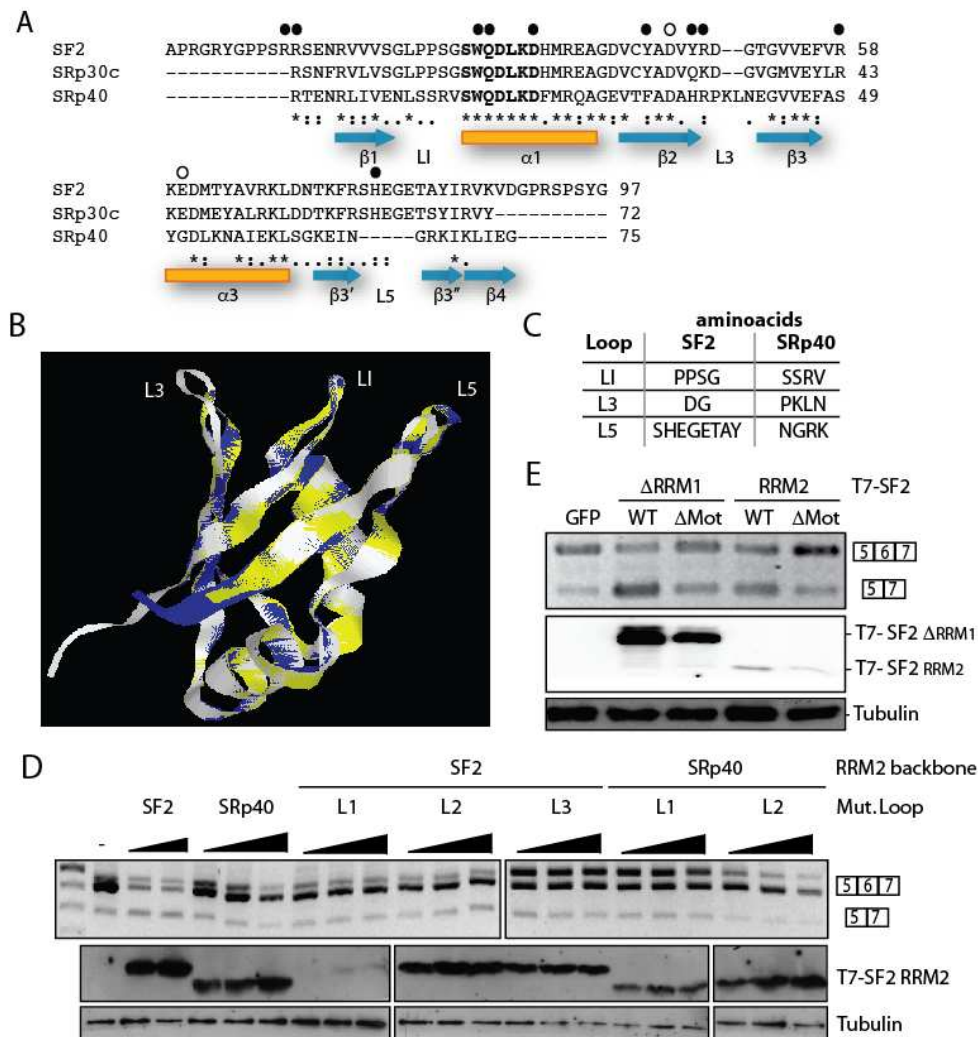
Fig. III. 4. Multiple sequences in SF2 are required for Fas alternative splicing regulation.(A) SF2, SRp30c and SRp40 RRM2 aminoacid sequences were aligned using CLUSTALW. Filled and open circles point to residues in SF2 directly involved or not involved in RNA binding, respectively (Tintaru et al., 2007). "*" indicates that the residue is conserved. ":" and "." mean that conserved and semi-conserved substitutions have been observed, respectively. Yellow boxes and blue arrows indicate that aminoacids in those columns are located in a-helices or b-strands, respectively. **(B)** Structure alignment of SF2 RRM2 (blue) (pdb: 2o3d) (Tintaru et al., 2007) and the modeled SRp30c (yellow) and SRp40 (grey) pseudo-RRMs. **(C)** Table with the aminoacid sequences of loops 1 (L1), 3 (L3) and 5 (L5) of SF2 and SRp40. **(D)** Two increasing amounts of SF2-SRp40 chimeric RRMs were co-transfected with Fas WT minigene and the RNA analyzed as in Fig. III. 1A. The levels of overexpression were monitored by T7-tag western blotting, using tubulin as loading control. **(E)** SF2 ΔRRM1 and RRM2 WT or with the mutated SWQDLKD domain were overexpressed and analyzed the effect on Fas minigene alternative splicing.

We also generated a mutant of SF2 RRM2 SWQDLKD motif and analyzed its ability to promote Fas exon 6 skipping. This motif is present in α-helix1 and can be found in the second RRM of all the two RRM-containing SR proteins. This motif has been described to be required for binding to RNA (Tintaru et al., 2007), recruitment of SF2 to the nuclear stress bodies upon heat shock (Chiodi et al., 2004), SF2-mediated translation enhancement (Sanford et al., 2004) and regulation (Sun et al., 2010), and SF2-mediated alternative 5'ss selection (Dauksaite and Akusjarvi, 2004) and splicing repression (Dauksaite and Akusjarvi, 2002). Consistent with its important role in several SF2 activities, upon mutation of the motif (e.g SWQDLKD to ALQNVKN) SF2 RRM2 was no longer able to promote exon 6 skipping (Fig. III. 4E). It is very likely that the mutation does not only impair SF2 RRM2 ability to bind RNA but it may also affect the overall folding of the RRM.

Taken together, the data suggest that different surfaces, structured and unstructured, of SF2 are important for activity, which could be based upon RNA binding and protein interactions. Further mutational studies are required to delineate the specific residues required for SF2 function.

SF2 acts indirectly or through several sequences in Fas

Alternative inclusion or skipping of coding sequences is regulated by the presence of cis-regulatory sequences known as splicing silencers or enhancers, which bind trans-acting factors that will ultimately promote changes in alternative splicing.

SF2 could be promoting exon 6 skipping by either enhancing or inhibiting intron 6 or 5 splicing. To determine which was the intron regulated by SF2, we first monitored the effect of either SF2 overexpression or knockdown on splicing of single intron-containing minigenes (Fig. III. 5). Intron 6, with most of its sequence deleted as described previously (50-1102) (Izquierdo et al., 2005) to facilitate the analysis, was constitutively spliced and was not affected by changes in SF2 levels. A two-intron containing minigene in which this same sequence had been deleted was still sensitive to both SF2 overexpression and knockdown (data not shown) indicating that the sequence deleted n intron 6 is not essential for the response to SF2. In contrast, intron 5 was removed more efficiently upon SF2 knockdown and splicing was inhibited upon overexpression of SF2, indicating that removal of intron 5, rather than removal of intron 6, is impaired by this SR protein.
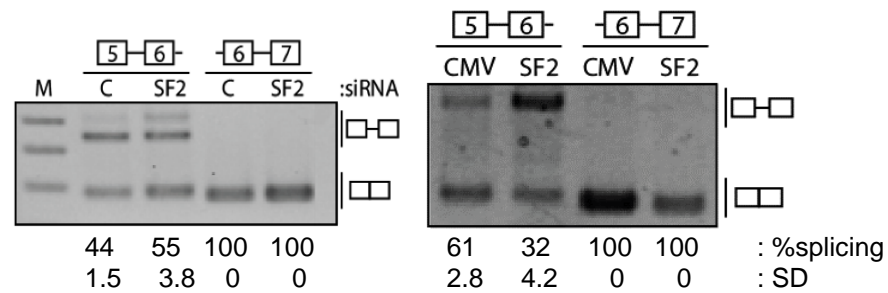
Fig. III. 5. SF2 inhibits intron 5 splicing. The splicing pattern of WT minigenes containing single introns was analyzed after SF2 knockdown (left panel) or overexpression (right panel). The diagonal line in the minigene scheme represents a deletion of part of intron 6, which does not alter Fas splicing (Izquierdo et al., 2005). Percentage of splicing and standard deviation are shown.

To determine the cis-acting sequences mediating the effect of SF2 RRM2 on intron 5 splicing, we looked for putative SF2 binding sites in Fas two intron-containing minigene using ESE finder tool (http://*rulai.cshl.edu/tools/ESE/*) (Cartegni et al., 2003),(Smith et al., 2006) (Fig. III. 6A). We found potential binding sites in both introns and in exons 5 and 7. Next, we generated deletion/substitution mutants of Fas minigene, with special emphasis on those sequences predicted to bind SF2.

We first monitored the effect of SF2 in Fas-α-globin chimeric minigenes, because α-globin exon 2 has many putative SF2 binding sites (data not shown) but its inclusion is not affected by increasing SF2 levels (Fig. III. 6B). A minigene in which Fas exon 5, containing two putative SF2 binding sites, was substituted by α-globin exon 1 was still sensitive to SF2 overexpression (Fig. III. 6C). Similarly, SF2 promoted exon 6 skipping of Fas mutants in which the predicted binding sites in intron 5 had been deleted (del1=22-47; del2=62-86) (Fig. III. 6C). We next tested the effects of SF2 overexpression in α-globin minigenes where exon 2 and the neighboring splice sites had been substituted by Fas exon 6 and its adjacent splice sites (Adj. Fas ss) (Fig. III. 6D) (Bonnal et al., 2008). Similar to the wt α-globin gene, increasing SF2 levels did not alter Adj. Fas ss alternative splicing pattern suggesting that Fas exon 6 and neighboring sequences do not have the necessary sequence elements for SF2-mediated exon skipping. However, when also the distal splice sites of Fas were present in the minigene (Bonnal et al., 2008), SF2 overexpression promoted exon 6 skipping (Fig. III. 6D). These data suggest that Fas distal splice sites might mediate SF2 effects, but when we replaced a-globin by Fas distal splice sites, SF2 was still not able to promote a-globin exon 2 skipping (data not shown).
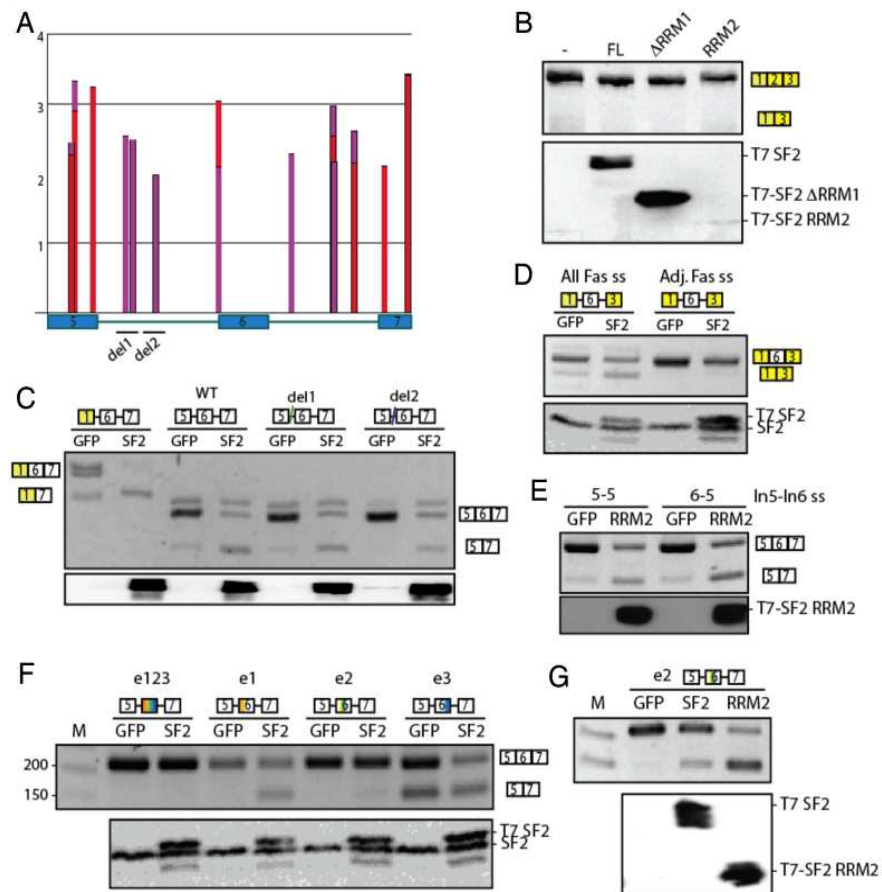
Fig. III. 6. SF2 promotes exon 6 skipping indirectly or by binding to multiple sequences in Fas.(A) Putative SF2 binding sites in Fas genomic sequence from exon 5 to +43 nucleotides in exon 7, with a partial deletion in intron 6 (nucleotides 50 to 1020) (X axis), predicted by ESEfinder 3.0. Blue boxes indicate exons. del1 and del2 show the approximate deleted sequences in mutants used in **(C)**. Red and pink bars indicate SF2 binding sites predicted by two different, but very similar, weight matrixes in ESEfinder. Y axis represents the score of each putative binding site. Effect of SF2 and SF2 deletion mutants overexpression in α-globin **(B)**, Fas intron 5 deletion mutants **(C)**, α-globin-Fas chimeric RNAs **(C)** and **(D)** and Fas splice site **(E)** and exon 6 **(F)** and **(G)** mutants.

To determine if the distal 3'ss is necessary to observe SF2 effects, we introduced intron 5 3'ss instead of intron 6 3'ss (mutant 5-5) or swapped intron 5 and intron 6 splice sites (mutant 6-5) (Bonnal et al., 2008) (Fig. III. 6E). If intron 6 3'ss was mediating exon 6 skipping, we would expect that overexpression of SF2 had no effect on mutant 5-5 and promoted exon 6 inclusion in the mutant 6-5. However, SF2 overexpression similarly promoted exon skipping in both mutants, indicating that intorn 6 3'ss does not mediate SF2 regulation of exon 6 alternative splicing.

Finally, although no binding sites for SF2 were predicted in Fas exon 6, we tested the effects of overexpressing this SR protein on mutants in which either the whole exon

6 (e123) or the first (e1), second (e2) or third (e3) thirds had been substituted by a random sequence (Izquierdo et al., 2005; Bonnal et al., 2008) (Fig. III. 6F). Only when either the whole exon or its middle part, which binds PTB, were mutated, increased levels of SF2 had no effect on Fas exon 6 inclusion. However, overexpression of SF2 RRM2, that has a stronger effect than the full length protein, promoted exon 6 skipping of the e2 mutant (Fig. III. 6G).

Collectively, these results suggest one of two possibilities. First, SF2 may act through two or more cis-acting sequences that we have not deleted simultaneously or, second, that the effect is indirect and that SF2 RRM2 may be inhibiting splicing in general. The fact that SF2 does not alter the splicing pattern of only those minigenes that show constitutive inclusion of the exon strongly suggests a general splicing inhibitory effect of SF2 that is only observed when the exon can be alternatively spliced.

SF2 might have a general inhibitory effect on splicing

To determine if SF2 and, specifically, its second RRM inhibit splicing in a more general manner, we tested the effect of overexpressing either SF2 full length, SF2 ΔRRM2 or SF2 RRM2 in two different minigenes.

Pax6 is a transcription factor containing paired–box and homeobox motifs that is specifically expressed in the developing eye and brain. Mutations in this gene account for 98% of aniridia cases, a congenital malformation of the eye that causes blindness (reviewed in Lee et al., 2008). Pax6 is alternatively spliced and the differential inclusion of exon 5a alters its DNA binding specificity (Epstein et al., 1994). In fact, constitutive inclusion of exon 5a caused a U to C transition at position -3 of intron 5 was described in patients suffering from aniridia (Epstein et al., 1994). SMN1, survival of motor neuron 1, is a component of the SMN complex, important for the proper translocation into the nucleus of most snRNPs (see Introduction). However, an SMN1 paralog, SMN2, encodes a truncated, partially functional protein due to skipping of its exon 7. SMN1 and SMN2 differential splicing pattern is due to a polymorphism at position 6 in exon 7 (Lorson et al., 1999).

We used minigenes to monitor the effects of SF2 and deletion mutants on both Pax6 mutant exon 5a (S. Bonnal, unpublished data) and SMN2 exon 7 (Lorson et al., 1999) alternative splicing (Fig. III. 7A). In the case of Pax6, which showed total exon 6 inclusion, SF2 was able to promote exon 6 skipping but neither SF2 ΔRRM2 or SF2 RRM2 alone were able to mimic the effect of the full length protein. However, similarly to what was observed with Fas exon 6, SF2 RRM2 caused exon 7 skipping in the SMN2 minigene, while SF2 full length had minimal effects and the construct lacking the

second RRM promoted exon 7 inclusion. These data suggest that SF2 can cause exon skipping through at least two different mechanisms, one in which the RRM2 is necessary and sufficient to promote this event (i. e. Fas and SMN2) and another one in which the second RRM is necessary but not sufficient to cause exon skipping (i. e. Pax).

