UNIVERSITY OF NOVA GORICA
GRADUATE SCHOOL

AN RNA-BASED THERAPEUTIC APPROACH FOR INHERITED MIS-SPLICING DISEASES

DISSERTATION

Andrea Dal Mas

Mentor: Dr. Franco Pagani, MD

Nova Gorica, 2014

ABSTRACT

Over the last 10 years several attempts have been carried out in order to provide efficient tools to redirect aberrant splicing in disease-causing genes. Designing a splicing correction approach to different disease-causing mutation is a process that required in-depth study of the specific splicing biology behind each exon and, more in general, behind each gene under analysis, in order to provide the most feasible method to revert the mechanism that leads to aberrant splicing. In this work, I engineered the 5'-tail of the U1 small nuclear RNA in order to create modified U1s –named Exon Specific U1s (ExSpe U1s)- that targeting the intronic region downstream the 5' splice site of a defective exon, rescue its splicing. To understand their therapeutic potential and mechanism of action, I focussed on two exons, SMN2 exon 7 and SPINK5 exon 11 associated with spinal muscular atrophy (SMA) and netherton syndrome (NS), respectively. In both cases, a C-to-T synonymous substitution induces skipping of the corresponding exon from the final mRNA. Similarly to what previously reported for the SMN mutation, I found that the exonic +9C>T transversion in NS creates an hnRNP A1 dependent exonic splicing silencer. In addition, I showed that this mutation interferes with two adjacent motifs recognized by Tra2β. Using minigene models, I have identified the most active ExSpe U1 molecules which were tested in the corresponding cellular models of the diseases. In SMA, I transduced with lentiviral particles expressing ExSpe U1s primary fibroblasts and I created stable Hek293 Flp-In clones that express ExSpeU1 from a single copy of the gene. In affected fibroblasts I was able to recover the correct SMN full length transcript and SMN protein level. In the Hek293 Flp-In system, a single chromosome integrated copy of ExSpe U1 gene is sufficient to modulate endogenous SMN2 exon 7 splicing. In NS primary keratinocytes, lentiviral mediated expression of ExSpe U1s recover

the correct full length SPINK5 isoform and the protein levels in a dose dependent manner.

In comparison to antisense oligonucleotide (AON)-mediated SMN2 splicing correction, I showed that ExSpeU1s exploit a completely different mechanism. Even if AONs and ExSpe U1 partially interact with the same intronic sequences, ExSpeU1 -but not AON- is sensitive to hnRNPA1 overexpression. Using U1 snRNA decoy, I sequestered the endogenous U1 particles, making them not available at canonical splice sites and I showed that the ExSpe U1 is able to functionally substitute the U1 snRNP promoting *per se* exon definition, while AON is not.

EMSA experiments demonstrated that ExSpe U1s and U1 snRNP recruit the same protein complexes and I investigated the contribution of U1-specific proteins in ExSpe U1-mediated splicing rescue creating ExSpe U1-A or U1-70K deficient molecules. While U1-A seems to not have a crucial role in splicing rescue, U1-70K is required to efficiently promote exon inclusion.

All together, these results indicate that ExSpeU1s are active in cellular models of spinal muscular atrophy and netherton syndrome and their capacity to act independently from the presence of the U1 particle make them new tools to increase exon inclusion in those genes in which mutations affect the definition of the exon and they could be an alternative or complementary strategy when an antisense-based rescue approach has no efficacy or has only a partial effect.

IZVLEČEK

V zadnjih desetih letih je bilo izvedenih veliko poskusov z namenom iskanja ustreznih orodij za preusmeritev napačnega spajanja eksonov v genih, ki povročajo bolezni. Načrtovanje tovrstnih pristopov pri različnih bolezenskih mutacijah zahteva poglobljen študij individualnih primerov mutacij napačnega spajanja ter prav tako na nivoju posameznih okvarjenih genov. Le na ta način lahko razvijemo ustrezne metode, ki popravijo mehanizem napačnega spajanja. V tem delu sem ustvaril 5'-rep U1 male jedrne RNA (U1 snRNA) z namenom tvorbe spremenjenih U1 molekul, poimenovanih Ekson specifične RNA (ExSpe U1), ki popravljajo napačno spajanje preko vezave na tarčne regije v intronskih zaporedjih navzdol od 5' spojitvenih mest okvarjenih eksonov. Z namenom razumevanja njihove terapevtske vloge in mehanizma delovanja sem se osredotočil na primera dveh okvarjenih eksonov: na ekson 7 gena SMN2 (vzrok spinalne mišične atrofije, SMA) ter na ekson 11 gena SPINK5 (vzrok Nethertonovega sindroma, NS). V obeh primerih sinonimna sprememba C v T povroči izpustitev omenjenih eksonov v končni mRNA. Podobno kot je bilo predhodno opisano za primer SMN mutacije, sem ugotovil, da točkovna mutacija v eksonski regiji +9C>T pri sindromu NS, povzroči nastanek od hnRNP A1 odvisnega eksonskega utiševalca spajanja. V nadaljevanju sem pokazal, da ima ta mutacija vpliv na dva sosednja motiva, prepoznana z Tra2beta. Z uporabo reporterskih minigenov sem določil dve najbolj aktivni ExSpe U1 molekuli, kar sem preveril s pripadajočimi celičnimi modeli preučevanih bolezni. Na primeru SMA sem izvedel transdukcijo primarnih fibroblastov in iPS celic z lentiviralnimi delci z izraženimi ExSpe U1 molekulami in ustvaril stabilno celično linijo Hek293 Flp-In klonov, ki izražajo ExSpeU1 na nivoju posameznega gena. V okvarjenih fibroblastih in iPS celicah sem ponovno pridobil pravilen SMN prepis gena in ustrezen funkcionalen protein. V

sistemu Hek293 Flp-In je kromosomska vključitev le ene kopije ExSpe U1 gena povzročila ustrezne spremembe pri popravilu spajanja eksona 7 v genu SMN2. Za primer NS sem uporabil primarne keratinocite, kjer je lentiviralno-inducirano izražanje ExSpe U1 povrnilo pravilno SPINK5 izoobliko prepisa gena ter vrednosti proteina, na način odvisen od količine odmerka.

V primerjavi z uporabo obratnosmiselnih oligonukleotidov (AON) za popravilo spajanja SMN2 sem pokazal, da ExSpe U1 molekule delujejo na povsem drugačen način. Kljub temu, da AON in ExSpe U1 molekule delno delujejo na enaka intronska zaporedja, so ExSpe U1 molekule v nasprotju z AON občutljive na povečano izražanje hnRNPA1. Z uporabo U1 snRNA vabe sem odbil endogene U1 delce na običajnih mestih spajanja in tako pokazal, da so ExSpe U1 v nasprotju z AON zmožne funkcijsko nadomestiti U1 male jedrne ribonukleinske delce (U1 snRNP).

Poizkusi EMSA so pokazali, da ExSpe U1 in U1 snRNP privabijo enake proteinske komplekse. Z uporabo spremenjenih ExSpe U1-A in U1-70K molekul sem raziskoval doprinos U1-specifičnih proteinov in posledično uravnavanje ExSpe U1-posredovanega popravljanja spajanja. Za molekulo U1-A sem določil pomembno vlogo pri povrnitvi pravilnega spajanja, med tem ko molekula U1-70K ni ključnega pomena pri posredovanju vključitve eksona.

V svojem delu sem z uporabo celičnih modelov spinalne mišične atrofije in Nethertonovega sindoma določil, da so ExSpe U1 molekule delujoče v teh sistemih ter njihovo zmogljivost samostojnega delovanja, brez sodelovanja ostalih U1 delcev. To zagotavlja njihovo pomembno vlogo kot orodja za povrnitev ustrezne vključitve eksonov v genih, kjer so mutacije spremenile genetsko opredelitev eksona in tako predstavljajo pomembno alternativo ali komplementarno strategijo v primerih neuspešne ali delno uspešne terapije z obratnosmernimi oligonukleotidi.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1 - Introduction

Chapter 2 - Results

Chapter 3 - Discussion

ABBREVIATIONS

The standard abbreviations used in this thesis follow IUPAC rules. The abbreviations are explained also in the text when they are used for the first time.

Aa	Amino acid
Bp	base pair
cDNA	complementary DNA
DMD	duchenne muscular dystrophy
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphate (A, G, C and T)
dsRBP	double strand RNA-binding protein
dsRNA	double strand RNA
DTT	dithiothreitol
EDTA	ethylenediamine tetra-acetic acid
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
GV	genomic variant
hnRNP	heterogeneous ribonuclear protein
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
Kb	kilobase
kDa	kilodalton
LEKTI	lympho-epitelial Kazal-type-related inhibitor
LV	lentiviral particles
N	nucleotide
NE	nuclear extract
NMD	nonsense-mediated decay
Nt	nucleotides
NS	netherton syndrome
PBS	phosphate buffer saline

PCR	polymerase chain reaction
PPT	polypyrimidine tract
R	purine (G or A)
RNA	ribonucleic acid
RNA PolII	RNA polymerase II
RNA PolIII	RNA polymerase III
RRM	RNA recognition motif
RS	arginine-serine rich motif
RT	room temperature
SDS	N-lauroylsarcosine sodium salt
SELEX	systematic evolution of ligands by exponential enrichment
SMA	spinal muscular atrophy
SMN	survival motor neuron
snRNA	small nuclar RNA
snRNP	small nuclear ribonucleoprotein particles
SPINK5	serine protease inhibitor Kazal-type 5
SR	arginine-serine rich protein
SRE	splicing regulatory element
Ss	splice site
TBE	tris-borate-EDTA buffer
U2AF	U2 snRNP auxiliary factor
Wt	wild type
Y	pyrimidine (T or C)
UTR	untranslated reg

1. INTRODUCTION

1.1 RNA processing and gene expression

All the processes that give rise to the production of mature functional RNA molecules are of a special importance in maintaining cellular integrity: they include fine-tuned regulated pathways of synthesis, processing and surveillance, in order to obtain protein-coding RNAs (messenger RNAs or mRNA) or non-coding RNAs. These RNAs participate in transcription (e.g. 7SK RNAs), RNA processing (small nuclear RNAs and small nucleolar RNAs), in translation (ribosomal and transfer RNAs) or gene regulation (miRNAs, siRNAs, piRNAs).

Precursor messenger RNAs (pre-mRNAs) are transcribed by RNA polII in the nucleus and undergo several post-transcriptional modifications in order to achieve correctly processed and stable mature mRNAs. The nuclear life of a mRNA passes through a complex network of dynamic interactions and co-transcriptional events that comprise 5'-capping, 3'-end processing, splicing and RNA editing that eventually lead to the formation of different mature transcripts (figure 1.1). The mRNA is then exported from the nucleus to the cytoplasm, where it serves as a template for protein synthesis (Maniatis and Reed 2002).

RNA maturation processes are perfectly synchronized by the cellular machineries for quality control, what guarantees the right conversion of the information carried by the DNA to the active RNA (Bentley 2005, Schmid and Jensen 2008).

1.1.1 5'-end capping

The capping event occurs co-transcriptionally and comprises the addition of 7-methylguanosine (m^7G) to the 5'-end of the nascent pre-mRNA. Its presence on the pre-mRNA molecules is fundamental for their stability

because it prevents the degradation promoted by 5'-3' exonucleases (Beelman and Parker 1995). Moreover, the presence of the 5'-cap mediates the export from the nucleus to the cytoplasm and promotes translation (Lewis, Izaurralde et al. 1996, Gross, Moerke et al. 2003).

1.1.2 3'-end processing

RNA processing at the 3'-end of pre-mRNA molecules is a two step reaction that comprises pre-mRNA cleavage and subsequent polyadenylation at the 3' end of the transcript. The enzyme CPSF binds specifically to the AAUAAA consensus sequence (Beaudoing, Freier et al. 2000), promoting the cleavage at a 5'-CA-3' dinucleotide located 10-35 bases downstream its binding site. After the cleavage, polyA polymerase adds a polyA tail composed of 200-250 adenine residues. During this process the intervention of PABII (polyA binding protein) that binds to the nascent polyA tail stabilizes and protects pre-mRNA from the exonucleases digestion (West, Gromak et al. 2004).

1.1.3 Splicing

The RNA splicing process is a co-transcriptional event that plays a critical role in eukaryotic cell metabolism. In fact, most of genes contain introns, that must be removed in order to join together the coding regions, the exons. This process takes place in the nucleus and it is executed using a multiprotein complex named spliceosome. Human genes typically have multiple introns and it has been estimated that at least 90% of them undergo alternative splicing. As a consequence of this event, a large number of mRNA transcripts can be generated from a single gene, increasing proteome diversity (Black 2003).

1.1.4 RNA editing

RNA editing is a post-transcriptional modification that occurs at the RNA level in order to allow the cell to increase the transcriptome variability by recoding genomic information (Benne, Van den Burg et al. 1986). It does not comprise alterations that can be ascribed to other processes (e.g. pre-mRNA

splicing or 3'-end formation). RNA editing is an event common in many organisms and consists of insertion/deletion of nucleotides or substitutions of bases by modification on RNA level without affecting the genomic information. The best-characterized events of RNA editing found in mammals convert the adenosine (A) into inosine (I) through a deamination reaction and the cytosine (C) into uracil (U) (Sommer, Kohler et al. 1991). The strong impact of RNA editing could be illustrated by the apolipoprotein B mRNA example. In the intestine this mRNA undergoes a tissue specific C>U editing that converts a CAA codon into a UAA stop codon, leading to the formation of a shorter intestinal ApoB48 lipoproduct in comparison to ApoB100 protein synthetized in the liver (Ashkenas 1997).

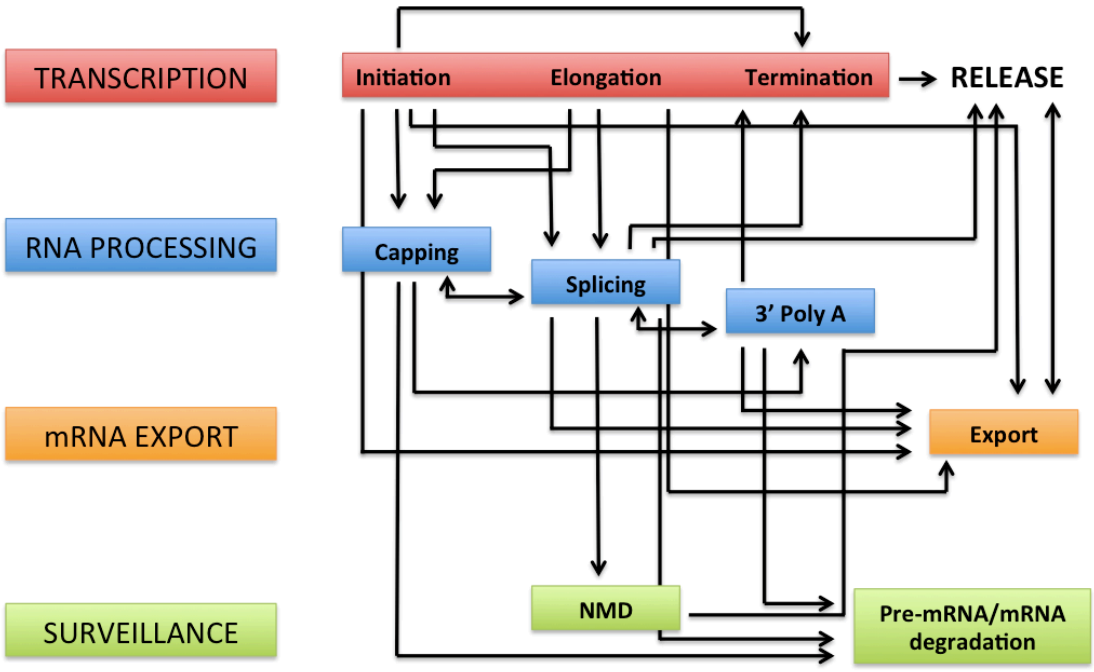


Figure 1.1. A complex network of coupled interactions in gene expression.

The major events of gene expression are indicated on the left and each transcriptional step is shown along the top in the red box. "Release" indicates release of the mature mRNA from the site of transcription. The splicing step makes part of the RNA processing but is also linked with other step of gene expression such as transcription, nonsense mediated

decay (NMD), mRNA export and degradation. Black arrows indicate physical or functional interactions between components of gene expression machine.
[Figure adapted from (Maniatis and Reed 2002)].

1.2 Pre-mRNA splicing

Eukaryotic pre-mRNAs are transcribed in the nucleus by RNA polII. Considering the fact that the translational machinery of the cell is not able to recognize the coding part of the transcript from the intronic regions, the excision of the non-coding sequence has to occur before the translation initiates. The splicing reaction is a rapid and precise mechanism through which the introns are accurately removed in order to put the exons in the correct protein-reading frame (Roca, Krainer et al. 2013) .
The definition of the exons and the proper realization of the splicing reaction are mediated by a macromolecular machinery named spliceosome, which is composed of small nuclear RNAs (snRNAs) and tens of associated proteins (Maniatis and Reed 2002, Wahl, Will et al. 2009)

1.2.1 The chemistry of a splicing reaction.

From a chemical point of view the splicing cut-and-paste catalytic reaction entails two S_N2-type transesterification steps that involve functional groups from three different regions in the pre-mRNA (Lamond 1993). The initial signals necessary to start the splicing reactions are very conserved motifs at the intron-exon borders and their proximity: the GU dinucleotide at the 5'-splice site (5'ss) or donor splice site, the AG dinucleotide at the 3'-splice site (3'ss) or acceptor splice site, a polypyrimidine tract and the A nucleotide at the branch point, both located upstream of the 3'ss (figure 1.2, A) (Moore and Sharp 1993).
In the first transesterification reaction the phosphate at the GU 5'ss is attacked attached by the 2'-hydroxil group of the adenine residues at the branch point site (BPS), leading to the cleavage of the 5' exon from the intron

and the formation of a lariat as intermediate (figure 1.2, B) (Ruskin, Krainer et al. 1984).

The second transesterification step consists of the attack of the 3'-hydroxil group of the detached exon on the phosphate at the AG 3'ss end of the intron (figure 1.2, C). This step results in the ligation of two exons through a phosphodiester bond and the released intron in a lariat form (figure 1.2, D). The lariat is then debranched to obtain a linear excised intron, which is usually rapidly degraded (Konarska, Grabowski et al. 1985, Moore and Sharp 1993).

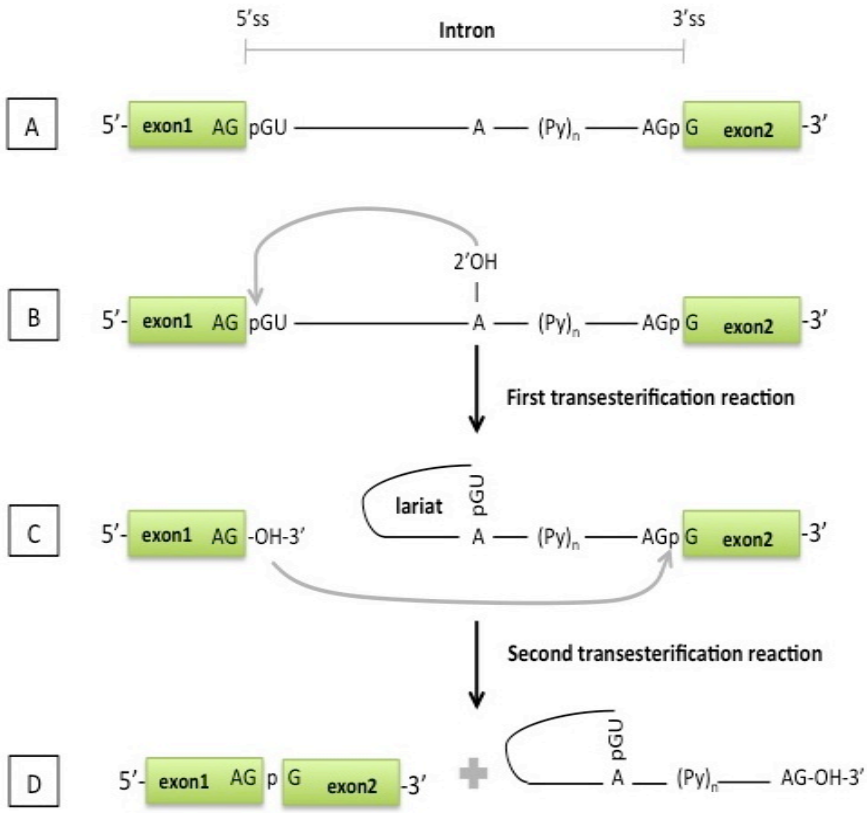


Figure 1.2. Splicing occurs in two transesterification reactions.