To further determine if SF2 RRM2 performs a general inhibitory effect on splicing, we co-transfected SF2 with a model pre-mRNA, IgM-M, containing one single intron and two alternative 3'ss, one from the IgM gene and the other from an adenovirus intron (AdML) (Pacheco et al., 2006). Strikingly, SF2 promoted retention of the intron and inhibition of the use of both 3'ss, suggesting a general inhibitory effect (Fig. III. 7B).

Since we could not find any cis-acting sequence mediating SF2 effect, we tested the effects of tethering SF2 to Fas exon 6. We co-transfected SF2 fused to the bacteriophage MS2 coat protein (Dauksaite and Akusjarvi, 2002), which has high affinity for MS2 aptamers, with a Fas minigene containing two MS2 stem-loops in Fas exon 6 replacing PTB binding site, URE6 (Izquierdo et al., 2005) (Fig. III. 7C). Interestingly, tethering SF2 to the alternatively spliced exon 6 led to enhanced inclusion of this exon, the effect being totally dependent on the presence of MS2 aptamers. The activity of splicing regulators depends on the position of their binding site relative to the alternatively spliced exon, as reported for an increasing number of functional RNA maps drawn (reviewed in Corrionero and Valcarcel, 2009). Therefore, it is possible that SF2 binding to sequences upstream or downstream of Fas exon 6 might promote exon 6 skipping.

We also determined if SF2 acts cooperatively with PTB to promote exon 6 skipping (Fig. III. 7D), since URE6 mutant (e2) was less responsive to SF2 (Fig. III. 6F). Initially, we co-transfected SF2 full length with PTB-MS2 fusion protein (Gromak et al., 2003) and monitored Fas WT alternative splicing. As expected, both factors caused *per se* exon 6 skipping (Fig. III. 7D, bottom). However, we did not detect any synergistic effect when both proteins were co-transfected. When we performed the same experiments using the Fas minigene containing MS2 aptamers replacing URE6 sequence, both SF2 and PTB-MS2 caused exon 6 skipping. Surprisingly, co-transfection of both proteins did not result in increased Fas exon 6 skipping (Fig. III. 7D, top). This effect was dependent on SF2 RRM1, since SF2 ΔRRM1 caused skipping of the exon even in the presence of PTB-MS2.

These results show that SF2 is able to antagonize PTB-MS2 activity only when it is tethered to Fas exon 6 containing MS2 aptamers, but not when it exerts its effects by directly binding to URE6. These data suggests that the way PTB interacts with the RNA is important for SF2 activity. It is possible that PTB binding through the fused MS2

coat protein to Fas exon 6 leaves the four PTB RRMs free to render other kind of interactions. Then, SF2, maybe through its first RRM, might be able to interact with PTB and inhibit its activity. On the other hand, PTB direct binding to the RNA might occlude the structures required for this interaction, suggesting that it is not the same to tether PTB, a possibly other proteins, to the pre-mRNA than allowing a direct binding of the protein to its cognate sequence. This should be taken into account when studying protein function in alternative splicing.
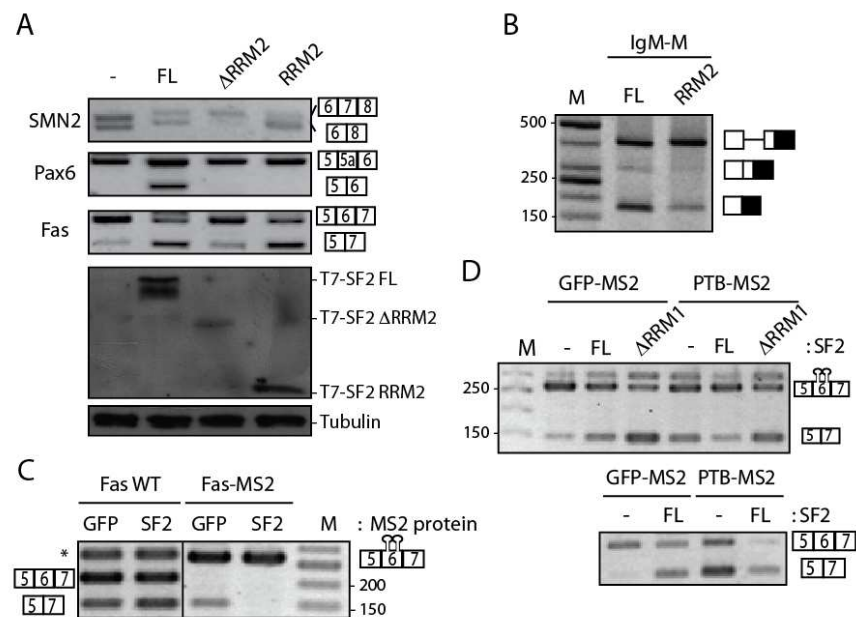
Fig. III. 7. SF2 RRM2 affects splicing. (A) Effect of SF2 FL, ΔRRM2 and RRM2 overexpression on SMN2 exon 7, Pax6 exon 5a and Fas minigenes alternative splicing, analyzed as in Fig. III.1A. Levels of SF2 overexpression were monitored as in Fig. III. 3D. **(B)** Effect of SF2 RRM2 overexpression on IgM-M minigene. White and black boxes correspond to IgM and AdML exons. The black line represents IgM intron 1. The alternative splicing pattern was analyzed using a forward primer specific for IgM exon M1 and a reverse primer for AdML exon 2 **(C)** Analysis of the effect of SF2-MS2 fusion protein (Dauksaite and Akusjarvi, 2002) in either Fas WT or Fas-MS2 (Izquierdo et al., 2005) minigenes. **(D)** PTB-MS2 was transfected in the absence or presence of SF2 FL or SF2 ΔRRM1, and the effects on Fas WT (bottom panel) and Fas-MS2 (top panel) minigenes were analyzed.

SF2 effects in in vitro assays

SF2 RRM2 promotes both Fas exon 6 and SMN2 exon 7 skipping and inhibits splicing of single-intron containing minigenes.

To determine if this effect is direct and, if so, to establish the molecular mechanism underlying this activity, we first performed in vitro splicing assays using in

vitro transcribed RNAs corresponding to Fas pre-mRNA sequences from exon 5 to nucleotide +25 of intron 6 and we incubated the reactions with increasing amounts of full length SF2 protein purified from HEK293 cells (kind gift of A. R. Krainer) (Sinha et al., 2010). Surprisingly, even at very high concentrations of SF2, the effect on Fas intron 5 splicing was minimal (Fig. III. 8A).
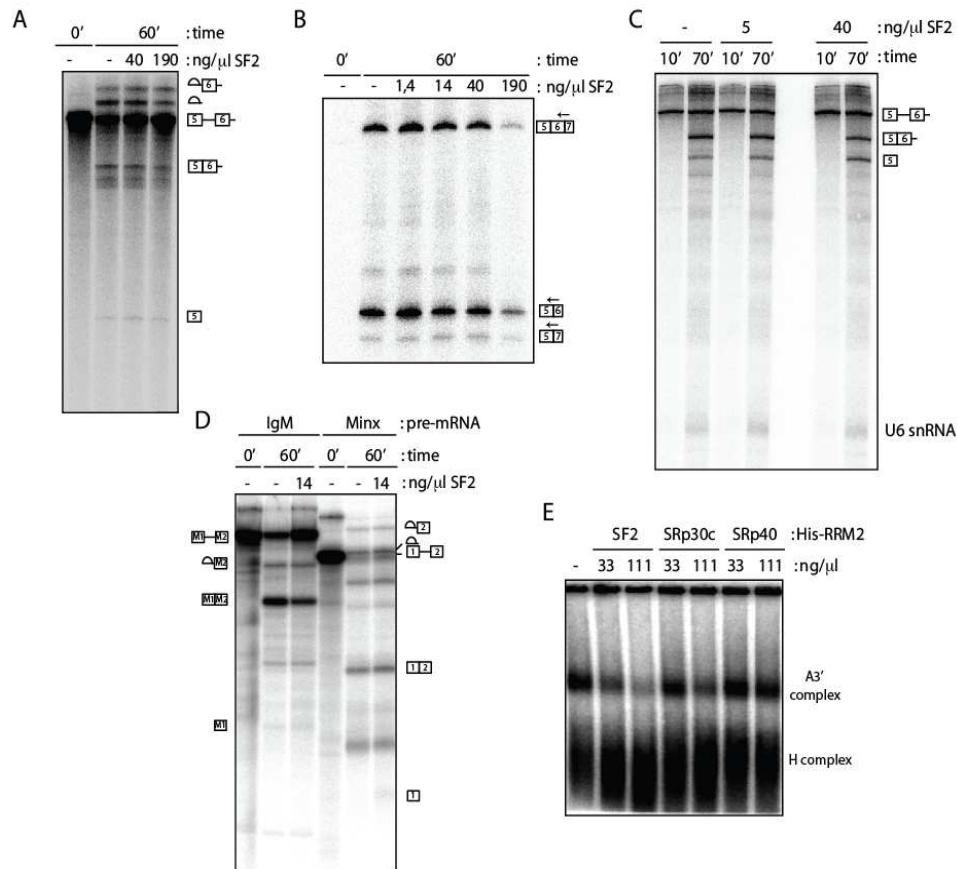
Fig. III. 8.Effects of SF2 in in vitro experiments. (A)[32]P-UTP uniformly labeled RNAs corresponding to exon 5 to 25 nucleotides of intron 6 were in vitro transcribed and incubated with HeLa cell nuclear extracts under splicing conditions (as described in Chapter II) in the absence or presence of ATP and absence or presence of recombinant His-SF2 (40 and 190 ng/μL) for 60'. The RNA species were purified, loaded in denaturing 13% polyacrylamide gels and exposed to a PhosphorImager screen. The pre-mRNA, splicing products and splicing intermediates are indicated. **(B)** In vitro transcribed RNA corresponding to Fas exon 5 to exon 7 (with deletion of intron 6 from positions 50 to 1102) was used as a substrate for in vitro splicing reactions in the absence or presence of ATP. Reactions were incubated in the absence or with increasing amounts of recombinant His-SF2 (1.4, 14, 40 and 190 ng/μL). **(C)** A PCR product corresponding the CMV promoter followed by Fas genomic sequence from exon 5 to 25 nucleotides of intron 6 was incubated with HeLa cell nuclear extracts under transcription-splicing coupling conditions (See Materials and Methods) in the absence or presence of SF2 (5 and 40 ng/μL). After 10', transcription was blocked with Actinomycin D and the reaction let proceed for other 60'. Then, the RNA species were purified, loaded in a denaturing 8%

polyacrylamide gel and exposed to a PhosphorImager screen. **(D)** IgM and AdML radiolabeled RNAs were used as substrates in splicing reactions carried out as in (A) for 0' or 60' in the absence or presence of 14 ng/μL SF2. The RNA species were analyzed as in (A). **(E)** Spliceosome assembly assays were performed using Fas WT RNA corresponding to exon 6, the last 68 nucleotides of intron 5 and the first 25 nucleotides of intron 6. The reactions were incubated in the absence or with 33 and 111 ng/μL of recombinant (*E. coli*) His-tagged SF2, SRp30c or SRp40 second RRMs for 15'. Afterwards, 5 mg/mL were added to the reactions and after 10' at room temperature, 6xDNA loading dye was added and the samples were loaded in a 1.5% low melting point agarose gel. The identities of the complexes are indicated.

To improve the sensitivity of the assay we performed splicing switch experiments as described (Izquierdo et al., 2005). In vitro transcribed non-radiolabeled RNA corresponding to Fas genomic sequences between exons 5 and 7, with the partial deletion of intron 6 mentioned above, was incubated under standard splicing conditions and the products analyzed by primer extension using splice-junction oligonucleotides (Izquierdo et al., 2005). Again, SF2 full length did not show any direct effect on Fas exon 6 alternative splicing in vivo (Fig. III. 8B). The effect observed in the last lane may suggest a general inhibitory effect of splicing by SF2, although more experiments and a loading control are required to confirm this result.

SF2 and other SR proteins have been described to interact with the pol II CTD, possibly increasing its effective concentration near the newly synthesized pre-mRNA (Misteli and Spector, 1999; Sapra et al., 2009). An in vitro transcription-splicing functionally coupled system (Das et al., 2006), which was shown to facilitate splicing regulation by SR proteins (Das et al., 2007), was used to monitor the effects of SF2 on Fas intron 5 splicing (Fig. III. 8C). Using this assay, we could detect some inhibitory effect at a concentration of SF2 of 40 ng/μl. Strikingly, and similar to what was observed in Fig. III. 8A, the splicing defect seemed to affect preferencitally the second step of the splicing reaction (compare ratios of spliced mRNA/free exon 5). However, when U2 snRNP assembly on Fas intron 5 3'ss was analyzed as described in Chapter II, the presence of SF2 and SRp30c RRM2s had a strong inhibitory effect, while SRp40 RRM2 did not affect U2 snRNP recruitment (Fig. III. 8D). Importantly, SF2/SRp30c effects might be mediated by altering U1 assembly to the downstream 5'ss, thus altering the exon definition proces.

In vitro splicing experiments suggest a second step inhibition while spliceosome assembly assays a defect in complex A formation. These apparent contradictory results may be explained by differences between the full length protein and the RRM2 alone and/or the presence of post-translational modifications in SF2 full length purified

in HEK293 cells absent in SF2 RRM2, expressed in *E. coli*. Also, the probes' sequences are different. Further studies will be required to determine if, first, the effect in Fas is direct and, second, how SF2 RRM2 acts as a general splicing inhibitor.

To further investigate the possible inhibitory effect of SF2, we had previously determined that SF2 RRM2 overexpression inhibits in vivo splicing of IgM-M, containing IgM genomic sequence followed by AdML exon 2 3'ss (Fig. III. 7B). For this reason, we tested the effects of SF2 full length on either IgM or AdML in in vitro splicing assays (Fig. IIIE). Interestingly, SF2 strongly inhibited IgM splicing at 14 ng/μl while AdML was not affected. This result suggest that SF2 RRM2 might be inhibiting IgM-M splicing in vivo by affecting IgM 5'ss recognition. However, it is also possible that higher levels of SF2 could inhibit AdML in vitro splicing. It will be of interest determine if SF2 RRM2 also affects IgM in vitro splicing and, thus, use this pre-mRNA to study SF2 RRM2 mechanism.

The mTOR pathway does not mediate SF2 effects on Fas exon 6 alternative splicing

The tumorigenic effect of SF2 is partially caused both by altering the splicing pattern of pro-apoptotic/proto-oncogenic genes (Karni et al., 2007; Moore et al., 2010) and by activating of the mammalian target of rapamycin (mTOR) pathway (Karni et al., 2007; Karni et al., 2008) by a still unidentified mechanism. The mTOR pathway is activated in response to nutrients, stress and growth factors like insulin. It is thought to be a master switch between anabolic and catabolic metabolisms. Activated mTOR, bound to raptor and other proteins forming the mTOR complex 1 (mTORC1), phosphorylates S6 kinase 1 (S6K1) and the eukaryotic initiation factor (eIF4E)-binding protein (4EBP1), inducing, among other cellular processes, an increase in the translation efficiency. Rapamycin blocks mTORC1 activity by directly binding to one of its components (reviewed in Fingar and Blenis, 2004).

To determine if the effect of SF2 was mediated through the activation of the mTOR pathway, we cotransfected NIH3T3 mouse fibroblasts with T7-SF2 and Fas WT minigene in the presence or absence of rapamycin (Fig. III. 9). In the absence of rapamycin, SF2 led to Fas exon 6 skipping concomitant with S6K1 p70 phosphorylation and the appearance of a new S6K1 isoform, S6K1 p31, as described in (Karni et al., 2007). Similarly, in the presence of 0,1 μM rapamycin (i.e. inhibition of mTOR pathway and S6K1 phosphorylation), we could still observe a quantitatively similar level of Fas exon 6 skipping and S6K1 p31 isoform upon SF2 overexpression. These data indicate that the effect of SF2 on Fas exon 6 alternative splicing is independent of the SF2-mediated activation of the mTOR pathway.
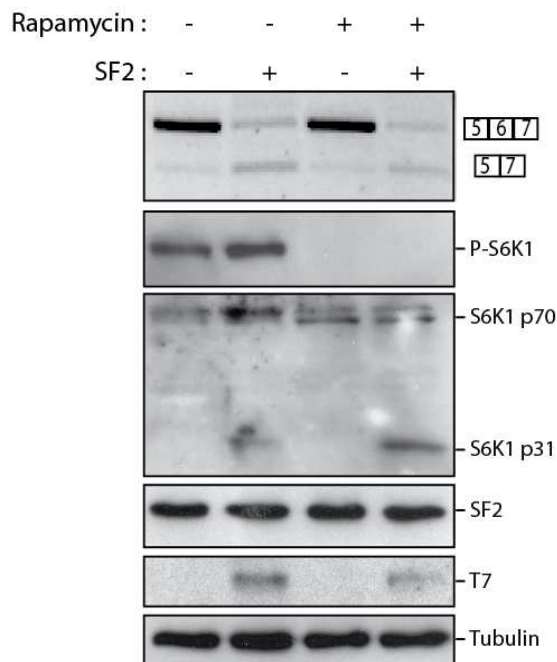
Fig. III. 9. SF2 activation of mTOR pathway is not involved in Fas exon 6 alternative splicing regulation. NIH3T3 cells were transfected with SF2 FL or the empty vector and Fas WT minigene. After 48 h, cells were treated with 0.1 μM of the mTOR pathway inhibitor rapamycin or DMSO as a control for 5 h. Fas splicing pattern was analyzed as in Fig. III. 1A. Western blots to monitor mTOR pathway activation levels and SF2 overexpression were carried out by using the indicated primary antibodies.