A. Schematic representation of the pre-mRNA with essential signals for splicing reaction. Exons are green boxes, introns are black lines, A is the adenine residue of the branch point, (Py)n is the polypyrimidine tract.

Positions of the GU and AG dinucleotides at the 5' and 3'-splice sites are indicated.

B. In the first transesterification step, an adenine residue of the branch site attacks the phospodiester bond at the 5'ss of exon 1 generating two splicing intermediates: free exon 1 and lariat-exon2.

C. During the second transesterification reaction, the 3'-OH group of the detached exon attacks the phosphate at the AG 3'-end of the intron.

D. Final products of splicing reaction are two exons joined together through phosphodiester bond and the intron in a lariat form.

1.2.2 The spliceosome.

Splicing of pre-mRNAs is carried out by the spliceosome, a macromolecular machinery that assembles onto nascent mRNA transcripts, catalyses splicing reactions and gets released from the mRNA upon splicing completion. This highly dynamic complex is composed of five uridine-rich (U) small nuclear ribonucleoparticles (snRNPs), which constitute its main building blocks. Each U1, U2, U4/U6 and U5 snRNPs consists of one snRNA (or two in case of U4/U6), a common set of seven Smith (Sm) proteins and a different number of complex-specific proteins (Will and Luhrmann 2001). Human spliceosomes also interact with a multitude (from 150 to 300) of non-snRNP proteins, such us heterogeneous nuclear RNPs (hnRNPs) and serine-arginine rich (SR) proteins (Wahl, Will et al. 2009). Some metazoan species and plants contain a second, minor spliceosome responsible for the excision of a rare class of introns, that is composed of structurally distinct but functionally analogous U11/U12 and U4atac/U6atac snRNPs, with the U5 snRNP which is common for the two spliceosomal machineries (Patel and Bellini 2008).

Spliceosome assembly occurs in a stepwise manner, leading to assembly/disassembly of the different snRNP particles and non-snRNP splicing factors on the pre-mRNA. RNA-protein, protein-protein and RNA-RNA base-pairing interactions and several conformational changes are carried out during the formation of the mature mRNA (Hong, Bennett et al.

1997, Das and Reed 1999, Kent, Ritchie et al. 2005, Tardiff and Rosbash 2006).

In mammals, different spliceosome assembly intermediates have been identified: the E, A and B complexes contain unspliced RNAs, while the C complex comprise the catalytic product of the first transesterification reaction (exon 1 and lariat exon 2). The products of the second transesterification reaction are contained in two different complexes: the spliced exon is within complex D and the lariat intron in the "i" complex (Wahl, Will et al. 2009).

The assembly of the spliceosome begins with the ATP-independent binding of the U1 snRNP through base pairing interactions of the 5'-tail of the U1 snRNA and the 5'ss of the intron (figure 1.3). Although this first step is energy independent, subsequent spliceosomal rearrangements need NTPs hydrolysis (Will, Rumpler et al. 1996).

The "early" complex (E), named also commitment complex, is characterized by the recognition of the splice sites of the exon. Based on *in vitro* studies, the at the 5'ss, the U1 snRNP particle interacts with the donor sequence in the pre-mRNA through RNA-RNA base-pairing and protein-RNA interaction via U1-70K and U1-C proteins. The 3'ss is individuated due to the interaction of the dimeric U2 Auxiliary Factor (U2AF) in which the larger U2AF65 subunit recognizes the polypirymidine tract and the smaller U2AF35 subunit the AG dinucleotide at the intron-exon border. The branch point nucleotide A is usually located 20-40 nucleotides upstream the acceptor splice site and it is specifically bound by the branch point binding protein (BBP/SF1) (Valcarcel, Gaur et al. 1996, Wu, Romfo et al. 1999, Schellenberg, Ritchie et al. 2008).

The second step in the spliceosomal assembly is the formation of the A complex (figure 1.3), in which BBP/SF1 is replaced by the binding of the U2 snRNP at the branch point. This splicing factors exchange requires hydrolysis of ATP and it is mediated by the U2 snRNA basepairing with the branch point sequence (BPS) and by the interactions of two U2 snRNP subunits SF3a and SF3b with the pre-mRNA (Gozani, Feld et al. 1996).

The transition from the complex A to the complex B occurs with the ATP-dependent addition of the U4/U6/U5 snRNPs, in which U4/U6 are basepaired. Formally, the complex B contains all the snRNPs necessary to promote the splicing reaction, but lacks a real catalytic centre. The spliceosome has to be activated and this operation occurs through rearrangements based on an intricate network of RNA-RNA interactions: the U1 snRNP is displaced from the 5'-splice site of the pre-mRNA by the U6 snRNP, causing the disruption of the U4/U6 basepairing. These conformational changes provide the structural basis to juxtapose the branch site and the donor splice site, promoting the formation of the activated B complex (B*) (Reed 2000, Boehringer, Makarov et al. 2004, Turner, Norman et al. 2004).

The first step of the two splicing transesterification reactions occurs in the B* complex and it generates two intermediates: the free 5' exon and the lariat-3'-exon (O'Keefe, Norman et al. 1996). This event is followed by the formation of the C complex in which the second transesterification reaction takes place. The U5 snRNP, along with the U2 and U6, has a key role in this step. The U5 snRNA, along with the U2 and U6 snRNAs, is involved in aligning the exons for the second transesterification step through a highly conserved stem loop (O'Keefe, Norman et al, 1996). In particular, a U5-associated protein named Prp8 is involved in the stabilisation of these interactions (Umen and Guthrie 1995).

It still remains unsolved which of the spliceosome components finally catalyses the splicing reaction, although the possibility that the mechanism is RNA-based has more evidences. *In vitro* experiments demonstrate that U2 and U6 snRNA are able to basepair even in absence of proteins. In the presence of a branch point they are able to promote the nucleofilic attack of the adenosine at the branch point (Valadkhan and Manley 2001, Valadkhan, Mohammadi et al. 2009).

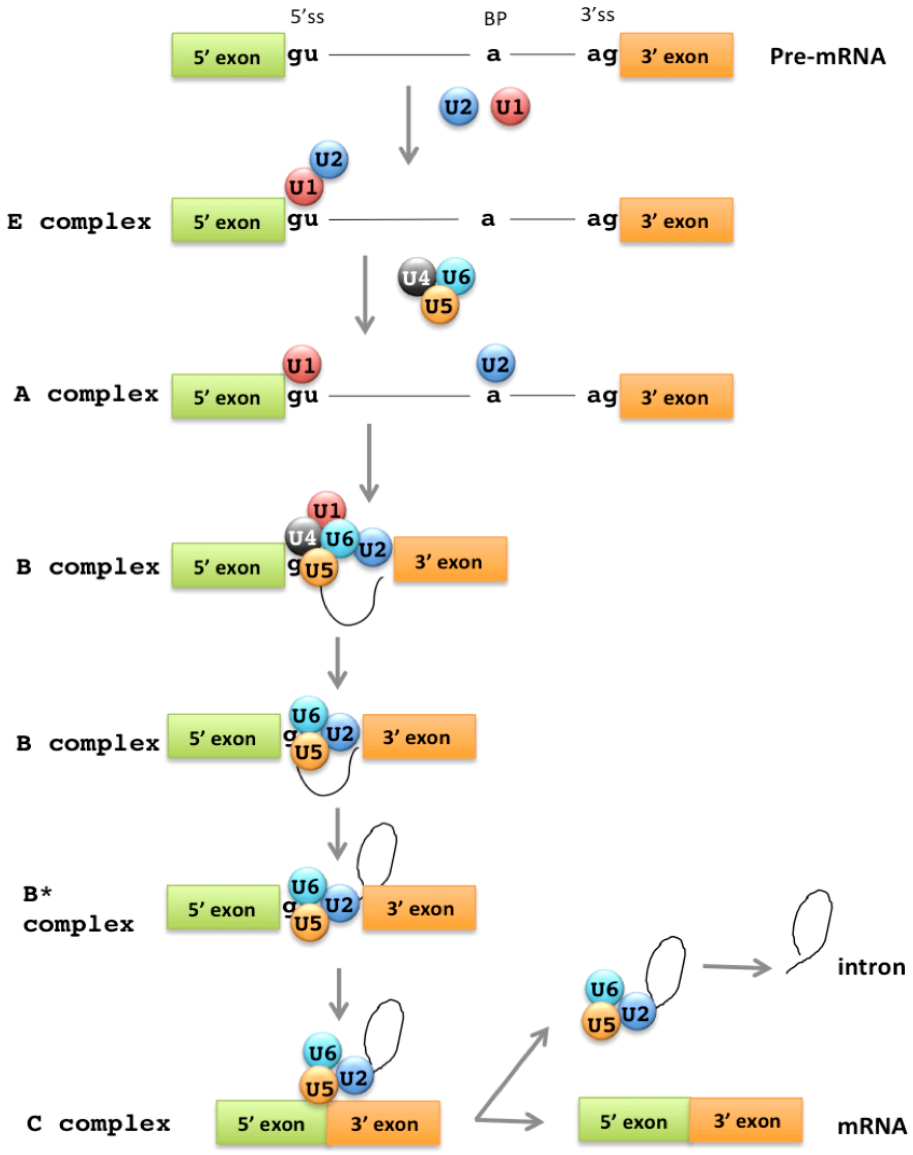


Figure 1.3. Spliceosome assembly step by step.

The spliceosome assembles onto the pre-mRNA in a stepwise manner. Exons are noted by *boxes*, and introns by *lines*. Consensus nucleotides are indicated above the line. The complex E contains U1 snRNP bound to the 5' splice site, SF1 bound to the branch point, and U2AF65 and U2AF35 bound to the pyrimidine tract and 3' splice site AG, respectively (not shown). In the complex A, SF1 is replaced by U2 snRNP at the branch point. The U4/U6/U5 tri-snRNP then enters to form the B complex. Finally, a rearrangement occurs to form the catalytically active C complex, in which U2 and U6 interact, and U6 replaces U1 at the 5' splice site, promoting the two splicing reactions.

1.2.3 U1 small nuclear RNA (U1 snRNA).

The human U1 snRNA gene (*RNU1*) is present in clustered repeated units of 45 kb located in the short arm of chromosome 1 (1p36). In addition, many different U1 snRNA pseudogenes have been identified in the human genome that could generate functional variants of U1 snRNAs (Lund and Dahlberg 1984, Bernstein, Manser et al. 1985).

RNU1 is a TATA-less promoter gene, which minimally contains an enhancer-like distal sequence element (DSE) and a fundamental proximal sequence element (PSE) that is recognized by the snRNA gene-specific factor PTF (PSE-binding transcription factor)/ SNAPc (snRNA activator protein complex) (figure 1.4) (Sadowski, Henry et al. 1993)

The *RNU1* transcript is not spliced and the 3'-end is not polyadenylated: what possibly prevents the association with the translation machinery. The 3'-end formation requires the presence of a conserved element, the 3'-box, located 9-19 nucleotides downstream the snRNA encoding region (Hernandez 2001, Egloff, O'Reilly et al. 2008)

RNU1 is polymerase II (PolII) dependent and its transcription needs TBP (TATA-box binding protein), TFIIB (transcription factor IIB), TFIIA, TFIIE and TFIIF proteins (Kuhlman, Cho et al. 1999).

The 3'-end formations occurs in a stepwise manner: the first event is the recognition of the *cis*-acting 3'-box, which is a 13-16 nucleotide long element that directs the production of a 3'-extended pre-snRNA that is subsequently processed, leading to the formation of the mature 3'-end after the transport to the cytoplasm (Kuhlman, Cho et al. 1999, Kiss 2004). Recently a large complex named *Integrator*, composed of Int11 and Int9, has been shown to play a role in pre-snRNA 3'-end formation (Baillat, Hakimi et al. 2005). Thus, findings indicate that the 3'-box is an RNA processing element analogous to the polyadenylation signal, commonly found in protein coding genes (Uguen and Murphy 2003, Egloff, O'Reilly et al. 2008).

It is widely accepted that the phosphorylation of the C-terminal domain (CTD) of the large subunit of the PolII is necessary for the 3'-box-dependent RNA 3'-end formation *in vivo*, what indicates that 3'end processing occurs co-

transcriptionally (Medlin, Uguen et al. 2003). In particular it has been demonstrated that the CTD phosphorylation is fundamental for recruiting the *Integrator* complex, whose binding is crucial for a correct 3'-end processing (Egloff, O'Reilly et al. 2007).

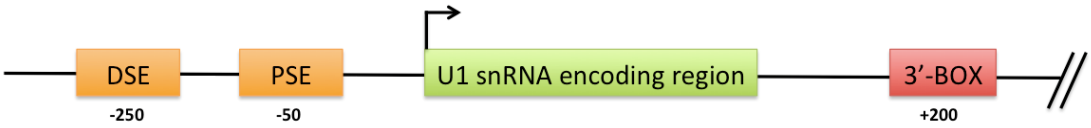


Figure 1.4. Genomic organization of human U1 snRNA.

Schematic representation of the U1 snRNA gene structure. Orange boxes are DSE and PSE *cis*-acting promoter elements, while the red ones is the 3'-end box. Their position, relative to transcription start site (indicated by the arrow above the green box) is indicated below each box.

1.2.3.1 U1 small nuclear ribonucleoparticle (U1 snRNP)

The *trans*-acting factors involved in the recognition of specific sequences on the pre-mRNA molecules are crucial players in splicing events. In particular, uridine-rich small nuclear ribonucleoparticles (U-snRNPs) are the constitutive elements of the spliceosome machinery and therefore responsible for the identification of the canonical consensus signals. Among these, the U1 snRNP represent the first important element that gives rise to spliceosome assembly after the recognition of the 5'ss in the pre-mRNA (Sperling, Azubel et al. 2008).

In humans the U1 snRNP consists of a 164-nucleotides long RNA molecule that recruits ten different proteins. Seven of these proteins are common to all the U snRNPs and are known as Sm proteins (Sm B, D1, D2, D3, E, F and G), while the other three proteins (U1-70K, U1-A and U1-C) are U1 specific (figure 1.5) (Stark, Dube et al. 2001).

The secondary structure of U1 snRNA is characterized by four stem-loops (I, II, III, IV) and two single stranded regions. In particular, the single stranded

5'-proturuding tail of the U1 snRNA basepairs with 9 bases at the 5'ss consensus motif in the pre-mRNA molecule (Mount, Pettersson et al. 1983). The other single stranded region located between stem-loops II and IV is the conserved Sm-binding site (AAUUUGUGG), which is bound in the cytoplasm by the heteroheptameric ring composed of Sm proteins (Raker, Hartmuth et al. 1999).

U1 specific proteins U1-70K and U1-A interact with stem-loop I and II respectively (Surowy, van Santen et al. 1989, Yuo and Weiner 1989, Hamm, Dathan et al. 1990), whereas U1-C is thought to enter in the U1-complex through a protein-protein interaction (Nelissen, Will et al. 1994).

The role of the U1 snRNP in the identification of the donor splice site has been discovered and proven several years ago (Mount, Pettersson et al. 1983, Rinke, Appel et al. 1984). Nevertheless, the recognition of the 5'ss can occur, at least for some exons, also in absence of the U1 particle at the donor site, although the splicing process is less efficient in these cases (Du and Rosbash 2002). The U1-C has been reported to contribute to the 5'ss recognition and it stabilizes the basepairing between the 5'-tail of the U1 snRNA and the donor site (Du and Rosbash 2002), promoting the formation of the E complex (Will, Rumpler et al. 1996).

In vitro experiments showed that U1-A is not required for splicing, as the deletion of stem loop II or U1A depletion from nuclear extracts does not impact U1 snRNP splicing efficacy (Heinrichs, Bach et al. 1990, Will, Rumpler et al. 1996).

Recently, new roles for the U1 snRNP particles have been identified. In addition to its role in splicing, this particle has been shown to be active in the protection of pre-mRNA from premature polyadenylation at cryptic polyadenylation signals (Berg, Singh et al. 2012).

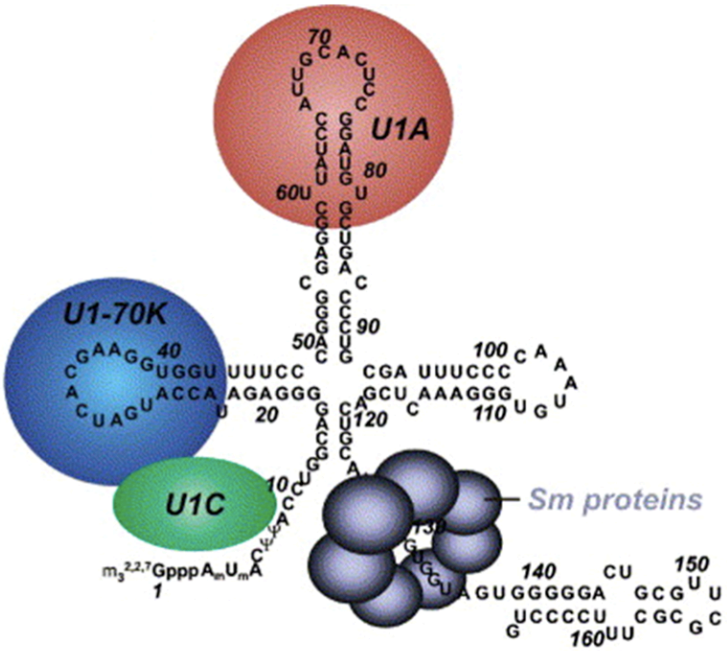


Figure 1.5. Schematic representation of the U1 small nuclear RNP.

The secondary structure of the U1 small nuclear RNA shows four stem loops and two single stranded regions. The U1-specific proteins U1-70K (blue circle) and U1-A (red circle) bind stem loop I and II respectively. U1-C is part of the complex through a protein-protein interaction with U1-70K. The seven Sm proteins (dark blue circles) bind the single stranded region between loops III and IV. The 5'-RNA-tail tri-methylated carries the consensus motif for the recognition of the donor splice site on the pre-mRNA molecule.

1.2.4 Canonical cis-acting elements in splice-sites selection.

The recognition of an exon occurs through the recognition of the splice sites (ss) that are small consensus sequences located at the exon/intron boundaries on the pre-mRNA molecules. The canonical consensus sequences are the 5'ss, the 3'ss, the polypyrimidine tract and the branch point sequence (BPS) (figure 1.6) (Barash, Calarco et al. 2010).

The 5'ss defines the exon/intron border at the 5' end of the intron. The first two dinucleotides of the intron (GU) are highly conserved, although they are not sufficient for the recognition of the exon and require the sequence surrounding the donor splice site, that is quite degenerated (Aebi, Hornig et al. 1987).

The other intron/exon border is the 3'ss region, located at the 3'-end of the intron and it is composed by three conserved elements: the A nucleotide at the branch point, the intronic polypirimydine tract and the AG dinucleotide that defines the end of the intron (Shapiro and Senapathy 1987).

The consensus sequences of the canonical splicing signals not always perfectly match the consensus sequence carried by the elements of the splicesome deputated to recognize them. Thus, accordingly to their degree of similarity to the general consensus sequences, the splice sites are defined as strong or weak. Several bioinformatics tools have been developed during the last 10 years in order to predict positions and strength of splice sites (Cartegni, Wang et al. 2003, Fairbrother, Yeo et al. 2004).

Given that canonical *cis*-acting elements are not perfectly conserved, it is important to underline that lots of intronic pre-mRNA sequences match the consensus sequence of 5' and 3'ss, sometimes even with higher strength than the real ones. These intronic splice sites usually define pseudo-exons that are normally not included in the mature transcripts (Sun and Chasin 2000). However, the fact that the spliceosome is able to recognize the correct splicing signals along a multitude of pseudo-splice sites means that the canonical *cis*-acting signals are not sufficient *per se* to promote the correct identification of an exon. The precise recognition of the intron/exon junction *in vivo* needs additional *cis*-acting regulatory sequences located both in

introns and exons that help the splicing machinery to execute the reaction (see paragraph 1.2.5).