SF2 may contribute to exon 6 skipping after Fas receptor activation

Fas receptor activation can be mimicked by treating cells with the agonistic anti-Fas antibody CH11 and can be monitored by determining the cleaved levels of the caspase-3 and -7 target protein poly(ADP-ribose) polymerase (PARP) (reviewed in Earnshaw et al., 1999). It has been previously shown that CH11 treatment of Jurkat cells promotes the anti-apoptotic form of Fas, by enhancing exon 6 skipping (Izquierdo, 2008a). This effect could be mediated by caspase-mediated cleavage of U2AF65 that would generate a dominant negative form of this splicing factor (Izquierdo, 2008a). However, many other splicing regulators are cleaved upon Fas receptor activation, including several hnRNPs (e.g. A0, A2/B1, A3, K), p54*nrb*, SF2 and SRp30c (Thiede et al., 2001).

To determine if SF2 could play a role in CH11-driven Fas exon 6 skipping, we determined the levels and the cleavage pattern of SF2 after treating Jurkat cells with 50 or 250 ng/mL CH11 for three hours (Fig. III. 10). Although we could not detect any SF2 cleavage product as described previously (Thiede et al., 2001), we observed an increase in the protein levels at 50 ng/mL CH11 and more obvious at 250 ng/mL CH11. In contrast, U2AF65 and PARP cleavage were similar at 50 and 250 ng/mL. We then determined the alternative splicing pattern of a transfected Fas WT minigene. Treatment with CH11 promoted Fas exon 6 skipping in a concentration-dependent manner concomitant with the increase on SF2 levels.

These data show that activation of the Fas receptor pathway by the CH11 antibody promotes exon 6 skipping and increases SF2 levels, suggesting a physiological role for SF2 during Fas-mediated apoptosis.
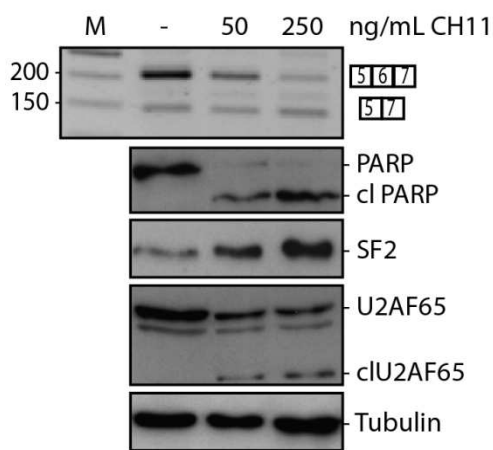
Fig. III. 10. Fas receptor activation promotes Fas exon 6 skipping concomitant with SF2 upregulation. Jurkat leukemia cells were incubated with increasing amounts of the Fas receptor agonistic antibody CH11 and Fas minigene alternative splicing was monitored as in Fig. III.1A. Activation of the receptor was monitored by following PARP and U2AF65 cleavage (clPARP, clU2AF65) by western blotting. SF2 levels were also determined by western blotting and tubulin was used as loading control.

Discussion

SF2 has been described to have a proto-oncogenic activity (Karni et al., 2007). Interestingly, knockdown of SF2 gave a quantitative increase in exon 6 inclusion, suggesting that this SR protein is a bona fide regulator of this alternative splicing event. Consistent with this, overexpression of SF2 led to exon 6 skipping and the generation of the anti-apoptotic form of the receptor.

SF2 has been extensively studied in the splicing field over the last two decades. In the last few years SF2 has become an important regulator of many different steps of gene expression (see Introduction). Consistent with its wide range of activities, SF2 levels are tightly regulated (Sun et al., 2010) and alterations in its expression have been associated with genome instability and aberrant apoptosis and cell growth (Li et al., 2005; Karni et al., 2007).

Interestingly, SF2 can be classified as a protooncogene, since its overexpression is sufficient to transform immortal rodent fibroblasts (Karni et al., 2007). How SF2 undergoes its tumorigenic activity is not fully understood, although recent studies are shedding some light on the mechanisms used by SF2 to enhance cell proliferation and escape apoptosis. For example, SF2 can promote the generation of specific alternatively spliced forms with anti-apoptotic activity (i. e. Bcl-xL) (Paronetto et al., 2007; Moore et al., 2010) and enhance translation by activating mTOR pathway (Karni et al., 2008; Michlewski et al., 2008), two important characteristics of tumor cells. Moreover, SF2 has been described to affect other processes such as miRNA

processing (Wu et al., 2010) or protein sumoylation (Pelisch et al., 2010) and it is very likely that other functions remain to be uncovered.

For many of these functions, the second RRM of SF2 has been shown to be necessary. This non-canonical RRM has been demonstrated to be necessary and sufficient on its own for the enhancement of sumoylation (Pelisch et al., 2010) and for the protein localization to the stress bodies upon thermal stress (Chiodi et al., 2004). Interestingly, our data shows that the RRM2 is also necessary and sufficient to promote Fas exon 6 skipping. We have shown that SRp30c RRM2 also promotes this event, consistent with its high similarity with SF2 second RRM. However, this effect is not specific of pseudo-RRMs, because SRp40 RRM2 does not alter exon 6 alternative splicing. We analyzed the structural requirements of SF2 RRM2 to regulate Fas splicing by generating SF2-SRp40 RRM2 chimeras. Any change in SF2 RRM2 sequences abrogated its activity, while introducing SF2-specific residues in SRp40 RRM2 did not confer any splicing regulatory activity to the chimeric protein. These data suggest that the integrity of SF2 RRM2 is necessary to alter Fas alternative splicing.

We also explored the cis-acting sequence motifs in Fas that mediate SF2 regulation of exon 6 inclusion. Although we mutated most of the sequences in the Fas minigene, these constructs remained responsive to SF2 overexpression. Importantly, endogenous Fas was also affected by changes in SF2 levels, indicating that was not an effect mediated by expression through a different promoter/polyadenylation site. We also observed that SF2 RRM2 promotes skipping of SMN2 exon 7. On the other hand, when we tethered SF2 to Fas exon 6, it promoted exon inclusion and SF2 could antagonize the effect of PTB-MS2 tethered to this sequence, indicating that SF2, under certain conditions, can have antagonistic effects on Fas exon 6 alternative splicing. It is possible that, when interacting directly with the pre-mRNA (i. e. tethering SF2 to exon 6) SF2 can promote exon inclusion, maybe by improving exon definition interactions or by enhancing directly U2AF65 interaction with the upstream pY tract. On the other hand, SF2 RRM2-mediated exon 6 skipping might be caused either indirectly or by inhibiting some other splicing factors, possibly without directly binding the pre-mRNA.

This hypothesis is supported not only by the fact that electrophoretic mobility shift assays using recombinant His-tagged SF2 RRM2 did not point to a specific sequence on Fas (data not shown), but also by the absence of reads in Fas after in vivo crosslinking and immunoprecipitation (CLIP) assays against SF2 (Sanford et al., 2008). Our data also suggest that SF2 ΔRRM2 domain can antagonize SF2 RRM2 effects in certain minigenes (i. e. SMN2), although the mechanism also remains unknown.

Finally, in vitro data did not help to elucidate SF2 mechanism of action. In vitro, splicing in vitro assays did not show any effect of SF2 on removal of either intron 5 or

exon 6 skipping. However, a functional transcription-splicing coupled assay suggested that SF2 could have direct effects on splicing. Moreover, both SF2 and SRp30 RRM2's inhibited spliceosome assembly (i. e. U2 snRNP recruitment) on Fas intron 5 3'ss. It is possible that RRM2 interacts with U2 snRNP components inhibiting its recruitment to the branch point.

It is also conceivable that SF2 RRM2 acts by affecting cytoplasmic functions, because it localizes throughout the cell with only some enrichment in nuclear speckles (Caceres et al., 1997). We showed that SF2 effect on Fas alternative splicing is not mediated by SF2 activation of the mTOR pathway (Fig. III. 9) or enhancement of protein sumoylation (data not shown, in collaboration with F. Pelisch and A. Srebrow). However, SF2 may be involved in other cellular pathways such as microRNA processing (Wu et al., 2010) and nonsense-mediated mRNA decay (NMD) (Sato et al., 2008). Thus, it will be of interest to test if inhibition of these processes (e. g. by knocking down Dicer or Upf1) alters the effect of SF2 overexpression on Fas exon 6 inclusion.

Furthermore, despite the relevance that the second RRM of SF2 seems to have for the many functions of SF2, its specific RNA and protein interactors are still unknown. SF2 RRM2 pull down assays either in vitro or in vivo and CLIP experiments could help to dissect the different roles of SF2 RRM2 and the mechanism by which it regulates Fas and SMN2 alternative splicing.

Our data show that SF2 promotes exon 6 skipping and the anti-apoptotic form of Fas. This might be one of the strategies that SF2 activates during its proto-oncogenic program, allowing cells to escape FasL-mediated apoptosis. Although the mechanism underlying this effect is still unknown, our results suggest that might be an unidentified means by which SF2 regulates splicing.

Many of the described functions of SF2 position this SR protein as a hub in a network of proto-oncogenic and anti-apoptotic processes. One of the nodes would be the generation of anti-apoptotic isoforms of certain apoptosis-related factors (Moore et al., 2010), either directly or indirectly. Consistent with this idea, SF2 promotes the anti-apoptotic form of Fas by enhancing exon 6 skipping. It would be of great interest to determine the role that the anti-apoptotic form of Fas plays in SF2 proto-oncogenic activity (e.g. decreasing cell sensitivity to Fas-mediated apoptosis). Moreover, a correlation between SF2 levels and Fas exon 6 skipping in human cancer cells will highlight the relevance of regulating this alternative splicing event for the tumorigenic activity of SF2.

Finally, SF2-mediated Fas exon 6 skipping can also be important during Fas-mediated apoptosis. We have observed that upon Fas receptor activation in Jurkat

cells, SF2 levels increase by an unidentified mechanism, concomitant with an increase in the anti-apoptotic form of the receptor. Although the physiological significance for this switch is at the moment unclear, increased levels of SF2 may be a mechanism of survival of cells receiving pro-apoptotic signals.

GENERAL DISCUSSION

Splicing of introns to generate translatable mature mRNAs is an essential process for cellular and organism viability. For many years, RNA was considered just a mere messenger of the information encoded by the genome. However, during the last decades, this molecule has risen as one of the main contributors to the complexity of multicellular organisms (reviewed in Mendes Soares and Valcarcel, 2006).

Although the RNA nature of the origin of life is still controversial, catalytically active RNAs or rybozymes may have been present on earth already three billion years ago (Gilbert et al., 1986). The possibility that few molecules of RNA originated all the existing life forms on earth is utterly impressive. Increases of complexity at organism and evolutionary levels have been driven presumably by continuous, subtle changes at a molecular scale along billions of years. In fact, one probably cannot find a better example of increased complexity along evolution than the process of splicing. It very likely emerged from ancestral RNAs similar to self-splicing class II introns, single molecules able to cut themselves via two transesterification reactions (reviewed in Patel and Steitz, 2003; Valadkhan, 2005). However, the splicing machinery in humans is composed by 150-300 snRNP and non-snRNP proteins and five snRNAs (Neubauer et al., 1998; Rappsilber et al., 2002; Zhou et al., 2002). Moreover, more than 90% of intron-containing genes are alternatively spliced and differentially regulated depending on the cell-type, developmental stage, cell environment, stress or presence of drugs, documenting the importance of alternative splicing on the determination of cell fate and response to stimuli (Wang et al., 2008a; Wilhelm et al., 2008).

Splice site recognition: High variability, strict rules

Pre-mRNAs contain highly degenerated sequences that delineate the intron-exon boundaries but nevertheless the splicing machinery recognizes almost unequivocally bona fide splice sites (Horowitz and Krainer, 1994; Zhang, 1998). The appropriate selection of splice sites is aided by the presence of auxiliary cis- and trans-acting factors that promote the recognition of these sequences (reviewed in Smith and Valcarcel, 2000). Moreover, the fidelity of the process is ensured by the recognition of each splice site multiple times during the spliceosome assembly pathway and by the presence of proofreading activities, which determine if the selected sequences and the assembly of the splicing machinery are correct (Mayas et al., 2006; Soares et al., 2006; Villa and Guthrie, 2005; Xu and Query, 2007).

In addition, our data suggest that the recognition of a splice site as such is substrate-specific (Chapter II). For example, the presence of a duplicated acceptor site normally leads to the use of the upstream AG during splicing catalysis. However, the

presence of two tandem AGs leads to complete exon 6 skipping in Fas. This specific effect underscores the functional relevance on splicing of the composition and architecture of the mRNP.

Our results show that initial recruitment of U2AF65 is not affected but yet it cannot promote U2 snRNP interaction with the branch point. The specific structure of U2AF65 binding to Fas intron 5 pY tract may explain the strict sequence requirements governing the recognition of this 3'ss, which is not limited to the duplication of the acceptor site (Fig. II. 1D). This specificity could be mediated by auxiliary sequences present in the surrounding intron and exon or be due to intrinsic characteristics of this particular 3'ss, which positions U2AF65 in a particular conformation.

Although we have determined some sequence modifications that can overcome the effect of the AG duplication, the sequences contributing to this transcript-specific effect of the mutation remain unknown. Interestingly, U2AF65 initial interaction with the pY tract in the AG mutant does not lead to U2 snRNP recruitment, suggesting a defect in U2AF65 conformation in the context of the AG mutation. An altered conformation of U2AF65 could affect the ability of its third RRM and RS domain to interact with SF1 or SF3b155 and branch point (Gozani et al., 1998; Selenko et al., 2003; Valcarcel et al., 1996), thus affecting U2 snRNP recruitment. Therefore, determination of U2AF65's precise contacts with the pre-mRNA (i. e. intron 5 3'ss and branch point) and SF1 interaction with the branch point could help to explain the effect of the AG mutation.

The presence of competing acceptor sites is relevant for the second step of splicing in AG-independent substrates, such as AdML, and can be modulated by second step factors such as SPF45 and hSlu7 (Chua and Reed, 1999b; Lallena et al., 2002). The fact that in Fas the defect happens during the first steps of spliceosome assembly suggests that U2AF35 may have a role in the splicing inhibition caused by the AG mutation. U2AF35 activity is not only limited to AG acceptor site recognition and enhancement of U2AF65 recruitment (Guth et al., 2001). It can also promote splicing without increasing U2AF65 interaction with the pY tract, although a consensus 3' splice site is required (Guth et al., 2001). Again, U2AF35 recruitment was not initially affected in the Fas AG mutant, but an alternative conformation may affect U2AF35-mediated splicing enhancement in an unknown manner not directly linked to U2AF65 recruitment or function.

U2AF interaction with bona fide 3'ss is controlled by the DEK oncogene (Soares et al., 2006). U2 snRNP recruitment to the branch point is also subject to proofreading activities, such as the ATPase Prp5p, which competes with the branch point-U2 snRNA base pairing, thus modulating splicing fidelity (Xu and Query, 2007). We show that also SF3b155, a U2 snRNP factor binding to both sides of the branch point in a sequence-

non-specific manner, may be required for proper base pairing interactions between U2 snRNA and the branch point (Results. Chapter I). Absence of SF3b155 crosslinking, caused by Spliceostatin A treatment, is concomitant with U2 snRNA base pairing with decoy sequences located upstream of the actual branch point. This mechanism implies differential sensitivity of introns to the drug depending on the basepairing potential of these decoy sequences with U2 branch point recognition site, leading to changes in specific alternative splicing events. SF3b155 knock down and SSA treatment in HeLa cells showed a significant overlap between alternative splicing changes monitored by splicing-sensitive microarrays, consistent with the inactivating effect of SSA on SF3b155 in vitro. The NMR structure of U11/U12 di-snRNP allowed to predict how this di-snRNP and possibly U2 snRNP interact with the branch point (Golas et al., 2005). Determination and mutation of the specific residues that interact with the pre-mRNA could help to understand the effects of the absence of SF3b155-pre-mRNA contacts on U2 snRNA base pairing.