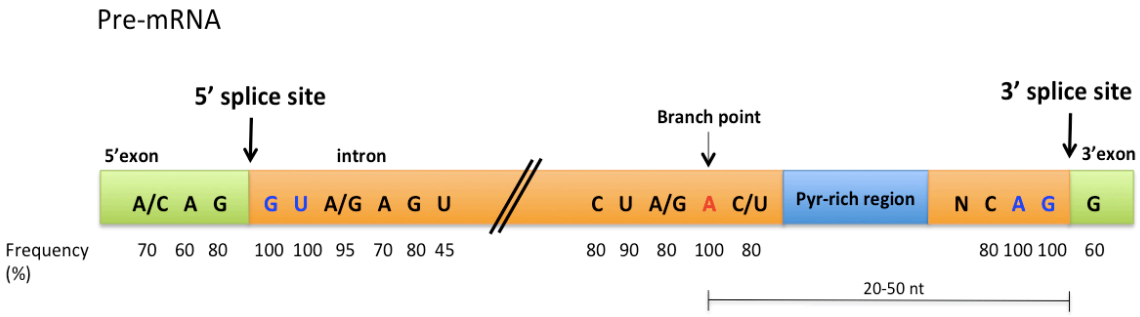


Figure 1.6. Schematic representation of exon-intron boundaries.

Two exons are represented by the green boxes. Between them are reported the consensus sequences presented within an intron. The arrows show the position of the 5' and 3'ss and the branch point. The polypyrimidine tract is indicated as a blue box.

The frequency of each nucleotide at a specific position is indicated in the bottom of the panel. The GU at the 5'ss, the AG at the 3'ss and the A at the branch point site are the only 100% conserved bases.

1.2.4.1 The 5'-splice site.

In the higher eukaryotes the motif surrounding the 5'-splice site (5'ss) is composed of 9 nucleotides MAG/GURAGU (M indicates A or C, R indicates purins and the slash the exon/intron boundary), spanning from exonic position -3 to intronic position +6. The intronic GU dinucleotide in position +1 and +2 is 100% conserved and mutations involving one of these bases completely disrupt the splice site, abolishing the splicing of the exon (figure 1.6) (Langford, Klinz et al. 1984). However, it has been reported in very few cases a C>U substitution at position +2 (Senapathy, Shapiro et al. 1990).

The recognition of the 5'ss is driven by a nearly perfect basepairing between the sequence surrounding the donor site and the 5'-tail of the U1 snRNA. This is the first event that leads to the assembly of the spliceosomal machinery upon the intron (Horowitz and Krainer 1994)-.

The contribution of the RNA:RNA basepairing of the 5'ss consensus sequence and the U1 snRNA has been shown to be crucial for the selection of the donor splice site. In fact, mutations that improve the recognition of weak splice sites has been described to promote the constitutive inclusion of alternatively spliced exons (Muro, Iaconcig et al. 1998).

1.2.4.2 The 3'-splice site

The three conserved intronic elements that constitute the 3'-splice site (3'ss) and that define the 3' border of the intron are the branch point, the polypyrimidine tract and the AG dinucleotide at the border of the intron (figure 1.6) (Reed 1989).

The branch point sequence (BPS): in human the sequence that defines the branch point is highly degenerated, with the exception of the A nucleotide that is 100% conserved. The general consensus sequence is the YNYURAC motif (R as purine, Y as pyrimidine) and most of them are located 20-50 nucleotides upstream the 3'ss (Reed and Maniatis 1988). Nevertheless, in the case of the rat α-tropomyosin gene intron 2 the BPS is located 172 nucleotides upstream the 3'ss and is essential for the regulation of alternative splicing (Smith and Nadal-Ginard 1989).

The splicing factor SF1 recognizes the branch point, discriminating between sequences containing or not the totally conserved adenosine (Berglund, Chua et al. 1997). In a second step, during the formation of the spliceosomal A complex, the U2 snRNA matches the BPS through an RNA:RNA interaction.

The polypyrimidine tract: the polypyrimidine tract is a stretch of pyrimidines (average length 8 bases) located between the branch point sequence and the AG dinucleotide at the intron/exon border. This intronic sequence is recognized by the U2 Auxiliary Factor (U2AF), through its 65kDa subunit, during the formation of the E complex (Kielkopf, Lucke et al. 2004). The length of the polypyrimidine tract could influence the proper recognition of the 3'ss. It has been shown that progressive shortening of the poly-uridine sequence affects the recognition of the branch point site, thus affecting spliceosome assembly and lariat formation. On the contrary, elongating the

length of the sequence can improve splicing efficiency (Roscigno, Weiner et al. 1993).

The AG dinucleotide: the terminal AG is located at the terminal part of the intron, at the exon/intron junction, downstream the polypyrimidine tract. The general consensus sequence is a YAG/G motif (Y as a pyrimidine, the diagonal line indicates the exon-intron border) in which only the AG are purely conserved. The 3'ss does not require a specific interaction with snRNPs, but it is recognized by U2AF through its 35 kDa subunit during the formation of the E complex (Wu, Romfo et al. 1999).

1.2.5 Auxiliary cis-acting sequences

Since the sequences surrounding the canonical *cis*-acting elements are largely degenerated and in introns there are several pseudo-signals that are never used, it is clear that canonical signals are not sufficient *per se* to correctly define the intron/exon junctions (Mardon, Sebastio et al. 1987). There are other additional *cis*-acting elements in pre-mRNAs that, in combination with the conventional ones, are able to precisely identify the exons, guiding the spliceosome to the proper splice sites. In fact, these sequences, located both in introns and in exons, act stimulatingly or inhibitory on the use of specific splice sites (Cartegni, Chew et al. 2002).

Accordingly to their position and their function, these elements are named exonic splicing silencer or enhancer (ESS and ESE, respectively) and intronic splicing silencer or enhancer (ISS and ISE, respectively). They have a random distribution on the pre-mRNA molecule, although studies revealed a higher density of enhancers in real exons than in pseudoexons and introns, and *viceversa* for sequences that act as silencers (figure 1.7) (Zhang and Chasin 2004). Intronic and exonic splicing auxiliary elements show highly degenerate sequences, making their identification difficult, even with the use of bioinformatics tools (Cartegni, Chew et al. 2002, Pagani and Baralle 2004).

In addition, it has been reported an overlapping functions of enhancer and silencer for some regulatory sequences, as in the case of composite exonic regulatory elements of splicing (CERES) identified in the cystic fibrosis

transmembrane conductance regulator gene (*CFTR*) exons 9 and 12 (Pagani, Stuani et al. 2003).

Moreover, splicing regulatory elements can operate in an additive way, as the presence of more than one copy may influence splicing through the improvement of the affinity of the associated splicing factor. Different splicing factors may cooperate to promote the recognition of an exon (Wang, Rolish et al. 2004, Dominguez and Allain 2006). This fine-tuned balance of splicing factors is of extreme importance for the alternative splicing mechanism that, in turn, contributes in a significant way to gene expression regulation (Black 2003).

1.2.5.1 Splicing enhancers

Splicing enhancers are regulatory sequences located both in introns and in exons that promote exon inclusion. Exonic splicing enhancers (ESEs) play a key role in alternative splicing regulation (Black 2003), but are implicated also in constitutive splicing events (Schaal and Maniatis 1999). Extensive studies of ESEs with bioinformatics tools demonstrated that most of them are recognized by members of the serine-arginine-rich (SR) family proteins that act on the pre-mRNA molecule, stimulating the inclusion of the exon in the mature transcript through recruiting splicing factors or antagonizing the effect of neighbouring silencer motifs (Graveley 2000, Cartegni, Chew et al. 2002)

Studies based on functional systematic evolution of ligands by exponential enrichment (SELEX) allowed the identification of purine-rich and non-purine-rich ESEs that were involved in the recruitment of SR proteins, whose binding consensus sequence is usually very short (6-8 nucleotides), degenerated and partially overlapping (Tian and Kole 1995, Schaal and Maniatis 1999, Boukis and Bruzik 2001, Cartegni, Wang et al. 2003). The frequencies of each nucleotide in specific positions in the consensus sequences were calculated and then used to create matrices in order to predict the location of putative SR protein specific exonic splicing enhancers, creating an online tool named ESEfinder (Cartegni, Wang et al. 2003).

Intronic splicing enhancers (ISEs) are reported and characterized in literature, but few large-scale studies have been carried out in comparison to ESEs (Zheng 2004). ISEs were described principally during the investigation of point mutations related to pathologies. Some of these mutations affect intronic regions 20-40 nt downstream the donor splice site of an exon, causing its exclusion from the mature transcript (McCarthy and Phillips 1998, Lew, Fei et al. 2004). However, ISEs are fundamental also for the constitutive recognition of an exon and may be involved in crucial alternative splicing in specific tissues or developmental stages (Venables 2007).
A well-described ISE is the G-run motif, which is usually present in clusters and promote the recognition of the nearby splice sites. Other important sequences are the UGCAUG hexanucleotides that are recurrent in regions downstream of exons of genes involved in neuron-specific pathways and act as enhancers binding the brain-specific splicing factor Fox-1 (Nakahata and Kawamoto 2005).

1.2.5.2 Splicing silencers

Splicing silencers are regulatory sequences located both in introns and exons that are responsible for splicing inhibition, and are called intronic splicing silencers (ISSs) or exonic splicing enhancers (ESSs) respectively. Although the motifs of silencers are usually rich in purines or pyrimidines (Fairbrother and Chasin 2000, Pozzoli and Sironi 2005), they could be bound by several splicing factors with generally inhibitory functions that belong to the heterogeneous nuclear ribonucleoprotein (hnRNP) family (Cartegni, Chew et al. 2002). Splicing silencers modulate splicing in two different ways: antagonizing the effect of adjacent splicing enhancers or recruiting splicing inhibitory factors that interfere with the spliceosomal machinery, leading to the skipping of the exon from the mRNA (Cartegni, Chew et al. 2002, Matlin, Clark et al. 2005).
ESSs were identified during studies on HIV-1 *tat* gene alternative splicing. In particular, the first ESS was discovered in tat exon 2, where it inhibits recognition of the upstream 3'ss (Amendt, Hesslein et al. 1994). ESSs are

involved not only in alternative splicing mechanisms, but also in constitutive splicing events (Szeszel-Fedorowicz, Talukdar et al. 2006).

ISSs usually serve as binding sites for Polypyrimidine tract binding protein (PTB) (Zheng 2004). PTB recognizes pyrimidine rich sequences and antagonizes the binding of U2AF upstream of an exon or creating a region of silencing across an exon that becomes downregulated (Sauliere, Sureau et al. 2006). In contrast, an ISS has been reported that is unexpectedly recognized by SR proteins, which exert an inhibitory activity(Pagani, Buratti et al. 2000). This particular situation indicates that, in general, splicing factors can behave both as silencers or enhancers, depending on the context and position of the regulatory sequences in a specific pre-mRNA (Pagani, Buratti et al. 2000).

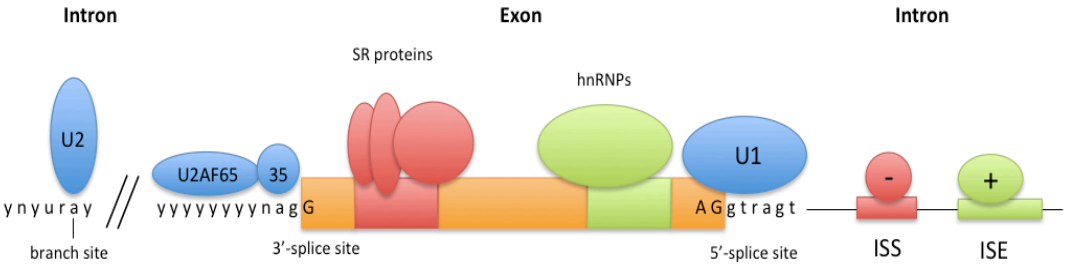


Figure 1.7. Regulatory elements in pre-mRNA splicing.

The scheme represents a possible distribution of canonical and auxiliary *cis*-acting elements in a pre-mRNA transcript. The canonical splicing signals are relatively degenerated and short. Only the GU dinucleotide at the 5'ss, the A at the branch point and the AG dinucleotide at the 3'ss are 100% conserved.

In blue are depicted the components of the spliceosome machinery: the U1 snRNP binds the 5'ss, the U2 snRNP binds at the branch site and the U2AF through its two components (U2AF65 and U2AF35 subunits) interacts with the polypyrimidine tract and the AG at the 3'ss, respectively.

Additional *cis*-acting elements, located both in introns and in exons can act as splicing silencer or enhancers (ISS, ISE, ESS, ESE), allowing the

proper identification of the exon, distinguishing it from the many cryptic sites that have identical sequences.

Trans-acting splicing factors interact with the regulatory sequences and, accordingly to their general functions, are divided into SR proteins family with enhancer activity and hnRNP protein family with inhibitory activity.

1.2.6 RNA secondary structure

RNA molecules have the natural tendency to fold in highly stable secondary and tertiary structures that were shown to play an important role in exon definition for particular transcripts (Buratti and Baralle 2004). The RNA secondary structure can influence the splicing process by interfering with the activity of canonical or auxiliary *cis*-acting elements. In the first case, structural features of the pre-mRNA molecule can hide or expose the 5'ss, the 3'ss or both of them, altering the accessibility of the spliceosome components to these canonical splicing signals. For instance, in the case of human *tau* exon 10 a stem-loop structure located at the 5'ss modulates the splicing of this exon. Mutations that affect the secondary structure of this region destabilize the stem-loop, what results in an increase of exon 10 inclusion (Grover, Houlden et al. 1999).

Moreover, the RNA secondary structure can also have a relevant impact on the functionality of splicing regulatory elements. In the case of fibronectin the *EDA* exon inclusion is modulated by a particular secondary structure that allows the exposition of an ESE element recognized by ASF/SF2 that has been shown to be crucial for stem-loop stabilization and, in turn, for the binding of SR proteins in the nearby region (Muro, Caputi et al. 1999).

In silico studies suggest that regulatory sequences are enriched in single stranded regions of the pre-mRNA molecule, which makes them perfect targets for splicing factors (Hiller, Zhang et al. 2007). This implies that the secondary structure of a specific RNA may play a key role in the determination of the functionality of *cis*-acting elements.

1.2.7 Trans-acting factors

Proteins bound to RNA are important for all aspects of pre-mRNA splicing. They are called *trans*-acting factors and can be divided into small nuclear ribonucleoproteins (snRNPs, see paragraph 1.2.3) and non-snRNPs *trans*-acting factors. Apart from their function in the splicing process, these proteins show common features such as the presence of RNA recognition motifs (RRMs) and/or protein binding domains. The SR and the hnRNP proteins families are the two major classes of RNA binding proteins that have different regulatory functions in pre-mRNA splicing process.

1.2.7.1 Heterogeneous ribonuclear proteins (hnRNPs).

The heterogeneous ribonuclear proteins (hnRNPs) family is a class of various RNA-binding proteins that interact with nascent pre-mRNAs until they are completely processed (Dreyfuss, Matunis et al. 1993). These factors predominantly localize in the nucleus, although some of them shuttle from the nucleus to the cytoplasm facilitating the nuclear export (Pinol-Roma and Dreyfuss 1992, Izaurralde and Mattaj 1995). More than 20 members of the hnRNP families have been identified and their diversity is enlarged by alternative splicing events that increase the number of different isoforms as well as post-translational modifications, such as phosphorylation, arginine methylation and SUMOylation (Dreyfuss, Kim et al. 2002, Martinez-Contreras, Cloutier et al. 2007).

The hnRNPs are ubiquitinously expressed in all the tissues, but the relative amounts of different proteins have a variability among cell types and show stage-specific expression pattern (Kamma, Portman et al. 1995). The common features of their structures are one or more RNA binding motifs (RRMs) and auxiliary domains, necessary for protein-protein interactions (Gly-rich, KH) (Honore, Rasmussen et al. 1995). The hnRNPs frequently mediate splicing repression by binding to ESSs or by steric interference with other splicing factors (Cartegni, Chew et al. 2002). However, this class of proteins may associate with splicing enhancers, depending on the position of the regulatory element (Caputi and Zahler 2002).

A classical example is hnRNP A1 that shows the highest binding affinity for the UAGGGY silencer motif, mediating exon repression (Kim, Merrill et al. 1997). Nevertheless, this protein is also involved in mRNA stability (Hamilton, Burns et al. 1997), telomerase maintenance (Zhang, 2006), cell proliferation (Iervolino, Santilli et al. 2002) and miRNA processing (Guil and Caceres 2007). It is known that it interacts with itself and/or other hnRNPs as well as with U2 and U4 snRNPs, what results in different manners of splicing inhibition (Mayeda and Krainer 1992, Caceres, Stamm et al. 1994). It can directly competes with positive regulatory splicing factors for the binding to the 5'ss, prevent the access to the 3'ss by steric hindrance binding to a nearby ESS or it can cause "looping out" of alternative exons cooperating with other hnRNP A1 protein complexes along the pre-mRNA molecule (Mayeda and Krainer 1992).

1.2.7.2 Serine-arginine rich proteins (SRs).

The SR proteins are a family of highly conserved nuclear phosphoproteins that have multiple roles, not only in the RNA splicing process, but also in the cell metabolism (Huang and Steitz 2005, Dominguez and Allain 2006). SRs have a modular structure that consists of one or two RNA recognition motifs (RRMs) and one characteristic domain rich in arginine and serine residues known as RS domain located at the C-terminus (Birney, Kumar et al. 1993). The RRMs are fundamental for the sequence-specific interaction between the SRs and the pre-mRNA, while the RS domain mediates protein-protein interactions with components of the splicing apparatus. In addition, the RS domain can also interact directly with the branch point site during the splicing process, promoting the formation of the spliceosome (Boucher, Ouzounis et al. 2001, Shen and Green 2004).

SR proteins are required for both alternative and constitutive splicing events (Sanford, Ellis et al. 2005). Their consensus binding sequence on the pre-mRNA molecule has shown to be highly degenerated, as demonstrated by SELEX studies performed on individual SR proteins (Liu, Zhang et al. 1998). The SRs can bind a target sequence of some other family members thus

contributing to functional redundancy and suggesting an overlapping between the functions exerted by different family members (Longman, Johnstone et al. 2000).

Two non-exclusive models have been proposed to explain the mechanism of action of SR proteins in pre-mRNA splicing process. The first is based on the capacity of this splicing factors to bind ESE and, through their RS domain, to recruit and stabilize the U1 snRNP and the U2AF binding at the 5'ss and 3'ss, respectively (Sanford, Ellis et al. 2005). The second model proposed that SR proteins, interacting with ESEs, can antagonize the effect of nearby silencer elements (Kan and Green 1999). The RS domain of SR proteins can be phosphorylated at different positions (Stamm 2008) and this post-transcriptional modification largely affects the splicing events in the nucleus by changing the RNA binding activity and their subnuclear localization (Misteli and Spector 1997). They are preferentially localized in nucleus, but are continuously shuttled from the nucleus to the cytoplasm, according to their phosphorylation status,(Caceres, Screaton et al. 1998).

1.2.8 Alternative splicing

The proteome diversity is guaranteed by several cellular events that allow the generation of different polypeptides from a single gene. These mechanisms include the usage of multiple transcription start sites, alternative pre-mRNA splicing, polyadenylation, RNA editing and post-translational modifications. Pre-mRNA splicing is considered the most important source of proteome diversity, as alternative splicing (AS) produces a large number of mRNAs that encode proteins with different or even opposite biological functions with crucial consequences on cells metabolism (Maniatis and Tasic 2002, Pan, Shai et al. 2008, Wang, Sandberg et al. 2008). It is a highly regulated process that plays a key role many cellular processes, such as sex determination, cell differentiation, cell transformation or apoptosis. It has been observed in almost all metazoan and it is estimated to occur in about 95% of all multi-exon human genes (Celotto and Graveley 2001).

Experiments on the *Dscam* gene in *D. melanogaster* provides information of the complexity of AS: this gene contains 116 exons, but only 17 are retained in the final mature transcript, implying a theoretical possibility to produce 36.016 different mRNAs. And, in fact, over 18.000 variants have been found in *D. melanogaster* (Schmucker, Clemens et al. 2000).