Determining possible genetic interactions between SF3b155 and DExD/H-box ATPases, which mediate conformational transitions of the spliceosome, such as Prp5p, could also help to further understand the proofreading activities on U2-snRNA-branch point base pairing interactions. Moreover, it would be of interest to determine the composition of spliceosomes stalled by SSA treatment as has been done for other drugs (Kuhn et al., 2009), therefore elucidating how the inappropriate interaction of U2 snRNP induced by the drug affects the recruitment of other components of the spliceosome.

To gain further insight into direct gene targets of SSA, bioinformatic analysis correlating splice site sensitivity to SSA treatment or SF3b155 knockdown and base pairing potential with U2 snRNA of anchoring sites could be carried out. Moreover, sequence comparison of those splice sites affected or not could suggest other important sequence features mediating the sensitivity to the drug or to SF3b155 knock down.

We have also shown that the presence of U2 bprs is not necessary for U2 snRNP recruitment and U2 snRNA base pairing with the pre-mRNA (Results, Chapter I). Moreover, RNase-mediated U2 snRNA inactivation affects also the interaction of SF3b155 with the pre-mRNA, indicating that SF3b155 is not stabilizing the interaction between inactivated U2 snRNA and the pre-mRNA. This unexpected base pairing was detected in AdML, which contains a strong pY tract and an almost-canonical branch point, but not in Fas –which has weaker 3' splice site signals. This suggests that initial, high affinity recognition of these sequences by U2AF65 and SF1, respectively, can promote the interaction of U2 snRNA with the pre-mRNA even in the absence of

canonical interactions. However, the Δbprs U2 complex can assemble also in weak splice sites, such Fas intron 5, indicating that U2 snRNP recruitment does not necessarily require either U2 snRNA or SF3b155 interactions with the pre-mRNA. Interestingly, Δbprs U2 complex does not seem to be a precursor of the 3' A complex. The latter is only moderately affected by SF1depletion, and its assembly can be restored by addition of a functional, truncated form of SF1 lacking the C-terminal proline-rich region (C4 SF1) (Guth et al., 2000; Rain et al., 1998). However, Δbprs U2 complex does not assemble in SF1-depleted extracts and cannot be rescued by addition of SF1 (data not shown). One interpretation of these result is that immunodepletion of SF1 could deplete other factors important for Δbprs U2 complex assembly. The identification of this factor may help to identify SF1 co-factors involved in U2 snRNP recruitment.

Splice sites are recognized by complex protein-protein, protein-RNA and RNA-RNA networks, many of them mediated by RNA recognition motifs (RRMs). Our work on SF2 RRM2 suggests that certain RRMs can have general inhibitory roles on splicing (Chapter III). Using mutational analysis of our Fas minigene and electrophoretic mobility shift assays, we could not find any cis-acting sequence that could be mediating the effect of SF2 RRM2. Moreover, an SF2-Fas pre-mRNA interaction was not detected by in vivo crosslinking followed by SF2 immunoprecipitation (CLIP) (Sanford et al., 2008). Our data suggest that SF2 RRM2 may prevent the correct assembly of the spliceosome, not only in Fas, but also in other substrates, indicating a more general effect of this domain, perhaps by direct binding to a spliceosome component. Consistent with this idea, SF2 has been identified in proteomic analysis of A and C complexes and the 12S U2 snRNP (reviewed in Jurica and Moore, 2003). Moreover, SF2 has also been described to bind to chromatin and pol II and may affect splicing co-transcriptionally.

If the effect of SF2 in splicing is general, why do introns show differential sensitivity to SF2 RRM2 overexpression? As we have shown for SF3b155, pre-mRNAs can be differentially affected to altered levels of core components of the spliceosome. However, SF2 is involved in many different processes in the cell, including as diverse functions as translation initiation efficiency, protein sumoylation or microRNA processing (Michlewski et al.; Pelisch et al., 2010; Wu et al., 2010). Although we have already discarded some of these pathways as mediators of the effect of SF2 on splicing, such as sumoylation (Pelisch et al., 2010) and mTOR pathway activation (Karni et al., 2008), others remain to be studied. Analysis of SF2 RRM2 wild type and

mutant interacting partners might help to understand the mechanism by which SF2 inhibits splicing.

Splicing-modulation therapies to treat complex diseases

The vast amount of cis- and trans-acting factors involved in splicing catalysis and regulation affords the possibility to precisely fine tune the sequences present in the final transcript. This high level of regulation contributes to the specific transcript signature characteristic for each cell type in a given environment. Inevitably, however, the enormous complexity achieved is a double-edged sword. Mutations either in regulatory sequences, splicing factors or core components of the spliceosome are a common cause of disease (Wang and Cooper, 2007; Lopez-Bigas et al., 2005). Moreover, aberrant, disease-causing splicing patterns can be also due to missregulation of several signaling pathways (reviewed in Shin et al., 2004).

Splicing-modulation therapies are mainly focused on altering specific alternative splicing events to ameliorate the symptoms of certain monogenic diseases. For example, increasing the levels of exon 7 inclusion of the SMN2 gene promotes the expression of a functional SMN protein able to compensate for mutations in SMN1 (Hua et al., 2008; Hua et al., 2010), which cause Spinal Muscular Atrophy (SMA) (Lefebvre et al.). Modulation of splicing to restore the reading frame in mutated genes is another approach that could be used for example to treat Duchenne Muscular Dystrophy (DMD), caused by mutations in the dystrophin gene (Disset et al., 2006). Modification of SR protein expression, especially of the oncogenic SF2, or its phosphorylation levels by inhibiting SR protein kinases, have been proposed to alter splicing events important for cancer or retrovirus replication (Bakkour et al., 2007; Fukuhara et al., 2006; Keriel et al., 2009; Muraki et al., 2004; Soret et al., 2005).

Cancer cells usually show aberrant splicing patterns that promote cell division, resistance to apoptosis, angiogenesis or tumor invasion (reviewed in Venables, 2006). Nevertheless, recent reports and results in this thesis suggest that direct targeting of the splicing machinery, and not just a single splicing event, can be used as a novel anticancer therapy. Spliceostatin A and pladienolide were the first molecules described to interact directly with SF3b, a U2 snRNP component, inhibit splicing and display an anti-tumor effect (Kaida et al., 2007; Kotake et al., 2007). Interestingly, both compounds were isolated from the fermentation broth of different bacteria and their functions in these organisms are still unknown (Kaida et al., 2007; Kotake et al., 2007).

Why are tumor cells more sensitive to drugs targeting the spliceosome than normal cells? There are two main possibilities to explain this observation. First,

because of their high metabolic rate, tumor cells may be more sensitive to splicing inhibition and decreased transcript levels (Furumai et al., 2010). It is not clear, however, how cells with a high metabolic rate would be more affected by splicing inhibition. Alternatively, these drugs could be particularly affecting the splicing pattern of specific genes important for cell cycle progression. Our data suggest that the latter option is more likely to explain the mechanism of action of Spliceostatin A, becuase it decreases the levels of certain cell cycle regulators (i. e. CCNA2, AURKA) by promoting NMD-associated alternative splicing in these genes. Consistentl with this, AURKA has been recently described to promote anti-apoptotic splicing of several genes, partially by increasing the levels of SF2 (Moore et al., 2010).

Remarkably, the splicing factor SF2 is a proto-oncogene. Increased levels of SF2 have been observed in several types of cancers and can promote tumor growth in nude mice (Karni et al., 2007), partially by activating certain anti-apoptotic splicing patterns (Moore et al., 2010). We have shown that the non-canonical RRM2 of SF2 is necessary and sufficient to promote the anti-apoptotic form of Fas. SF2 RRM2 is sufficient to modify the splicing pattern and splicing efficiency of certain number of events. It will be interesting to determine if SF2 RRM2 by itself can regulate the anti-apoptotic program activated by SF2 and which is the exact contribution of the RRM2 to the proto-oncogenic activity of SF2.

The next question is whether drugs targeting different components of the spliceosome would all have antitumor activity. Both pladienolide and Spliceostatin A bind the same U2 snRNP subcomplex and inhibit tumor progression (Kaida et al., 2007; Kotake et al., 2007). It would be of interest to determine if pladienolide also affects SF3b155 interaction with the pre-mRNA, thus altering U2 snRNA interaction with the branch point. However, mutations in core components of the spliceosome, such as Prp8 or Prp31, cause a dominant version of retinitis pigmentosa and only photoreceptor neurons seem to be affected by the haploinsufficiency of these factors (Mordes et al., 2007). Also mutations on the SMN1 gene, important for snRNP maturation and translocation to the nucleus, cause spinal muscular atrophy (Lefebvre et al., 1995). These observations indicate that an anticancer activity can only be achieved by affecting the function of certain snRNP components, but not others.

High-throughput technologies to analyze tissue specific transcriptomes will help to determine aberrantly expressed transcripts in tumors and other diseased tissues (Wang et al., 2008). Comparative analysis of tumor samples treated or not with Spliceostatin A could point to genes responsible of tumorigenesis, tumor progression and metastasis for specific cancers. For example, meayamycin, a FR901464 derivative that also inhibits splicing, greatly affected growth of some human breast, colon and

prostate cancer cell lines, while its effect was rather minimal in lung and other prostate cancer cell lines (Albert et al., 2009). The differential effects of this drug in several types of tumors could indicate the existence tumor-specific splicing signatures, and consequently tumor characteristics, pointing to different candidate genes for cancer-specific therapy.

To conclude, the work of this thesis underscores the relevance of an adequate balance of alternative splicing isoformsfor appropriate regulation of pivotal processes such as cell division and apoptosis.

MATERIALS AND METHODS

This section aims to implement the Materials and Methods sections of chapters I and II. Only the homemade buffers are included in this section, the commercial ones being omitted. The concentrations shown are final concentrations (1X) unless otherwise stated.

Materials

Nuclear extracts:

- HeLa cell nuclear extract (Cilbiotech, Cat. No. CC-01-20-50).

Buffers for DNA precipitation:

- TNE: 20 mM Tris pH7.5, 1mM EDTA, 100 mM NaCl.

Reagents for DNA staining:

- Ethidium Bromide: 1 μg/ml

- SYBR Safe DNA gel stain (Invitrogen, Cat. No. S33102), dilution 1/10000.

Buffers for enzymes:

- RQ1 DNase buffer: 40 mM Tris HCl pH 8.0, 10 mM NaCl, 6 mM $MgCl_2$, 10 mM $CaCl_2$.

- Protein Kinase buffer: 100 mM Tris 7.5, 12.5 mM EDTA, 150 mM NaCl, 1% SDS).

Buffers for immunoprecipitation assays:

- Low-salt buffer: 150 mM NaCl, 50 mM Tris pH 8.0, 1% NP40.

- High-salt buffer: 500 mM NaCl, 50 mM Tris pH 8.0, 1% NP40.

Buffers for purification of recombinant proteins:

- Buffer X: 20 mM Tris pH 8.4, 1M NaCl, 0.2 mM EDTA, 1mM DTT.

- Protein elution buffer: 50 mM glutathione, 100 mM Hepes-KOH pH 8.0, 1 mM DTT.

- Lysis buffer: 50 mM NaH_2PO_4, 800 mM NaCl, 1% Triton X-100, 10 mM b-mercaptoethanol, 10% glycerol, protease inhibitors.

- Buffer D: 20 mM Hepes-KOH pH 7.9, 0.2 mM EDTA, 20% glycerol, 1 mM DTT, 0.01% NP40.

Gel loading dyes:

- Agarose loading dye 6X: 20 mM Tris pH 7.5, 0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol.

- Sequencing gel-loading dye: 98% deionized formamide, 10 mM EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue.

- SDS loading buffer 4X: 200 mM Tris pH 6.8, 400 mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol.

Protein Gel staining solutions:

- Coomassie staining solution: 0.25% Coomassie Brilliant Blue, 10% acetic acid and 50% methanol.

- Coomassie destaining solution: 15% methanol, 10% acetic acid.

Electrophoresis buffers:

- SDS gel electrophoresis buffer: 25 mM Tris base, 0.25 M glycine, 0.1% SDS.

- TBE: 89 mM Tris base, 89 mM Boric acid, 2 mM EDTA.

- Tris-Glycine: 50 mM Tris base, 50 mM glycine.

Other buffers:

- RNA elution buffer: 0.5 M NH_4Ac, 10 mM $MgAc_2$, 1 mM EDTA, 0.1% SDS.

- Semi-dry system Western Blot transfer buffer: 20% Methanol, 250 mM Tris base, 2.5 M glycine, 1% SDS.

Methods

Protein Purification and analysis

Protein induction in bacteria

100 ml stationary phase cultures of BL21 *E. coli* previously transformed with the desired contruct were diluted to 1 litre of LB medium with ampicilin (100 µg/ml) and grown at 37º C with shaking until the cells reached an OD_{600}=0.6. Induction was achieved by adding 0.1 mM IPTG to the culture and incubating for 3 hours at 37ºC.

Production of His fusion proteins in bacteria

The cells were sedimented by centrifugation at 4000 rpm for 10 min at 4º C and the resuspended in 20 ml of Lysis buffer. Cells were disrupted by sonicating 4'-5' in intervals of 30'' ON- 30'' OFF and 40% amplitude. The lysates were centrifuged at 16000 g for 20' at 4ºC. The supernatant was loaded in Ni2+-NTA columns containing 1 mL of resin slurry pre-washed with 5 mL deionized water and 10 mL Lysis buffer. The columns were washed three times with 5 mL Lysis buffer-10 mM imidazole and once with 5 mL Lysis buffer-20 mM imidazole. His-tagged fusion proteins were eluted in Lysis buffer-0.4M imidazole and 200 µl fractions were collected. The protein profiles of the eluted fractions were estimated by Bradford assay. The fractions containing the peak of protein concentration were dialyzed against 1 litre of buffer D 0.1 M KCl overnight at 4º C. The proteins were snap frozen in liquid N_2 and stored at –80º C until use.

The recombinant protein concentration was estimated by comparison with known concentrations of bovine serum albumin (BSA). Different amounts of the protein of interest or BSA were mixed with SDS loading buffer, boiled for 5 min and loaded on

denaturing SDS 12% polyacrylamide gels (37.5:1 acrylamide to bis-acrylamide). Electrophoresis was performed in SDS running buffer at 200 V for 45 min using Biorad minigel system. The gel was then stained with coomassie staining solution and destained in destaining solution, transferred to 3 MM paper and dried.

Production of GST fusion proteins in bacteria

The cells were sedimented by centrifugation at 4000 rpm for 10 min at 4º C and the resuspended in 20 ml of buffer X. Cells were sedimented again by centrifugation at 2000 rpm for 7 min and the resuspended in the same volume of buffer X supplemented with protease inhibitors. Cells were broken as described above and the lysates centrifuged at 10000 rpm for 20 min at 4º C. The supernatant was incubated with 50 mg of Glutathione agarose beads slurry -prewashed with buffer X supplemented with protease inhibitors- for 30 min at 4º C with shaking. The beads were sedimented by centrifugation at 1000 rpm for 5 min and resuspended in 5 ml buffer X supplemented with protease inhibitors and applied to a 10 ml Biorad plastic column holder and packed by gravity. Beads were then washes with 10 mL of buffer X with protease inhibitors and the 10 mL of buffer X without protease inhibitors. GST fusion proteins were eluted in elution buffer and fractions of 200 µl were collected. Eluted proteins were analyzed and stored as described above.

Western blotting

HeLa cell protein extracts were separated by polyacrylamide gel electrophoresis and transferred to nitrocellulose membrane previously washed with western transfer buffer for 1 h at 250 mAmp using Biorad© semi-dry system. Transferred membranes were blocked with PBS + tween-20 0.1% (v/v) with 10% (w/v) milk for 1 h and incubated with a specific primary antibody in PBS + tween-20 0.1% (v/v) with 5% milk (w/v) for 1 hour. The membranes were washed three times with PBS + tween-20 0.1% (v/v) during 5 min and incubated again with the corresponding secondary antibody coupled to horseradish peroxidase. The membrame was developed using Western Lightning™ Chemiluminescence Reagent Plus (PerkinElmer Life Sciences, Inc.) and exposed to photographic film.

Electrophoretic mobility shift assays

Recombinant proteinswere incubated in buffer D pH 8, 0.1 M KCl with 10 fmol of ^{32}P radiolabeled RNA in the presence of 300 ng/µl of tRNA in a total volume of 9 µl. After 20 min at 30ºC, RNA-protein complexes were separated from free RNA by electrophoresis on non-denaturing 5% polyacrylamide gels in 0.5X TBE. The gels where initially run at 50 V for 15 min and then at 250 V for 2 hours at 4ºC. The gel was then transferred to 3 MM Whatman paper, dried and analyzed by Phosphorimager.

RNA oligonucleotide affinity pull-down experiments

60 µg of commercial WT (5'-UAAAAUGUCCAAUGUUCCAACCUACAG/G-3') or AG (5'-UAAAAUGUCCAAUGUUCCAACCUAGAG/G-3') RNA oligonucleotides (Integrated DNA Technologies) were oxidized at its 5' end by incubating with 5 mM Na m-periodate (Sigma, Cat. No. S-1147) and 0.1 M Na acetate pH 5.2, for 1 h at RT in the dark. Oxidized oligonucleotides were EtOH precipitated and resuspended in 50 µl 0.1 M Na acetate, pH 5.2. 100 µl of 1:1 slurry of adipic acid dihydrazide-agarose beads (Sigma, Cat. No. A-0802) pre-washed with 0.1 M Na acetate were added to the resuspended RNA olinucleotides and left overnight at 4ºC in a rotating wheel. The beads were then washed three times with 2 M NaCl and three times more with buffer D pH8, 0.1 M KCl. Subsequently, the washed bound beads-RNA oligonucleotides were introduced in a 250-µl splicing reaction and incubated at 30ºC for 40'. Beads were then washed three times with 1 ml buffer D pH 8 with different KCl concentrations (100, 300 and 600 mM). Each washing step was followed by an 80 µl elution with buffer D pH 8 at the same concentration of the washes after. Eluted proteins were run in a 10% SDS polyacrylamide gel and Coomasie stained for analysis. Differential bands between both RNA oligonucleotides were cut and sequenced by mass spectrometry. Results were confirmed by Western Blotting.

Coupled transcription-splicing assay

This assay was performed as described in (Das et al., 2006). Briefly, a PCR product containing the CMV promoter followed by the minigene of interest was generated and purified by column (Qiagen). 200 ng of PCR product was incubated with 15 µL HeLa cell nuclear extracts, 0.5 mM ATP, 20 mM CP, 3.2 mM $MgCl_2$ and 0.5 µL 32P-UTP (3000 Ci/mmol, 10 mCi/mL) in a final volume of 25 µL. Reactions were incubated 10' at 30º. Then, transcription was stopped by adding 1 µL of alpha-amanitin (10ng/µl) and a 5 µL-aliquot was extracted and kept on ice. The rest of the sample was incubated for another 60' at 30ºC. Reactions were stopped by PK digestion and the RNA species purified by ethanol precipitation. RNA species were subsequently analyzed by electrophoresis in 8% denaturing polyacrylamide gels exposed to PhosphorImager screens.

BIBLIOGRAPHY

Adam, S.A., Nakagawa, T., Swanson, M.S., Woodruff, T.K., and Dreyfuss, G. (1986). mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. Mol Cell Biol 6, 2932-2943.

Agris, P.F., KiKuchi, Y., Gross, H.J., Takano, M., and Sharp, G.C. (1984). Characterization of the autoimmune antigenic determinant for ribonucleoprotein (RNP) antibody. Immunol Commun 13, 137-149.

Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell 13, 67-76.

Akerman, M., and Mandel-Gutfreund, Y. (2006). Alternative splicing regulation at tandem 3' splice sites. Nucleic Acids Res. 34, 23-31.

Akgul, C., Moulding, D.A., and Edwards, S.W. (2004). Alternative splicing of Bcl-2-related genes: functional consequences and potential therapeutic applications. Cell Mol Life Sci 61, 2189-2199.

Akiva, P., Toporik, A., Edelheit, S., Peretz, Y., Diber, A., Shemesh, R., Novik, A., and Sorek, R. (2006). Transcription-mediated gene fusion in the human genome. Genome Res 16, 30-36.

Albert, B.J., McPherson, P.A., O'Brien, K., Czaicki, N.L., Destefino, V., Osman, S., Li, M., Day, B.W., Grabowski, P.J., Moore, M.J., et al. (2009). Meayamycin inhibits pre-messenger RNA splicing and exhibits picomolar activity against multidrug-resistant cells. Mol Cancer Ther 8, 2308-2318.

Allers, J., and Shamoo, Y. (2001). Structure-based analysis of protein-RNA interactions using the program ENTANGLE. J Mol Biol 311, 75-86.

Allo, M., Buggiano, V., Fededa, J.P., Petrillo, E., Schor, I., de la Mata, M., Agirre, E., Plass, M., Eyras, E., Elela, S.A., et al. (2009). Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nat Struct Mol Biol 16, 717-724.

Bakkour, N., Lin, Y.L., Maire, S., Ayadi, L., Mahuteau-Betzer, F., Nguyen, C.H., Mettling, C., Portales, P., Grierson, D., Chabot, B., et al. (2007). Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. PLoS Pathog 3, 1530-1539.

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. Nature 465, 53-59.

Bass, B.L., and Weintraub, H. (1988). An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55, 1089-1098.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M.andSonnhammer, E. L. (2002).The Pfam protein families database. Nucleic Acids Res30, 276-280.

Batsche, E., Yaniv, M., and Muchardt, C. (2006). The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat Struct Mol Biol 13, 22-29.

Beckmann, J.S., and Trifonov, E.N. (1991). Splice junctions follow a 205-base ladder. Proc Natl Acad Sci U S A 88, 2380-2383.

Behrens, S.E., Tyc, K., Kastner, B., Reichelt, J., and Luhrmann, R. (1993). Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. Mol Cell Biol 13, 307-319.

Berget, S.M. (1995). Exon recognition in vertebrate splicing. J Biol Chem 270, 2411-2414.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A 74, 3171-3175.

Berglund, J.A., Abovich, N., and Rosbash, M. (1998). A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. Genes Dev 12, 858-867.

Berglund, J.A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997). The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell 89, 781-787.

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., and R., L. (2008). Isolation of an active step I spliceosome and composition of its RNP core. Nature 452, 846-850.

Bird, G., Zorio, D.A., and Bentley, D.L. (2004). RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. Mol Cell Biol 24, 8963-8969.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72, 291-336.

Blencowe, B.J., and Khanna, M. (2007). Molecular biology: RNA in control. Nature 447, 391-393.

Blumenthal, T. (2005). Trans-splicing and operons. WormBook, 1-9.

Bonnal, S., Martinez, C., Forch, P., Bachi, A., Wilm, M., and Valcarcel, J. (2008). RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. Mol Cell 32, 81-95.

Boucher, L., Ouzounis, C.A., Enright, A.J., and Blencowe, B.J. (2001). A genome-wide survey of RS domain proteins. RNA 7, 1693-1701.

Bourgeois, C.F., Lejeune, F., and Stevenin, J. (2004). Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. Prog Nucleic Acid Res Mol Biol 78, 37-88.

Boutz, P.L., Stoilov, P., Li, Q., Lin, C.H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., Jr., and Black, D.L. (2007). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. Genes Dev 21, 1636-1652.

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., and et al. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. Cell 68, 799-808.

Brosi, R., Groning, K., Behrens, S.E., Luhrmann, R., and Kramer, A. (1993a). Interaction of mammalian splicing factor SF3a with U2 snRNP and relation of its 60-kD subunit to yeast PRP9. Science 262, 102-105.

Brosi, R., Hauri, H.P., and Kramer, A. (1993b). Separation of splicing factor SF3 into two components and purification of SF3a activity. J Biol Chem 268, 17640-17646.

Bruzik, J.P., Van Doren, K., Hirsh, D., and Steitz, J.A. (1988). Trans splicing involves a novel form of small nuclear ribonucleoprotein particles. Nature 335, 559-562.

Buratti, E., and Baralle, F.E. (2004). Influence of RNA secondary structure on the pre-mRNA splicing process. Mol Cell Biol 24, 10505-10514.

Buratti, E., Baralle, M., and Baralle, F.E. (2006). Defective splicing, disease and therapy: searching for master checkpoints in exon definition. Nucleic Acids Res 34, 3494-3510.

Buratti, E., Brindisi, A., Pagani, F., and Baralle, F.E. (2004). Nuclear factor TDP-43 binds to the polymorphic TG repeats in CFTR intron 8 and causes skipping of exon 9: a functional link with disease penetrance. Am J Hum Genet 74, 1322-1325.

Burge, C.B., Padgett, R.A., and Sharp, P.A. (1998). Evolutionary fates and origins of U12-type introns. Mol Cell 2, 773-785.

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

Caceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L., and Krainer, A.R. (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. J Cell Biol 138, 225-238.

Caceres, J.F., Screaton, G.R., and Krainer, A.R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. Genes Dev 12, 55-66.

Cammas, A., Pileur, F., Bonnal, S., Lewis, S. M., Lévêque, N., Holcik, M., and Vagner S. (2007). Cytoplasmic relocalization of heterogeneous nuclear ribonucleoprotein A1 controls translation initiation of specific mRNAs. Mol Biol Cell 18, 5048-59.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., and Krainer, A.R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31, 3568-3571.

Cascino, I., Fiucci, G., Papoff, G., and Ruberti, G. (1995). Three functional soluble forms of the human apoptosis-inducing Fas molecule are produced by alternative splicing. J Immunol 154, 2706-2713.

Champlin, D.T., Frasch, M., Saumweber, H., and Lis, J.T. (1991). Characterization of a Drosophila protein associated with boundaries of transcriptionally active chromatin. Genes Dev 5, 1611-1621.

Chang, W.C., Chen, H.H., and Tarn, W.Y. (2008). Alternative splicing of U12-type introns. Front Biosci 13, 1681-1690.

Chen, L., Park, S.M., Tumanov, A.V., Hau, A., Sawada, K., Feig, C., Turner, J.R., Fu, Y.X., Romero, I.L., Lengyel, E., and Peter, M.E. (2010). CD95 promotes tumour growth. Nature 465, 492-496.

Cheng, J., Zhou, T., Liu, C., Shapiro, J.P., Brauer, M.J., Kiefer, M.C., Barr, P.J., and Mountz, J.D. (1994). Protection from Fas-mediated apoptosis by a soluble form of the Fas molecule. Science 263, 1759-1762.

Cheng, J., Zhou, T., Liu, C., Shapiro, J.P., Brauer, M.J., Kiefer, M.C., Barr, P.J., Mountz, J.D. (1994). Protection from Fas-mediated apoptosis by a soluble form of the Fas molecule. . Science. 263, 1759-1762.

Chern, T.M., van Nimwegen, E., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., and Zavolan, M. (2006). A simple physical model predicts small exon length variations. PLoS Genet 2, e45.

Chiodi, I., Corioni, M., Giordano, M., Valgardsdottir, R., Ghigna, C., Cobianchi, F., Xu, R.M., Riva, S., and Biamonti, G. (2004). RNA recognition motif 2 directs the recruitment of SF2/ASF to nuclear stress bodies. Nucleic Acids Res 32, 4127-4136.

Cho, E.J., Takagi, T., Moore, C.R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. Genes Dev 11, 3319-3326.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell 12, 1-8.

Chu, C. S., Trapnell, B. C., Murtagh, J. J., Moss, J., Dalemans, W., Jallat, S., Mercenier, A., Pavirani, A., Lecocq, J. P., and Cutting, G. R. (1991) Variable deletion of exon 9 coding sequences in cystic fibrosis transmembrane conductance regulator gene mRNA transcripts in normal bronchial epithelium. EMBO J10, 1355-1363.

Chua, K., and Reed, R. (1999a). Human step II splicing factor hSlu7 functions in restructuring the spliceosome between the catalytic steps of splicing. Genes Dev 13, 841-850.

Chua, K., and Reed, R. (1999b). The RNA splicing factor hSlu7 is required for correct 3' splice-site choice. Nature 402, 207-210.

Chua, K., and Reed, R. (2001). An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing. Mol Cell Biol 21, 1509-1514.

Clery, A., Blatter, M., and Allain, F.H. (2008). RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol 18, 290-298.

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C., and Duncan, P.I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J 15, 265-275.

Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and disease. Cell 136, 777-793.

Corrionero, A., and Valcarcel, J. (2009). RNA processing: Redrawing the map of charted territory. Mol Cell 36, 918-919.

Corsini, L., Bonnal, S., Basquin, J., Hothorn, M., Scheffzek, K., Valcarcel, J., and Sattler, M. (2007). U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. Nat Struct Mol Biol 14, 620-629.

Corsini, L., Hothorn, M., Stier, G., Rybin, V., Scheffzek, K., Gibson, T.J., and Sattler, M. (2009). Dimerization and protein binding specificity of the U2AF homology motif of the splicing factor Puf60. J Biol Chem 284, 630-639.

Cramer, P., Caceres, J.F., Cazalla, D., Kadener, S., Muro, A.F., Baralle, F.E., and Kornblihtt, A.R. (1999). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol Cell 4, 251-258.

Crick, F. (1970). Central dogma of molecular biology. Nature 227, 561-563.

Crick, F.H. (1958). On protein synthesis. Symp Soc Exp Biol 12, 138-163.

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

Crispino, J.D., Mermoud, J.E., Lamond, A.I., and Sharp, P.A. (1996). Cis-acting elements distinct from the 5' splice site promote U1-independent pre-mRNA splicing. RNA 2, 664-673.

Darnell, J.E., Jr. (1978). Implications of RNA-RNA splicing in evolution of eukaryotic cells. Science 202, 1257-1260.

Das, B.K., Xia, L., Palandjian, L., Gozani, O., Chyung, Y., and Reed, R. (1999). Characterization of a protein complex containing spliceosomal proteins SAPs 49, 130, 145, and 155. Mol Cell Biol 19, 6796-6802.

Das, R., Dufu, K., Romney, B., Feldt, M., Elenko, M., and Reed, R. (2006). Functional coupling of RNAP II transcription to spliceosome assembly. Genes Dev 20, 1100-1109.

Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P., and Reed, R. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. Mol Cell 26, 867-881.

Dauksaite, V., and Akusjarvi, G. (2002). Human splicing factor ASF/SF2 encodes for a repressor domain required for its inhibitory activity on pre-mRNA splicing. J Biol Chem 277, 12579-12586.

Dauksaite, V., and Akusjarvi, G. (2004). The second RNA-binding domain of the human splicing factor ASF/SF2 is the critical domain controlling adenovirus E1A alternative 5'-splice site selection. Biochem J 381, 343-350.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. Mol Cell 12, 525-532.

de la Mata, M., and Kornblihtt, A.R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nat Struct Mol Biol 13, 973-980.

DeCerbo, J., and Carmichael, G.G. (2005). Retention and repression: fates of hyperedited RNAs in the nucleus. Curr Opin Cell Biol 17, 302-308.

Deutsch, M., and Long, M. (1999). Intron-exon structures of eukaryotic model organisms. Nucleic Acids Res 27, 3219-3228.

Dietrich, R.C., Incorvaia, R., and Padgett, R.A. (1997). Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. Mol Cell 1, 151-160.

Disset, A., Bourgeois, C.F., Benmalek, N., Claustres, M., Stevenin, J., and Tuffery-Giraud, S. (2006). An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements. Hum Mol Genet 15, 999-1013.

Dix, I., Russell, C.S., O'Keefe, R.T., Newman, A.J., and Beggs, J.D. (1998). Protein-RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae*. RNA 4: 1675-1686.

Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol 3, 195-205.

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. Annu Rev Biochem 62, 289-321.

Du, H., and Rosbash, M. (2002). The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. Nature 419, 86-90.

Dye, M.J., Gromak, N., and Proudfoot, N.J. (2006). Exon tethering in transcription by RNA polymerase II. Mol Cell 21, 849-859.

Earnshaw, W.C., Martins, L.M., and Kaufmann, S.H. (1999). Mammalian caspases: structure, activation, substrates, and functions during apoptosis. Annu Rev Biochem 68, 383-424.

Ebbinghaus, C., Scheuermann, J., Neri, D., and Elia, G. (2004). Diagnostic and therapeutic applications of recombinant antibodies: targeting the extra-domain B of fibronectin, a marker of tumor angiogenesis. Curr Pharm Des 10, 1537-1549.

Ellis, J.D., Lleres, D., Denegri, M., Lamond, A.I., and Caceres, J.F. (2008). Spatial mapping of splicing factor complexes involved in exon and intron definition. J Cell Biol 181, 921-934.

Epstein, J.A., Glaser, T., Cai, J., Jepeal, L., Walton, D.S., and Maas, R.L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. Genes Dev 8, 2022-2034.

Fedorova, L., and Fedorov, A. (2003). Introns in gene evolution. Genetica 118, 123-131.

Feng, Y., Valley, M.T., Lazar, J., Yang, A.L., Bronson, R.T., Firestein, S., Coetzee, W.A., and Manley, J.L. (2009). SRp38 regulates alternative splicing and is required for Ca(2+) handling in the embryonic heart. Dev Cell 16, 528-538.

Fingar, D.C., and Blenis, J. (2004). Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression. Oncogene 23, 3151-3171.

Fisher, G.H., Rosenberg, F.J., Straus, S.E., Dale, J.K., Middleton, L.A., Lin, A.Y., Strober, W., Lenardo, M.J., and Puck, J.M. (1995). Dominant interfering Fas gene mutations impair apoptosis in a human autoimmune lymphoproliferative syndrome. Cell 81, 935-946.

Fong, N., Ohman, M., and Bentley, D.L. (2009). Fast ribozyme cleavage releases transcripts from RNA polymerase II and aborts co-transcriptional pre-mRNA processing. Nat Struct Mol Biol 16, 916-922.