Many different models of AS are present in nature (figure 1.8). A single cassette exon can reside between two different constitutively spliced exons and can be either spliced out from the mature mRNA or retained (exon skipping and exon inclusion, respectively). Moreover, multiple cassette exons can be located between the two constitutive exons and the splicing machinery chooses between them (mutually exclusive exons). Exons can have different 5' or 3'ss that can be alternatively used. In addition, intronic sequences can be included in the mature messenger RNA and, in turn, translated into proteins whose functionality depends on the maintenance of the correct protein reading frame.

Different 5' starting point (first exons) can be used by the selection of different promoters, as well as the usage of alternative polyadenylation signals in combination with alternative splicing, that can regulate the usage of the 3'-terminal exon. All these events can be found in the same transcriptional unit providing a panel with a variety of different mRNAs derived from a single gene. On the transcriptome level, AS can produce variants of mRNAs with different stability: for example a mRNA with a prolonged lifespan can increase the availability of the corresponding protein, while ones with a premature termination codon (PTC) will follow the nonsense mediated decay (NMD) pathway (Faustino and Cooper 2003).

The mechanisms that promote which splice site will be used and/or which exon will be included in mature mRNA in different cell types or developmental stages have been studied in recent years, although it is clear that further studies are required to fully understand the molecular mechanisms at the basis of this phenomena. The intricacy of AS process resides not only in the complexity of the *cis*-acting regulatory elements, but also in the combinatorial control of different arrays of splicing factors whose

relative abundance and antagonistic effects have a key role in this process (Caceres, Stamm et al. 1994, Paradis, Cloutier et al. 2007, Dreumont, Hardy et al. 2010).

In the last years it has became more clear that splicing decisions are tightly coupled to epigenetic factors, such as RNA PolII elongation rate, nucleosome positioning and chromatin remodelling that can have an impact on exon definition and fate. Thus, the emerging picture of AS events underlines its extreme complexity and its precise regulation, what makes it a central event in the regulation of gene expression and cell fate (Blaustein, Pelisch et al. 2005, Luco, Pan et al. 2010).

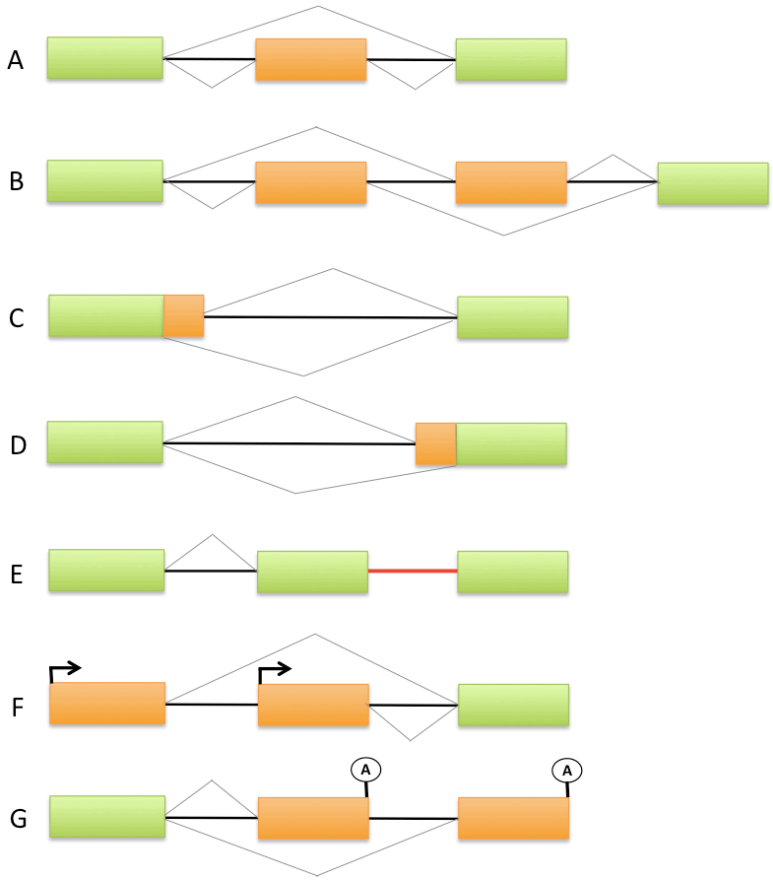


Figure 1.8. Pattern of alternative splicing.

The scheme represents different models of alternatively spliced exons. Constitutive exons are green boxes, alternative exons are the orange ones and introns are lines. The lines above and below the boxes indicated the possible splicing events. Arrows indicate the usage of alternative promoters, while circles the usage of alternative polyadenylation sites. AS of internal exons includes the cassette exon (A), mutually exclusive exons (B), alternative 5'-splice site (C), alternative 3'-splice site (D) and intron retention (E). The usage of alternative promoters leads to the selection of one of the multiple first exons (F), while the usage of alternative terminal exons generates mature transcripts with different polyadenylation sites (G).

1.3 Impact of mis-splicing events in clinical practice.

The development of fast sequencing methods have allowed in the past 20 years the discovery of genomic variants in the human DNA during the analysis of genes associated to diseases (Faustino, 2003; Pagani, 2003; Pagani, 2004; Buratti, 2006; Wessagowit, 2005; Baralle, 2009; Tazi, 2009). Nevertheless, genomic variants are not always ascribable to a clear molecular impairment of some biological functions (change of aminoacids, creation of premature termination codons, shift of the reading frame). Variations that seem to have no effect on the phenotype, may exert unpredictable and deleterious effect on pre-mRNA molecules causing severe diseases (Pagani and Baralle 2004, Buratti, Baralle et al. 2006). It has been estimated that 15% of pathological human mutations cause the splicing-associated diseases (Baralle, Lucassen et al. 2009). Although most of mutations affecting the splicing process are at the basis of inherited genetic pathologies, alterations of AS mechanism has been correlated also to acquired diseases like cancer (Ubby, Bussani et al. 2013). The study of the correlation between genomic variants that could potentially affect splicing and the occurrence of human diseases has become a central issue in medical practice. In fact, although is relatively easy to predict if a mutation affects splicing signals and will cause aberrant splicing, it is not so obvious in case of synonymous mutations located in exons or intronic polymorphisms, located far away from splice sites. They could be erroneously classified as benign substitutions, but frequently these mutations destroy or modify splicing regulatory elements, leading to the mis-recognition of the exon by the spliceosome machinery (Pagani and Baralle 2004, Baralle, Lucassen et al. 2009).

Mutations that impair the splicing mechanism could cause skipping of the exon, creation of a cryptic splice site, creation of pseudoexons or retention of intronic sequences (Hammond and Wood 2011). This could result in severe defects at the protein level through generation of non-functional proteins produced by out-of-frame transcripts or formation of premature termination codons (PTC) that drive the mRNAs through the nonsense-mediated decay (NMD) pathway (Lacroix, Lacaze-Buzy et al. 2012).

To provide tools to predict genomic variants affecting splicing, several *in silico* methods have been developed in order to identify splicing anomalies not only at canonical splicing signals but also in splicing regulatory elements (SREs) (Cartegni, Wang et al. 2003, Fairbrother, Yeo et al. 2004, Desmet, Hamroun et al. 2009).

1.3.1 Splice sites mutations.

Splicing signals are frequent targets of mutations in genetic disorders and cancer. Most of them are single point mutations affecting the conserved consensus sequences both at the 5'ss and at the 3'ss. Consensus sequences are limited to 9 bases at the 5' exon/intron border and 4 bases at the 3' intron/exon junction. It has been shown that mutations occurring in one of the first two bases (GU) in intron, immediately downstream the 5'ss, as well as the AG in the intron upstream the 3'ss completely abolish splicing (Krawczak, Thomas et al. 2007). However, pathological variations in nucleotides flanking the conserved AG/GU dinucleotides at 3' and 5'ss may be difficult to assess, because the consensus of the region show some variability (Krawczak, Thomas et al. 2007).

It was shown that variations at the 3'ss affecting the conserved -2 and -1 sites (AG) are the most frequent and have the effect of completely disrupt splicing; nevertheless, mutations affecting position +1 were also observed (Krawczak, Thomas et al. 2007).

Most common mutations at the 5'ss affect the invariant GU at position +1 and +2, followed by mutations in position +5 (Pohlenz, Dumitrescu et al. 2002). These variations are supposed to reduce the complementarity between the donor splice site on the pre m-RNA molecule and the 5'-tail of the U1snRNA, which is the first event that has to occur for spliceosome assembly (Black 2003). Although a mis-recognition of the 5'ss usually leads to exon skipping, different additional events can take place, ranging from defective donor splice site recognition to cryptic splice site activation, full intron inclusion or modification in RNA secondary structure (Cooper, Wan et al. 2009). It is important to highlight that despite defective donor splice site, recognition

leads to a decreased exon inclusion, yet some mRNA is produced and the protein product is functional. On other cases, the mature transcripts are usually not produced or the protein product is not functional. Therefore it is of a great importance to test each mutation to assess its effect on the splicing process and to develop diagnostic tools and therapeutic approaches (Garcia-Blanco, Baraniak et al. 2004, Spurdle, Couch et al. 2008, Tournier, Vezain et al. 2008, Hammond and Wood 2011). For instance, hybrid minigene systems have been broadly used for evaluation of splicing associated mutations in different gene models, such as *CFTR*, *NF1*, *ATM* and others (Pagani, Buratti et al. 2003). At the same time, this method allows the discovery and the characterization of novel regulatory elements, either intronic or exonic, which may participate in splice site selection (Pagani, Stuani et al. 2003, Cooper, Wan et al. 2009).

1.3.2 Mutations in auxiliary splicing regulatory sequences.

Nucleotide substitutions that involve auxiliary *cis*-acting elements can be classified as loss or gain of function mutations, relative to the fact if the splicing element is destroyed/weakened or created/enhanced, respectively (Faustino and Cooper 2003, Garcia-Blanco, Baraniak et al. 2004). Several mutations affecting exonic and intronic regulatory sequences, as well as composite exonic regulatory sequences of splicing (CERES) have been identified in different gene models (Pagani, Buratti et al. 2002, Pagani, Stuani et al. 2003, Wessagowit, Kim et al. 2005, Vidal, Cachia et al. 2009, Covaciu, Grosso et al. 2011).