Forch, P., Puig, O., Kedersha, N., Martinez, C., Granneman, S., Seraphin, B., Anderson, P., and Valcarcel, J. (2000). The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing. Mol Cell 6, 1089-1098.

Fortner, D.M., Troy, R.G., and Brow, D.A. (1994). A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. Genes Dev 8, 221-233.

Fox-Walsh, K.L., Dou, Y., Lam, B.J., Hung, S.P., Baldi, P.F., and Hertel, K.J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proc Natl Acad Sci U S A 102, 16176-16181.

Fu, Y.H., Pizzuti, A., Fenwick, R.G., Jr., King, J., Rajnarayan, S., Dunne, P.W., Dubel, J., Nasser, G.A., Ashizawa, T., de Jong, P., and et al. (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science 255, 1256-1258.

Fukuhara, T., Hosoya, T., Shimizu, S., Sumi, K., Oshiro, T., Yoshinaka, Y., Suzuki, M., Yamamoto, N., Herzenberg, L.A., and Hagiwara, M. (2006). Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. Proc Natl Acad Sci U S A 103, 11329-11333.

Furumai, R., Uchida, K., Komi, Y., Yoneyama, M., Ishigami, K., Watanabe, H., Kojima, S., and Yoshida, M. (2010). Spliceostatin A blocks angiogenesis by inhibiting global gene expression including VEGF. Cancer Sci.

Fushimi, K., Ray, P., Kar, A., Wang, L., Sutherland, L.C., and Wu, J.Y. (2008). Up-regulation of the proapoptotic caspase 2 splicing isoform by a candidate tumor suppressor, RBM5. Proc Natl Acad Sci U S A 105, 15708-15713.

Gama-Carvalho, M., Krauss, R.D., Chiang, L., Valcarcel, J., Green, M.R., and Carmo-Fonseca, M. (1997). Targeting of U2AF65 to sites of active splicing in the nucleus. J Cell Biol 137, 975-987.

Gilbert, W., Marchionni, M., and McKnight, G. (1986). On the antiquity of introns. Cell 46, 151-153.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett 582, 1977-1986.

Golas, M.M., Sander, B., Will, C.L., Luhrmann, R., and Stark, H. (2005). Major conformational change in the complex SF3b upon integration into the spliceosomal U11/U12 di-snRNP as revealed by electron cryomicroscopy. Mol Cell 17, 869-883.

Goncalves, V., Matos, P., and Jordan, P. (2009). Antagonistic SR proteins regulate alternative splicing of tumor-related Rac1b downstream of the PI3-kinase and Wnt pathways. Hum Mol Genet 18, 3696-3707.

Gooding, C., Clark, F., Wollerton, M.C., Grellscheid, S.N., Groom, H., and Smith, C.W. (2006). A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones. Genome Biol 7, R1.

Görlich, D., and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm.Annu Rev Cell Dev Biol 15, 607-60.

Gozani, O., Feld, R., and Reed, R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev 10, 233-243.

Gozani, O., Potashkin, J., and Reed, R. (1998). A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol Cell Biol 18, 4752-4760.

Grainger, R.J., and Beggs, J.D. (2005). Prp8 protein: at the heart of the spliceosome. RNA 11, 533-557.

Graveley, B.R. (2005). Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. Cell 123, 65-73.

Green, M.R. (1986). Pre-mRNA splicing. Annu Rev Genet 20, 671-708.

Griffiths, A., Miller, J. Suyuki, D., Lewontin, R., and Gelbart, W. (1999). Introduction ot Genetic Analysis, 7th ed. (New York, NY: W. H. Freeman & Co.).

Gromak, N., Rideau, A., Southby, J., Scadden, A. D. J., Gooding, C., Hüttelmaier, S., Singer, R. H., and Smith, C. W. J. (2003). The PTB interacting protein raver 1 regulates α-tropomyosin alternative splicing. EMBO J 22, 6356-6364.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35, 849-857.

Gui, J.F., Tronchere, H., Chandler, S.D., and Fu, X.D. (1994). Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. Proc Natl Acad Sci U S A 91, 10824-10828.

Guil, S., Long, J. C., and Cáceres, J. F. (2006). hnRNP A1 relocalization to the stress granules reflects a role in the stress response.Mol Cell Biol 26, 5744-58.

Guth, S., Martinez, C., Gaur, R.K., and Valcarcel, J. (1999). Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). Mol Cell Biol 19, 8263-8271.

Guth, S., Tange, T.O., Kellenberger, E., and Valcarcel, J. (2001). Dual function for U2AF(35) in AG-dependent pre-mRNA splicing. Mol Cell Biol 21, 7673-7681.

Hall, S.L., and Padgett, R.A. (1994). Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. J Mol Biol 239, 357-365.

Hastings, M.L., Allemand, E., Duelli, D.M., Myers, M.P., and Krainer, A.R. (2007). Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF(65). PLoS One 2, e538.

Hawkins, J.D. (1988). A survey on intron and exon lengths. Nucleic Acids Res 16, 9893-9908.

He, Y., and Smith, R. (2009). Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. Cell Mol Life Sci 66, 1239-1256.

Hector, R.E., Nykamp, K.R., Dheur, S., Anderson, J.T., Non, P.J., Urbinati, C.R., Wilson, S.M., Minvielle-Sebastia, L., and Swanson, M.S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. EMBO J 21, 1800-1810.

Heinrichs, V., Bach, M., Winkelmann, G., and Luhrmann, R. (1990). U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. Science 247, 69-72.

Helfman, D.M., and Ricci, W.M. (1989). Branch point selection in alternative splicing of tropomyosin pre-mRNAs. Nucleic Acids Res 17, 5633-5650.

Hertel, K.J. (2008). Combinatorial control of exon recognition. J Biol Chem 283, 1211-1215.

Hetzer, M., and Mattaj, I.W. (2000). An ATP-dependent, Ran-independent mechanism for nuclear import of the U1A and U2B" spliceosome proteins. J Cell Biol 148, 293-303.

Hiller, M., Huse, K., Szafranski, K., Jahn, N., Hampe, J., Schreiber, S., Backofen, R., and Platzer, M. (2004). Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity. Nat Genet 36, 1255-1257.

Hiller, M., Szafranski, K., Huse, K., Backofen, R., and Platzer, M. (2008). Selection against tandem splice sites affecting structured protein regions. BMC Evol Biol 8, 89.

Hiller, M., Zhang, Z., Backofen, R., and Stamm, S. (2007). Pre-mRNA secondary structures influence exon recognition. PLoS Genet 3, e204.

Hoet, R.M., Koornneef, I., de Rooij, D.J., van de Putte, L.B., and van Venrooij, W.J. (1992). Changes in anti-U1 RNA antibody levels correlate with disease activity in patients with systemic lupus erythematosus overlap syndrome. Arthritis Rheum 35, 1202-1210.

Hogg, R., McGrail, J.C., and O'Keefe, R.T. (2010). The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. Biochem Soc Trans 38, 1110-1115.

Hong, W., Bennett, M., Xiao, Y., Feld Kramer, R., Wang, C., and Reed, R. (1997). Association of U2 snRNP with the spliceosomal complex E. Nucleic Acids Res 25, 354-361.

Horowitz, D.S., and Krainer, A.R. (1994). Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing. Trends Genet 10, 100-106.

Hua, Y., Sahashi, K., Hung, G., Rigo, F., Passini, M.A., Bennett, C.F., and Krainer, A.R. (2010). Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. Genes Dev 24, 1634-1644.

Hua, Y., Vickers, T.A., Okunola, H.L., Bennett, C.F., and Krainer, A.R. (2008). Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. Am J Hum Genet 82, 834-848.

Huang, Q. and Pederson, T. A human U2 RNA mutant stalled in 3' end processing is impaired in nuclear import. (1999). Nucleic Acids Res.27, 1025–1031.

Huang, Y., and Steitz, J.A. (2001). Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. Mol Cell 7, 899-905.

Huang, Y., Yario, T.A., and Steitz, J.A. (2004). A molecular link between SR protein dephosphorylation and mRNA export. Proc Natl Acad Sci U S A 101, 9666-9670.

Huppler, A., Nikstad, L.J., Allmann, A.M., Brow, D.A., and Butcher, S.E. (2002). Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. Nat Struct Biol 9, 431-435.

Itoh, N., Yonehara, S., Ishii, A., Yonehara, M., Mizushima, S., Sameshima, M., Hase, A., Seto, Y., and Nagata, S. (1991). The polypeptide encoded by the cDNA for human cell surface antigen Fas can mediate apoptosis. Cell 66, 233-243.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell 78, 657-668.

Izquierdo, J.M. (2008a). Fas Splicing Regulation during Early Apoptosis Is Linked to Caspase-mediated Cleavage of U2AF65. Mol Biol Cell 19, 3299-3306.

Izquierdo, J.M. (2008b). Hu antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition. J Biol Chem 283, 19077-19084.

Izquierdo, J.M., Majos, N., Bonnal, S., Martinez, C., Castelo, R., Guigo, R., Bilbao, D., and Valcarcel, J. (2005). Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. Mol Cell 19, 475-484.

Jurica, M.S., and Moore, M.J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Mol Cell 12, 5-14.

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., H, W., Kitahara, T., Yoshida, T., Nakajima, H., et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat Chem Biol 3, 576-583.

Kalifa, Y., Huang, T., Rosen, L.N., Chatterjee, S., and Gavis, E.R. (2006). Glorund, a Drosophila hnRNP F/H homolog, is an ovarian repressor of nanos translation. Dev Cell 10, 291-301.

Kanadia, R.N., Clark, V.E., Punzo, C., Trimarchi, J.M., and Cepko, C.L. (2008). Temporal requirement of the alternative-splicing factor Sfrs1 for the survival of retinal neurons. Development 135, 3923-3933.

Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D., and Krainer, A.R. (2007). The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat Struct Mol Biol 14, 185-193.

Karni, R., Hippo, Y., Lowe, S.W., and Krainer, A.R. (2008). The splicing-factor oncoprotein SF2/ASF activates mTORC1. Proc Natl Acad Sci U S A 105, 15323-15327.

Kawashima, T., Pellegrini, M., and Chanfreau, G.F. (2009). Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. RNA 15, 2236-2247.

Kent, O.A., Reayi, A., Foong, L., Chilibeck, K.A., and MacMillan, A.M. (2003). Structuring of the 3' splice site by U2AF65. J Biol Chem 278, 50572-50577.

Kent, O.A., Ritchie, D.B., and Macmillan, A.M. (2005). Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway. Mol Cell Biol 25, 233-240.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet 11, 345-355.

Keriel, A., Mahuteau-Betzer, F., Jacquet, C., Plays, M., Grierson, D., Sitbon, M., and Tazi, J. (2009). Protection against retrovirus pathogenesis by SR protein inhibitors. PLoS One 4, e4533.

Kielkopf, C.L., Lucke, S., and Green, M.R. (2004). U2AF homology motifs: protein recognition in the RRM world. Genes Dev 18, 1513-1526.

Kim, E., Goren, A., and Ast, G. (2008). Alternative splicing: current perspectives. Bioessays 30, 38-47.

Kim, U., Wang, Y., Sanford, T., Zeng, Y., and Nishikura, K. (1994). Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. Proc Natl Acad Sci U S A 91, 11457-11461.

Knapp, G., Beckmann, J.S., Johnson, P.F., Fuhrman, S.A., and Abelson, J. (1978). Transcription and processing of intervening sequences in yeast tRNA genes. Cell 14, 221-236.

Koizumi, J., Okamoto, Y., Onogi, H., Mayeda, A., Krainer, A.R., and Hagiwara, M. (1999). The subcellular localization of SF2/ASF is regulated by direct interaction with SR protein kinases (SRPKs). J Biol Chem 274, 11125-11131.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet 41, 376-381.

Konarska, M.M., Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1985). Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature 313, 552-557.

Konarska, M.M., and Sharp, P.A. (1986). Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell 46, 845-855.

Konarska, M.M., and Sharp, P.A. (1987). Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell 49, 763-774.

Konarska, M.M., Vilardell, J., and Query, C.C. (2006). Repositioning of the reaction intermediate within the catalytic center of the spliceosome. Mol Cell 21, 543-553.

Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol. 3, 570-579.

Kotovic, K.M., Lockshon, D., Boric, L., and Neugebauer, K.M. (2003). Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. Mol Cell Biol 23, 5768-5779.

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

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

Kramer, A., Gruter, P., Groning, K., and Kastner, B. (1999). Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. J Cell Biol 145, 1355-1368.

Kramer, A., and Utans, U. (1991). Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. EMBO J 10, 1503-1509.

Krause, M., and Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in C. elegans. Cell 49, 753-761.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31, 147-157.

Kuersten, S., Lea, K., MacMorris, M., Spieth, J., and Blumenthal, T. (1997) Relationship between 3′ end formation and SL2-specific *trans*-splicing in polycistronic *Caenorhabditis elegans* pre-mRNA processing. RNA3, 269-278.

Kuhn, A., van Santen, M., Schwienhorst, A., Urlaub, H., and R., L. (2009). Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. RNA 15, 153-175.

Kuyumcu-Martinez, N.M., Wang, G.S., and Cooper, T.A. (2007). Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. Mol Cell 28, 68-78.

Labourier, E., Rossi, F., Gallouzi, I.E., Allemand, E., Divita, G., and Tazi, J. (1998). Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine-serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor. Nucleic Acids Res 26, 2955-2962.

Lacadie, S.A., and Rosbash, M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. Mol Cell 19, 65-75.

Lallena, M.J., Chalmers, K.J., Llamazares, S., Lamond, A.I., and Valcarcel, J. (2002). Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. Cell 109, 285-296.

Lamond, A.I., Konarska, M.M., Grabowski, P.J., and Sharp, P.A. (1988). Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. Proc Natl Acad Sci U S A 85, 411-415.

Lardelli, R.M., Thompson, J.X., Yates, J.R., 3rd, and Stevens, S.W. (2010). Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. RNA 16, 516-528.

Lareau, L.F., Brooks, A.N., Soergel, D.A., Meng, Q., and Brenner, S.E. (2007). The coupling of alternative splicing and nonsense-mediated mRNA decay. Adv Exp Med Biol 623, 190-211.

Lasda, E.L., Allen, M.A., and Blumenthal, T. (2010). Polycistronic pre-mRNA processing in vitro: snRNP and pre-mRNA role reversal in trans-splicing. Genes Dev 24, 1645-1658.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 19, 6860-6869.

Lee, H., Khan, R., and O'Keefe, M. (2008). Aniridia: current pathology and management. Acta Ophthalmol 86, 708-715.

Lee, J.E., and Cooper, T.A. (2009). Pathogenic mechanisms of myotonic dystrophy. Biochem Soc Trans 37, 1281-1286.

Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., and et al. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. Cell 80, 155-165.

Lejeune, F., and Maquat, L.E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr Opin Cell Biol 17, 309-315.

Lemaire, R., Prasad, J., Kashima, T., Gustafson, J., Manley, J.L., and Lafyatis, R. (2002). Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins. Genes Dev 16, 594-607.

Lev-Maor, G., Sorek, R., Shomron, N., and G., A. (2003). The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. Science 300, 1246-1247.

Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. (1996). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. Genes Dev 10, 1683-1698.

Li, H., Wang, J., Mor, G., and Sklar, J. (2008). A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. Science 321, 1357-1361.

Li, X., Wang, J., and Manley, J.L. (2005). Loss of splicing factor ASF/SF2 induces G2 cell cycle arrest and apoptosis, but inhibits internucleosomal DNA fragmentation. Genes Dev 19, 2705-2714.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456, 464-469.

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X.D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. Nat Struct Mol Biol 15, 819-826.

Lin, S., Xiao, R., Sun, P., Xu, X., and Fu, X.D. (2005). Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation. Mol Cell 20, 413-425.

Listerman, I., Sapra, A.K., and Neugebauer, K.M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. Nat Struct Mol Biol 13, 815-822.

Liu, C., Cheng, J., and Mountz, J.D. (1995). Differential expression of human Fas mRNA species upon peripheral blood mononuclear cell activation. Biochem J 310 (Pt 3), 957-963.

Liu, H.X., Cartegni, L., Zhang, M.Q., and Krainer, A.R. (2001a). A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nat Genet 27, 55-58.

Liu, Z., Luyten, I., Bottomley, M.J., Messias, A.C., Houngninou-Molango, S., Sprangers, R., Zanier, K., Krämer, A., and Sattler, M. (2001b). Structural basis for recognition of the intron branch site RNA by splicing factor 1. Science 294, 1098-1102.

Llorian, M., Schwartz, S., Clark, T.A., Hollander, D., Tan, L.Y., Spellman, R., Gordon, A., Schweitzer, A.C., de la Grange, P., Ast, G., and Smith, C.W. (2010). Position-

dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. Nat Struct Mol Biol 17, 1114-1123.

Loomis, R.J., Naoe, Y., Parker, J.B., Savic, V., Bozovsky, M.R., Macfarlan, T., Manley, J.L., and Chakravarti, D. (2009). Chromatin binding of SRp20 and ASF/SF2 and dissociation from mitotic chromosomes is modulated by histone H3 serine 10 phosphorylation. Mol Cell 33, 450-461.

Lopez-Bigas, N., Audit, B., Ouzounis, C., Parra, G., and Guigo, R. (2005). Are splicing mutations the most frequent cause of hereditary disease? FEBS Lett 579, 1900-1903.

Lorson, C.L., Hahnen, E., Androphy, E.J., and Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc Natl Acad Sci U S A 96, 6307-6311.

Lynch, M., and Richardson, A.O. (2002). The evolution of spliceosomal introns. Curr Opin Genet Dev 12, 701-710.

Ma, L., and Horvitz, H.R. (2009). Mutations in the Caenorhabditis elegans U2AF large subunit UAF-1 alter the choice of a 3' splice site in vivo. PLoS Genet 5, e1000708.

Ma, X.M., and Blenis, J. (2009). Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol 10, 307-318.

Madhani, H.D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. Annu Rev Genet 28, 1-26.

Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. Nature 416, 499-506.

Manley, J.L., and Krainer, A.R. (2010). A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev 24, 1073-1074.

Marchand, V., Mereau, A., Jacquenet, S., Thomas, D.,Mougin, A., Gattoni, R., Stevenin, J., and Branlant, C.(2002). A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNPA1cooperative binding. J. Mol. Biol. 323, 629-652.

Maris, C., Dominguez, C., and Allain, F.H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS J 272, 2118-2131.

Maroney, P. A., Denker, J. A., Darzynkiewicz, E., Laneve, R., and Nilsen, T. W. (1995) Most mRNAs in the nematode *Ascaris lumbricoides* are *trans*-spliced: a role for spliced leader addition in translational efficiency. RNA1, 714–723.

Matera, A.G., Terns, R.M., and Terns, M.P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol 8, 209-220.

Matlin, A.J., Clark, F., and Smith, C.W. (2005). Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 6, 386-398.

Mayas, R.M., Maita, H., and Staley, J.P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat Struct Mol Biol 13, 482-490.

Mayeda, A., and Krainer, A. R. (1992). Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2.Cell 68, 365-75.

McAfee, J.G., Soltaninassab, S.R., Lindsay, M.E., and LeStourgeon, W.M. (1996). Proteins C1 and C2 of heterogeneous nuclear ribonucleoprotein complexes bind RNA in a highly cooperative fashion: support for their contiguous deposition on pre-mRNA during transcription. Biochemistry 35, 1212-1222.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev 11, 3306-3318.

McGlincy, N.J., and Smith, C.W. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem Sci 33, 385-393.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P.H., and Higuchi, M. (1996). A mammalian RNA editing enzyme. Nature 379, 460-464.

Mendes Soares, L.M., and Valcarcel, J. (2006). The expanding transcriptome: the genome as the 'Book of Sand'. Embo J 25, 923-931.

Merendino, L., Guth, S., Bilbao, D., Martinez, C., and Valcarcel, J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. Nature 402, 838-841.

Mewes, H. W., Albermann, K., Bahr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K.,Kleine, K., Maierl, A., Oliver, S. G., et al. (1997). Overview of the yeast genome. Nature 387, 7-65.

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

Michel, F., and Ferat, J.L. (1995). Structure and activities of group II introns. Annu Rev Biochem 64, 435-461.

Michlewski, G., Sanford, J.R., and Caceres, J.F. (2008). The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. Mol Cell 30, 179-189.

Mili, S., Shu, H. J., Zhao, Y., and Piñol-Roma, S. (2001). Distinct RNP complexes of shuttling hnRNP proteins with pre-mRNA and mRNA: candidate intermediates in formation and export of mRNA.Mol Cell Biol 21, 7307-19.

Miller, J.W., Urbinati, C.R., Teng-Umnuay, P., Stenberg, M.G., Byrne, B.J., Thornton, C.A., and Swanson, M.S. (2000). Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. EMBO J 19, 4439-4448.

Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol Cell 3, 697-705.

Mitra, J., and Enders, G.H. (2004). Cyclin A/Cdk2 complexes regulate activation of Cdk1 and Cdc25 phosphatases in human cells. Oncogene 23, 3361-3367.

Mochizuki, H., Nishi, T., Bruner, J.M., Lee, P.S., Levin, V.A., and Saya, H. (1992). Alternative splicing of neurofibromatosis type 1 gene transcript in malignant brain tumors: PCR analysis of frozen-section mRNA. Mol Carcinog 6, 83-87.

Mohr, A., Zwacka, R.M., Jarmy, G., Buneker, C., Schrezenmeier, H., Dohner, K., Beltinger, C., Wiesneth, M., Debatin, K.M., and Stahnke, K. (2005). Caspase-8L expression protects CD34+ hematopoietic progenitor cells and leukemic cells from CD95-mediated apoptosis. Oncogene 24, 2421-2429.

Mongelard, F., Labrador, M., Baxter, E.M., Gerasimova, T.I., and Corces, V.G. (2002). Trans-splicing as a novel mechanism to explain interallelic complementation in Drosophila. Genetics 160, 1481-1487.

Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136, 688-700.

Moore, M.J., and Sharp, P.A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature 365, 364-368.

Moore, M.J., Wang, Q., Kennedy, C.J., and Silver, P.A. (2010). An alternative splicing network links cell-cycle control to apoptosis. Cell 142, 625-636.

Mordes, D., Luo, X., Kar, A., Kuo, D., Xu, L., Fushimi, K., Yu, G., Sternberg, P.J., and Wu, J. (2006). Pre-mRNA splicing and retinitis pigmentosa. Mol Vis. 12, 1259-1271.

Mordes, D., Yuan, L., Xu, L., Kawada, M., Molday, R.S., and Wu, J.Y. (2007). Identification of photoreceptor genes affected by PRPF31 mutations associated with autosomal dominant retinitis pigmentosa. Neurobiol Dis 26, 291-300.

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. & Bordonne, R. (2002). Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. Mol. Cell9, 891-901.

Munoz, M.J., Perez Santangelo, M.S., Paronetto, M.P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J.J., et al. (2009). DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. Cell 137, 708-720.

Muraki, M., Ohkawara, B., Hosoya, T., Onogi, H., Koizumi, J., Koizumi, T., Sumi, K., Yomoda, J., Murray, M.V., Kimura, H., et al. (2004). Manipulation of alternative splicing by a newly developed inhibitor of Clks. J Biol Chem 279, 24246-24254.

Murphy, W.J., Watkins, K.P., and Agabian, N. (1986). Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for trans splicing. Cell 47, 517-525.

Nagai, K., Oubridge, C., Jessen, T.H., Li, J., and Evans, P.R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature 348, 515-520.

Nagata, S. (1997). Apoptosis by death factor. Cell 88, 355-365.

Nahkuri, S., Taft, R.J., and Mattick, J.S. (2009). Nucleosomes are preferentially positioned at exons in somatic and sperm cells. Cell Cycle 8, 3420-3424.

Nakajima, H., Hori, Y., Terano, H., Okuhara, M., Manda, T., Matsumoto, S., and Shimomura, K. (1996). New antitumor substances, FR901463, FR901464 and FR901465. II. Activities against experimental tumors in mice and mechanism of action. J Antibiot (Tokyo) 49, 1204-1211.

Nayler, O., Stamm, S., and Ullrich, A. (1997). Characterization and comparison of four serine- and arginine-rich (SR) protein kinases. Biochem J 326 (Pt 3), 693-700.

Nelson, K.K., and Green, M.R. (1989). Mammalian U2 snRNP has a sequence-specific RNA-binding activity. Genes Dev 3, 1562-1571.

Nembaware, V., Wolfe, K.H., Bettoni, F., Kelso, J., and Seoighe, C. (2004). Allele-specific transcript isoforms in human. FEBS Lett 577, 233-238.

Nesic, D., Tanackovic, G., and Kramer, A. (2004). A role for Cajal bodies in the final steps of U2 snRNP biogenesis. J Cell Sci 117, 4423-4433.

Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A., and Mann, M. (1998). Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex.Nat Genet. 20, 46-50.

Newby, M.I., and Greenbaum, N.L. (2002). Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. Nat Struct Biol 9, 958-965.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457-463.

Oberstrass, F.C., Auweter, S.D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D.L., and Allain, F.H. (2005). Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science 309, 2054-2057.

Ogle, J.M., and Ramakrishnan, V. (2005). Structural insights into translational fidelity. Annu Rev Biochem 74, 129-177.

Oh, J.J., Razfar, A., Delgado, I., Reed, R.A., Malkina, A., Boctor, B., and Slamon, D.J. (2006). 3p21.3 tumor suppressor gene H37/Luca15/RBM5 inhibits growth of human lung cancer cells through cell cycle arrest and apoptosis. Cancer Res 66, 3419-3427.

Okamoto, Y., Onogi, H., Honda, R., Yasuda, H., Wakabayashi, T., Nimura, Y., and Hagiwara, M. (1998). cdc2 kinase-mediated phosphorylation of splicing factor SF2/ASF. Biochem Biophys Res Commun 249, 872-878.

Pacheco, T.R., Coelho, M.B., Desterro, J.M., Mollet, I., and Carmo-Fonseca, M. (2006). In vivo requirement of the small subunit of U2AF for recognition of a weak 3' splice site. Mol Cell Biol 26, 8183-8190.

Paronetto, M.P., Achsel, T., Massiello, A., Chalfant, C.E., and Sette, C. (2007). The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. J Cell Biol 176, 929-939.

Pascual, M., Vicente, M., Monferrer, L., and Artero, R. (2006). The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. Differentiation 74, 65-80.

Patel, A.A., and Steitz, J.A. (2003). Splicing double: insights from the second spliceosome. Nat Rev Mol Cell Biol 4, 960-970.

Patel, N. A., Kaneko, S., Apostolatos, H. S., Bae, S. S., Watson, J. E., Davidowitz, K., Chappell, D. S., Birnbaum, M. J., Cheng, J. Q., and Cooper, D. R. (2005). Molecular and genetic studies imply Akt-mediated signaling promotes protein kinase CbetaII alternative splicing via phosphorylation of serine/arginine-rich splicing factor SRp40.J Biol Chem 280, 14302-9.

Patterson, J.B., and Samuel, C.E. (1995). Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. Mol Cell Biol 15, 5376-5388.

Pelisch, F., Gerez, J., Druker, J., Schor, I.E., Munoz, M.J., Risso, G., Petrillo, E., Westman, B.J., Lamond, A.I., Arzt, E., and Srebrow, A. (2010). The serine/arginine-rich protein SF2/ASF regulates protein sumoylation. Proc Natl Acad Sci U S A 107, 16119-16124.

Peter, M.E., Budd, R.C., Desbarats, J., Hedrick, S.M., Hueber, A.O., Newell, M.K., Owen, L.B., Pope, R.M., Tschopp, J., Wajant, H., et al. (2007). The CD95 receptor: apoptosis revisited. Cell 129, 447-450.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29, e45.

Philips, A.V., Timchenko, L.T., and Cooper, T.A. (1998). Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. Science 280, 737-741.

Pikielny, C.W., Rymond, B.C., and Rosbash, M. (1986). Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. Nature 324, 341-345.

Piñol-Roma, S., and Dreyfuss, G. (1992). Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature355, 730-2

Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K., Li, J., and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature 458, 475-480.

Query, C.C., McCaw, P.S., and Sharp, P.A. (1997). A minimal spliceosomal complex A recognizes the branch site and polypyrimidine tract. Mol Cell Biol 17, 2944-2953.

Query, C.C., Moore, M.J., and Sharp, P.A. (1994). Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. Genes Dev 8, 587-597.

Rain, J.C., Rafi, Z., Rhani, Z., Legrain, P., and Kramer, A. (1998). Conservation of functional domains involved in RNA binding and protein-protein interactions in human and Saccharomyces cerevisiae pre-mRNA splicing factor SF1. RNA 4, 551-565.

Raker, V.A., Plessel, G., and Luhrmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. Embo J 15, 2256–2269

Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res 12, 1231-1245.

Reed, R. (1989). The organization of 3' splice-site sequences in mammalian introns. Genes Dev 3, 2113-2123.

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

Rhode, B.M., Hartmuth, K., Westhof, E., and Luhrmann, R. (2006). Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. EMBO J 25, 2475-2486.

Rideau, A.P., Gooding, C., Simpson, P.J., Monie, T.P., Lorenz, M., Huttelmaier, S., Singer, R.H., Matthews, S., Curry, S., and Smith, C.W. (2006). A peptide motif in Raver1 mediates splicing repression by interaction with the PTB RRM2 domain. Nat Struct Mol Biol 13, 839-848.

Rieux-Laucat, F., Le Deist, F., Hivroz, C., Roberts, I.A., Debatin, K.M., Fischer, A., and de Villartay, J.P. (1995). Mutations in Fas associated with human lymphoproliferative syndrome and autoimmunity. Science 268, 1347-1349.

Roca, X., and Krainer, A.R. (2009). Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA. Nat Struct Mol Biol 16, 176-182.

Roca, X., Olson, A.J., Rao, A.R., Enerly, E., Kristensen, V.N., Borresen-Dale, A.L., Andresen, B.S., Krainer, A.R., and Sachidanandam, R. (2008). Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics. Genome Res 18, 77-87.

Roesler, J., Izquierdo, J.M., Ryser, M., Rosen-Wolff, A., Gahr, M., Valcarcel, J., Lenardo, M.J., and Zheng, L. (2005). Haploinsufficiency, rather than the effect of an excessive production of soluble CD95 (CD95{Delta}TM), is the basis for ALPS Ia in a family with duplicated 3' splice site AG in CD95 intron 5 on one allele. Blood 106, 1652-1659.

Rogaev, E.I., Grigorenko, A.P., Faskhutdinova, G., Kittler, E.L., and Moliaka, Y.K. (2009). Genotype analysis identifies the cause of the "royal disease". Science 326, 817.

Rooke, N., Markovtsov, V., Cagavi, E., and Black, D. L. (2003). Roles for SR proteins and hnRNP A1 in the regulation of c-srcexon N1. Mol. Cell. Biol. 23, 1874-1884.

Rossi, F., Labourier, E., Forne, T., Divita, G., Derancourt, J., Riou, J.F., Antoine, E., Cathala, G., Brunel, C., and Tazi, J. (1996). Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. Nature 381, 80-82.

Roth, M.B., Murphy, C., and Gall, J.G. (1990). A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. J Cell Biol 111, 2217-2223.

Roth, M.B., Zahler, A.M., and Stolk, J.A. (1991). A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. J Cell Biol 115, 587-596.

Roybal, G.A., and Jurica, M.S. (2010). Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. Nucleic Acids Res.

Ruberti, G., Cascino, I., Papoff, G., and Eramo, A. (1996). Fas splicing variants and their effect on apoptosis. Adv Exp Med Biol 406, 125-134.

Ruskin, B., and Green, M.R. (1985). Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. Nature 317, 732-734.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38, 317-331.

Ruskin, B., Zamore, P., and MR., G. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell 52., 207-219.

Russell, J.H. (1995a). Activation-induced death of mature T cells in the regulation of immune responses. Curr Opin Immunol 7, 382-388.

Russell, J.H. (1995b). Activation-induced death of mature T cells in the regulation of immune responses. Curr. Opin. Immunol. 7, 382-388.

Saeki, T., Ouchi, M., and Ouchi, T. (2009). Physiological and oncogenic Aurora-A pathway. Int J Biol Sci 5, 758-762.

Sanford, J.R., Coutinho, P., Hackett, J.A., Wang, X., Ranahan, W., and Caceres, J.F. (2008). Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. PLoS One 3, e3369.

Sanford, J.R., Gray, N.K., Beckmann, K., and Caceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. Genes Dev 18, 755-768.

Sapra, A.K., Anko, M.L., Grishina, I., Lorenz, M., Pabis, M., Poser, I., Rollins, J., Weiland, E.M., and Neugebauer, K.M. (2009). SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. Mol Cell 34, 179-190.

Sato, H., Hosoda, N., and Maquat, L.E. (2008). Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. Mol Cell 29, 255-262.

Schindler, S., Szafranski, K., Hiller, M., Ali, G.S., Palusa, S.G., Backofen, R., Platzer, M., and Reddy, A.S. (2008). Alternative splicing at NAGNAG acceptors in Arabidopsis thaliana SR and SR-related protein-coding genes. BMC Genomics 9, 159.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell 101, 671-684.

Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. Nat Struct Mol Biol 16, 990-995.

Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Krämer, A., and Sattler, M. (2003). Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. Mol Cell. 11, 965-976.

Seraphin, B., and Rosbash, M. (1989). Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. Cell 59, 349-358.

Sharma, S., Falick, A.M., and Black, D.L. (2005). Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex. Mol Cell 19, 485-496.

Sharma, S., Kohlstaedt, L.A., Damianov, A., Rio, D.C., and Black, D.L. (2008). Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. Nat Struct Mol Biol 15, 183-191.

Sharp, P.A. (1987). Splicing of messenger RNA precursors. Science 235, 766-771.

Sharp, P.A., and Burge, C.B. (1997). Classification of introns: U2-type or U12-type. Cell 91, 875-879.

Shatsky, M., Nussinov, R., and Wolfson, H.J. (2004). A method for simultaneous alignment of multiple protein structures. Proteins 56, 143-156.

Shen, H., and Green, M.R. (2004). A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. Mol Cell 16, 363-373.

Shen, H., and Green, M.R. (2006). RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. Genes Dev 20, 1755-1765.

Shen, H., Kan, J.L., and Green, M.R. (2004). Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol Cell 13, 367-376.

Shi, Y., Reddy, B., and Manley, J.L. (2006). PP1/PP2A phosphatases are required for the second step of Pre-mRNA splicing and target specific snRNP proteins. Mol Cell 23, 819-829.

Shin, C., Feng, Y., and Manley, J.L. (2004). Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. Nature 427, 553-558.

Shuman, S. (2001). Structure, mechanism, and evolution of the mRNA capping apparatus. Prog Nucleic Acid Res Mol Biol 66, 1-40.

Singh, N.N., Androphy, E.J., and Singh, R.N. (2004). An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy. Biochem Biophys Res Commun 315, 381-388.

Singh, R., and Valcarcel, J. (2005). Building specificity with nonspecific RNA-binding proteins. Nat Struct Mol Biol 12, 645-653.

Sinha, R., Allemand, E., Zhang, Z., Karni, R., Myers, M.P., and Krainer, A.R. (2010). Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. Mol Cell Biol 30, 2762-2774.

Sinha, R., Nikolajewa, S., Szafranski, K., Hiller, M., Jahn, N., Huse, K., Platzer, M., and Backofen, R. (2009). Accurate prediction of NAGNAG alternative splicing. Nucleic Acids Res 37, 3569-3579.

Smith, C.W., Chu, T.T., and Nadal-Ginard, B. (1993). Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. Mol Cell Biol 13, 4939-4952.

Smith, C.W., and Nadal-Ginard, B. (1989). Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. Cell 56, 749-758.

Smith, C.W., and Valcarcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci 25, 381-388.

Smith, P.J., Zhang, C., Wang, J., Chew, S.L., Zhang, M.Q., and Krainer, A.R. (2006). An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum Mol Genet 15, 2490-2508.

Soares, L.M., Zanier, K., Mackereth, C., Sattler, M., and Valcarcel, J. (2006). Intron removal requires proofreading of U2AF/3' splice site recognition by DEK. Science 312, 1961-1965.

Soret, J., Bakkour, N., Maire, S., Durand, S., Zekri, L., Gabut, M., Fic, W., Divita, G., Rivalle, C., Dauzonne, D., et al. (2005). Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. Proc Natl Acad Sci U S A 102, 8764-8769.

Spadaccini, R., Reidt, U., Dybkov, O., Will, C., Frank, R., Stier, G., Corsini, L., C, W.M., Lührmann, R., and M, S. (2006). Biochemical and NMR analyses of an SF3b155-p14-U2AF-RNA interaction network involved in branch point definition during pre-mRNA splicing. RNA. 2006 Mar;12(3):410-25. 12, 410-425.

Spellman, R., Llorian, M., and Smith, C.W. (2007). Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. Mol Cell 27, 420-434.

Spellman, R., and Smith, C.W. (2006). Novel modes of splicing repression by PTB. Trends Biochem Sci 31, 73-76.

Spies, N., Nielsen, C.B., Padgett, R.A., and Burge, C.B. (2009). Biased chromatin signatures around polyadenylation sites and exons. Mol Cell 36, 245-254.

Spingola, M., Grate, L., Haussler, D., and Ares, M., Jr. (1999). Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae. RNA 5, 221-234.

Sridharan, V., and Singh, R. (2007). A conditional role of U2AF in splicing of introns with unconventional polypyrimidine tracts. Mol Cell Biol 27, 7334-7344.

Staley, J.P., and Woolford, J.L., Jr. (2009). Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr Opin Cell Biol 21, 109-118.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., and Soreq, H. (2005). Function of alternative splicing. Gene 344, 1-20.

Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., and Abelson, J. (2002). Composition and functional characterization of the yeast spliceosomal penta-snRNP. Mol Cell 9, 31-44.

Strasser, A., Jost, P.J., and Nagata, S. (2009). The many roles of FAS receptor signaling in the immune system. Immunity 30, 180-192.

Straus, S.E., Jaffe, E.S., Puck, J.M., Dale, J.K., Elkon, K.B., Rosen-Wolff, A., Peters, A.M., Sneller, M.C., Hallahan, C.W., Wang, J., et al. (2001). The development of lymphomas in families with autoimmune lymphoproliferative syndrome with germline Fas mutations and defective lymphocyte apoptosis. Blood 98, 194-200.

Sun, J.S., and Manley, J.L. (1995). A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. Genes Dev 9, 843-854.

Sun, Q., Mayeda, A., Hampson, R. K., Krainer, A. R. andRottman, F. M. (1993) General splicing factor SF2/ASFpromotes alternative splicing by binding to an exonic splicingenhancer. Genes Dev. 7, 2598- 2608.

Sun, S., Zhang, Z., Sinha, R., Karni, R., and Krainer, A.R. (2010). SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. Nat Struct Mol Biol 17, 306-312.

Sutton, R.E., and Boothroyd, J.C. (1986). Evidence for trans splicing in trypanosomes. Cell 47, 527-535.

Swanson, M.S., Nakagawa, T.Y., LeVan, K., and Dreyfuss, G. (1987). Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. Mol Cell Biol 7, 1731-1739.

Takahashi, T., Tanaka, M., Brannan, C.I., Jenkins, N.A., Copeland, N.G., Suda, T., and Nagata, S. (1994). Generalized lymphoproliferative disease in mice, caused by a point mutation in the Fas ligand. Cell 76, 969-976.

Tarn, W. Y., and Steitz, J. A. (1996a). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. Cell 84, 801-11.

Tarn, W. Y., and Steitz, J. A. (1996b). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science 273, 1824-32.

Taylor, R.C., Cullen, S.P., and Martin, S.J. (2008). Apoptosis: controlled demolition at the cellular level. Nat Rev Mol Cell Biol 9, 231-241.

Tejedor, J.R., and Valcarcel, J. (2010). Gene regulation: Breaking the second genetic code. Nature 465, 45-46.

Thickman, K.R., Swenson, M.C., Kabogo, J.M., Gryczynski, Z., and Kielkopf, C.L. (2006). Multiple U2AF65 binding sites within SF3b155: thermodynamic and spectroscopic characterization of protein-protein interactions among pre-mRNA splicing factors. J Mol Biol 356, 664-683.

Thiede, B., Dimmler, C., Siejak, F., and Rudel, T. (2001). Predominant identification of RNA-binding proteins in Fas-induced apoptosis by proteome analysis. J Biol Chem 276, 26044-26050.

Teigelkamp S, Newman AJ, Beggs JD. (1995). Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. EMBO J 14, 2602-12.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcarcel, J., and Guigo, R. (2009). Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol 16, 996-1001.

Tintaru, A.M., Hautbergue, G.M., Hounslow, A.M., Hung, M.L., Lian, L.Y., Craven, C.J., and Wilson, S.A. (2007). Structural and functional analysis of RNA and TAP binding to SF2/ASF. EMBO Rep 8, 756-762.

Tsai, K.W., Chan, W.C., Hsu, C.N., and Lin, W.C. (2010). Sequence features involved in the mechanism of 3' splice junction wobbling. BMC Mol Biol 11, 34.

Tsai, K.W., Tarn, W.Y., and Lin, W.C. (2007). Wobble splicing reveals the role of the branch point sequence-to-NAGNAG region in 3' tandem splice site selection. Mol Cell Biol 27, 5835-5848.

Tsai, K.W., Tseng, H.C., and Lin, W.C. (2008). Two wobble-splicing events affect ING4 protein subnuclear localization and degradation. Exp Cell Res 314, 3130-3141.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science 302, 1212-1215.

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature 444, 580-586.

Umen, J. G., and Guthrie, C. (1995). Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. RNA 1, 584-97.

Vagner, S., Vagner, C., and Mattaj, I.W. (2000). The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. Genes Dev 14, 403-413.

Valadkhan, S. (2005). snRNAs as the catalysts of pre-mRNA splicing. Curr Opin Chem Biol 9, 603-608.

Valadkhan, S. (2007). The spliceosome: a ribozyme at heart? Biol Chem 388, 693-697.

Valadkhan, S., and Manley, J.L. (2001). Splicing-related catalysis by protein-free snRNAs. Nature 413, 701-707.

Valadkhan, S., and Manley, J.L. (2003). Characterization of the catalytic activity of U2 and U6 snRNAs. RNA 9, 892-904.

Valcarcel, J., Gaur, R.K., Singh, R., and Green, M.R. (1996). Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science 273, 1706-1709.

Valcarcel, J., and Gebauer, F. (1997). Post-transcriptional regulation: the dawn of PTB. Curr Biol 7, R705-708.

van Venrooij, W.J., Hoet, R., Castrop, J., Hageman, B., Mattaj, I.W., and van de Putte, L.B. (1990). Anti-(U1) small nuclear RNA antibodies in anti-small nuclear ribonucleoprotein sera from patients with connective tissue diseases. J Clin Invest 86, 2154-2160.

Vella, V., Pandini, G., Sciacca, L., Mineo, R., Vigneri, R., Pezzino, V., and Belfiore, A. (2002). A novel autocrine loop involving IGF-II and the insulin receptor isoform-A stimulates growth of thyroid cancer. J Clin Endocrinol Metab 87, 245-254.

Venables, J.P. (2006). Unbalanced alternative splicing and its significance in cancer. Bioessays 28, 378-386.

Vidal, V.P., Verdone, L., Mayes, A.E., and Beggs, J.D. (1999). Characterization of U6 snRNA–protein interactions. RNA 5: 1470-1481.

Villa, T., and Guthrie, C. (2005). The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes Dev 19, 1894-1904.

Wahl, M.C., Will, C.L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell 136, 701-718.

Wang, C., Chua, K., Seghezzi, W., Lees, E., Gozani, O., and Reed, R. (1998a). Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. Genes Dev 12, 1409-1414.

Wang, E., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S., Schroth, G., and CB., B. (2008a). Alternative isoform regulation in human tissue transcriptomes. Nature 456:, 470-476.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008b). Alternative isoform regulation in human tissue transcriptomes. Nature 456, 470-476.

Wang, G.S., and Cooper, T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet 8, 749-761.

Wang, H.Y., Lin, W., Dyck, J.A., Yeakley, J.M., Songyang, Z., Cantley, L.C., and Fu, X.D. (1998b). SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. J Cell Biol 140, 737-750.

Wassarman, D.A., and Steitz, J.A. (1992). Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. Science 257, 1918-1925.

Watanabe-Fukunaga, R., Brannan, C.I., Copeland, N.G., Jenkins, N.A., and Nagata, S. (1992). Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. Nature 356, 314-317.

Watson, F.L., Puttmann-Holgado, R., Thomas, F., Lamar, D.L., Hughes, M., Kondo, M., Rebel, V.I., and Schmucker, D. (2005). Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science 309, 1874-1878.

Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J., and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453, 1239-1243.

Will, C.L., and Luhrmann, R. (1997). Protein functions in pre-mRNA splicing. Curr Opin Cell Biol 9, 320-328.

Will, C.L., and Luhrmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. Curr Opin Cell Biol 13, 290-301.

Will, C.L., and Luhrmann, R. (2005). Splicing of a rare class of introns by the U12-dependent spliceosome. Biol Chem 386, 713-724.

Will, C.L., Schneider, C., MacMillan, A.M., Katopodis, N.F., Neubauer, G., Wilm, M., Luhrmann, R., and Query, C.C. (2001). A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. EMBO J 20, 4536-4546.

Will, C.L., Urlaub, H., Achsel, T., Gentzel, M., Wilm, M., and Luhrmann, R. (2002). Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. EMBO J 21, 4978-4988.

Wu, H., Sun, S., Tu, K., Gao, Y., Xie, B., Krainer, A.R., and Zhu, J. (2010). A splicing-independent function of SF2/ASF in microRNA processing. Mol Cell 38, 67-77.

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

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

Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature 402, 832-835.

Xu, R.M., Jokhan, L., Cheng, X., Mayeda, A., and Krainer, A.R. (1997). Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. Structure 5, 559-570.

Xu, X., Yang, D., Ding, J.H., Wang, W., Chu, P.H., Dalton, N.D., Wang, H.Y., Bermingham, J.R., Jr., Ye, Z., Liu, F., et al. (2005). ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. Cell 120, 59-72.

Xu, Y.Z., and Query, C.C. (2007). Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. Mol Cell 28, 838-849.

Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.S., Zhang, C., Yeo, G., Black, D.L., Sun, H., et al. (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Mol Cell 36, 996-1006.

Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.D., and Gage, F.H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol 16, 130-137.

Zahler, A.M., Lane, W.S., Stolk, J.A., and Roth, M.B. (1992). SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev 6, 837-847.

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

Zhang, G., Guo, G., Hu, X., Zhang, Y., Li, Q., Li, R., Zhuang, R., Lu, Z., He, Z., Fang, X., et al. (2010). Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. Genome Res 20, 646-654.

Zhang, M.Q. (1998). Statistical features of human exons and their flanking regions. Hum Mol Genet 7, 919-932.

Zhang, Z., and Krainer, A.R. (2004). Involvement of SR proteins in mRNA surveillance. Mol Cell 16, 597-607.

Zhou, Z., Licklider, L.J., Gygi, S.P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. Nature 419, 182-185.

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

Zorio, D.A., and Blumenthal, T. (1999). Both subunits of U2AF recognize the 3' splice site in Caenorhabditis elegans. Nature 402, 835-838.

Zuo, P., and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes Dev 10, 1356-1368.

ACKNOWLEDGEMENTS

Sísifo, el héroe absurdo, condenado a empujar eternamente una roca hasta lo alto de una montaña, desde donde la piedra volvía a caer por su propio peso…

Estoy totalmente de acuerdo con Camus cuando afirma que no hay castigo más terrible que el trabajo inútil, sin esperanza, y el ser consciente de ello. También no hay más verdad que lo que pensamos mientras descendemos de nuevo a por la roca es lo que somos, cómo somos. Sin duda, he recogido innumerables veces esa piedra y la he vuelto a subir y, para ser sincera, muchas veces la roca ganó, yo era la roca.

Sin embargo, el tormento de Sísifo era inútil y eterno y, estando confinado en los infiernos, Sísifo se encontraba lejos de los placeres de la vida, sin compañía. Aquí ya no puedo aplicar este mito a mi experiencia durante estos últimos cinco años. Ahora, desde lo alto de la colina, sé que este camino ha llegado a su fin, que el recoger y subir esta piedra continuamente me dio la experiencia suficiente para afrontar nuevas caídas y volver a empujar con más fuerza, más ímpetu y más sabiduría. No estaría aquí, conquistando este monte si no hubiese sido por vosotros, que de alguna manera u otra habéis hecho cada descenso y cada subida más fáciles.

Quería agradecer a aquéllos que ya desde pequeñita, y siendo yo como era (y todavía en parte cómo soy, lo siento), me habéis apoyado en todas mis decisiones y me habéis enseñado que hay que luchar por lo que uno quiere, que nada es fácil, pero que con suficiente tesón y perseverancia se puede conseguir.

Gracias a aquéllos que me habéis acompañado durante estos cinco años. A los que me habéis enseñado todo lo que sé y que me habéis guiado durante este camino, sugiriéndome y ayudándome a encontrar nuevas rutas, más firmes y con menos pendientes, que me han permitido llegar a destino de una manera más segura y, afortunadamente, exitosa. También gracias por no haber tenido reparos en decir honestamente vuestra opinión sobre mi trabajo, ha sido lo mejor, profesional y personalmente.

Tengo que agradecer a todos los que habéis soportado mis quejas y mis frustraciones, cómo antes mencioné, la roca ganó muchas veces. Pero hablar con vosotros permitió darme cuenta que todos somos en cierta medida Sísifo, todos tenemos nuestras rocas, y que unas pesan más que otras. Sin duda, la mía fue más ligera gracias a vosotros y hacía el descenso en su busca con gozo.

También agradezco a todos aquellos que habéis contribuido a alegrar mi paso por esta etapa de mi vida, ya sea con un café o una cerveza en la mano, viendo o jugando a fútbol o volley o cuidándome hasta en los más mínimos detalles, ya fuese en casa, en el laboratorio o en la distancia.

Desde aquí me doy cuenta de que éste es sólo un alto en el camino, una nueva etapa y una nueva roca me esperan. Si me atrevo a afrontar este otro reto es por vosotros. Gracias.