A well-studied example in which the effect of an exonic regulatory sequence has been analysed is represented by spinal muscular atrophy (SMA), a disease in which the severity of the pathology is related to the degree of functional SMN protein deficiency. In humans there are two SMN genes: the vast majority of SMA patients have deletions of SMN1 gene and a single C-to-T transition in position +6 of SMN2 exon 7. This synonymous substitution alters the splicing of SMN2 pre-mRNA, causing frequent skipping of exon 7 that leads, in turn, to the production of an instable and not functional protein

that lacks the last 16 aminoacids (D'Amico, Mercuri et al. 2011). Two models have been proposed to explain exon 7 skipping: one is that the mutation +6C>T disrupts an ASF/SF2-dependent ESE (Cartegni, Chew et al. 2002), and the other is that it creates an ESS that recruits the splicing suppressor hnRNP A1 (Kashima, Rao et al. 2007, Kashima, Rao et al. 2007).

An additional example of a genomic variation affecting the non-canonical splicing regulatory sequence is represented by the dystrophin gene. Mutations within this gene are responsible for Duchenne muscular dystrophy (DMD). DMD is caused by loss of function mutations and, while more than 65% of DMD mutations are genomic deletions, a large number of exonic and intronic point variations that cause the disease through the defects in splicing mechanism (Cooper, Wan et al. 2009). Interestingly, a particular T-to-A substitution within exon 31 not only creates a premature termination codon, but also introduce and ESS that is bound by hnRNP A1, resulting in partial exon skipping. The mRNA that lacks exon 31 loses the coding sequence for one spectrin-like domain, but retains the correct reading frame that produces a partially functional protein, which gives the reason of the milder form of the disease in patients with this particular mutation (Disset, Bourgeois et al. 2006).

1.4 Tools for splicing manipulation

Over the last 10 years several attempts have been carried out in order to provide efficient tools to redirect aberrant splicing in disease-causing genes. Gene therapy approaches propose gene replacement as a therapeutic strategy to correct not only mutations disrupting the functionality of a gene but also proteins impairment due to splicing mutations at the RNA level. Design of a splicing correction approach to treat different disease-causing mutations is a process that requires in-depth study of the specific splicing biology behind each exon and, more in general, behind each gene under analysis, in order to provide the most feasible method to revert the mechanism that leads to aberrant splicing (Hammond and Wood 2011).

- *Improvement of exon inclusion.*

Spinal muscular atrophy (SMA) is an autosomal recessive neurodegenerative disorder affecting α-motoneurons in the anterior horn of the spinal cord. This pathology affects 1 in 6000 newborns with a carrier frequency of 1 in 40. It is formally caused by homozygous deletion or inactivation of SMN1 gene that encodes for survival motor neuron (SMN) protein (Lefebvre, Burglen et al. 1995). However, all the patients retain at least two copies of a homologous gene, named SMN2, which encode for the same SMN protein. *SMN2* is nearly identical to *SMN1*, with the exception of a single C-to-T synonymous substitution in position +6 of exon 7 that impairs its inclusion in mature transcript due to an unbalanced interaction of the splicing factors and the exon (Qin and Zhou 2008). In absence of full length SMN1 mRNAs, the 10% of full length transcripts produced by *SMN2* is not sufficient to provide the amount of SMN protein necessary for the normal cellular functions. Thus, although SMA is not *per se* a splicing disease, rescuing exon 7 skipping in SMN2 gene is a reliable therapeutic strategy (Hua, Vickers et al. 2007, Lorson, Rindt et al. 2010, Kolb and Kissel 2011).

In contraposition to pathologies whose aetiology is clearly linked to a splicing defect, there are some others in which only for a small number of patients the

onset of the disease is due to a defect in the splicing process. For example, in cystic fibrosis pathology in which only a minor group of patients showed mutations among the CFTR gene that have been shown to impair the correct recognition of exons causing their exclusion from the mature transcript (Pagani, Buratti et al. 2000, Pagani, Stuani et al. 2003). Another example of a pathologies caused by a variety of different mutations in which only a few affect the splicing mechanism are coagulation blood disorders. Genomic analysis of coagulation factor IX gene sequence in haemophilia B severely affected patients has revealed a multitude of genomic variants located in introns, exons or nearby splice sites that has been linked to skipping of different exons, impairing the production of Factor IX protein with coagulant activity (Fernandez Alanis, Pinotti et al. 2012).

For these kinds of pathology in which the manifestation of the disease is caused by splicing impairment only for a limited group of patients, a personalized therapeutic approach that takes into consideration each specific splicing mutation, could be an efficient way to provide the correct treatment (Baralle, Lucassen et al. 2009).

- *Looking for exon skipping.*

In other pathologies, such as Duchenne muscular dystrophy (DMD), the phenotype of the patient could be ameliorated causing the skipping of one or more exons. DMD is a genetic disease that affects 1 in 3500 male newborns, causing progressive muscular weakness and infiltration of fibrotic tissue into skeletal muscles. Death usually occurs from cardiovascular failure in early adulthood. The pathology is caused by defects in the dystrophin protein that is encoded by DMD gene, the longest one in the human genome, located at Xp21.2-p21.1 locus (Emery 2002). It is 2.4 Mb long and encodes a transcript composed of 79 exons, alternatively spliced to give rise to 7 different protein isoforms. Despite the extreme size of DMD gene, the average size of the produced mRNAs is about 14 Kb. This indicates the presence of very long introns in pre-mRNA molecules that makes splicing a very complex process. Aberrant splicing due to mutations in this gene leads, in the majority of cases,

to the creation of a premature termination codon (PTC) that, in turn, sends the transcript to nonsense-mediated decay (NMD). Patients that have deletions in the C-terminal part of the protein that causes its total disruption develop the most severe form of dystrophy, named Duchenne muscular dystrophy (DMD). Nevertheless, in-frame mutations in central repeated domains of the dystrophin transcript has been linked to a milder form of the disease, named Becker muscular dystrophy (BMD). For DMD the splicing-therapy approach looks for in-frame skipping of exons in order to restore the correct reading frame altered by mutations that impair the splicing process, partially recovering dystrophin protein function. Thus, several attempts have been carried out in order to modulate splicing for switching the severe form of DMD to the milder BMD form (Garcia-Blanco, Baraniak et al. 2004, Cooper, Wan et al. 2009, Hammond and Wood 2011).

- *Encouraging the choice of the correct 5' splice site.*

Splicing mutations can also lead to the creation of new splice sites, named cryptic splice sites, that are recognized with more efficiency by the spliceosome machinery. That causes the inclusion of intronic sequence as part of exons or the exclusion of coding sequences from the mature mRNA. Hutchinson-Gilford progeria syndrome is a rare dominant autosomal disease affecting 1:4.000.000 newborns. Its clinical maniphestations results in accelerated aging and the mortality is primarily due to accelerated atherosclerosis. A single C-to-T base substitution within LMNA exon 11 activates a cryptic splice site which produces a transcript shorter than the normal one and encodes for a LMNA protein that laks 50 aa, named progerin, which is toxic to cells. Progerin impairs the levels of functional B1 laminin, causing disruption of the nuclear lamina (Prokocimer, Barkan et al. 2013).

1.4.1 Antisense oligonucleotides (AOs)

One of the most used techniques to redirect splicing is based on the use of antisense oligonucleotides (AOs) that target specific regions in the pre-mRNA molecule. They are able to efficiently bind specific sequences masking them and avoiding the interactions of regulatory regions with their corresponding *trans*-acting factors. Steric block of sequences carrying a mutation or a regulatory element results in changes of the pattern of splicing. This approach was used for the first time to redirect aberrant splicing of β-globin (Dominski and Kole 1993). So far, it has been extensively used for correction of different mis-splicing diseases as AON are able, depending by their targeted region, to promote exon inclusion or cause exon skipping (Hua, Vickers et al. 2007, Aartsma-Rus, Fokkema et al. 2009), to avoid the recognition of cryptic exons (Svasti, Suwanmanee et al. 2009) or to alter the balance of alternative spliced exons (Khoo, Roca et al. 2007).

In order to exert their corrective activity AOs need some peculiar features: first they have to be easily internalized in the cell and they must be able to reach the nucleus, where the splicing reaction takes place. Another important point is their specificity to the target sequence, as off-targets events may result in severe alterations in cell metabolism and cellular toxicity. Thirdly, they must be able to resist the endogenous nucleases in order to maximize their potential duration of action. As normal ribonucleotides are normally exposed to the degradation activity of ribonucleases (especially RNase H), several modified AOs have been developed in order to improve their stability, cellular uptake and affinity to the desired target sequence. Chemical modifications have introduced modified AOs as splicing modulating tools commonly used for conventional basepairing with RNA molecules: 2'-O-methylphosphorothioate (2'OMePS), 2'-O-methoxyethyl phosphorothioate (2'MOE PS), phosphorodiamidate morpholino (PMO) and peptide nucleic acid (PNA), that are not substrates for RNase H.

The decision to induce the skipping of a specific exon from the mature transcripts has to be carefully evaluated in order to prove the therapeutic efficacy of the approach. It has to be evaluated if the exclusion of the exon

alters the protein-coding frame, as well as the effect on the function of the protein lacking the domain encoded by the skipped exon. To induce the skipping of an exon the most suitable target sequences are the canonical splice signals, since their steric blockage makes the region not accessible to the spliceosome components

An example of pathology in which antisense oligonucleotides have been successfully used to redirect splicing is Duchenne muscular dystrophy in which the AO-mediated skipping of exon 51 restores the frameshift caused by exon 50 deletion, promoting the production of functional dystrophin protein.

Conversely, AOs can be used also to promote exon inclusion, where the exons are predominantly skipped from the mature mRNA. In case of SMN2 gene, the proper recognition of exon 7 is impaired by a synonymous exonic transversion. In this case, the improvement of exon 7 inclusion is a reliable therapeutic strategy for treatment of spinal muscular atrophy, where AOs have been designed to target an intronic splicing silencer located downstream of exon 7 donor splice site. Steric block of the hnRNP A1-dependent ISS restores the correct mRNA splicing and increases the cellular levels of full-length SMN protein.

Moreover, AOs have been used to prevent the usage of the cryptic splice sites, which have been found with the same frequency in exons and introns. When mutations weaken the canonical splice sites they are used instead, what leads to the formation of aberrant transcripts lacking portion of exons or retaining intronic fragments. The most frequent consequences are the formation of premature termination codons (PTCs) or frameshifts. A protypical example of a cryptic splice site activation is provided by ATM exon 55 in which an exonic mutation is responsible for Ataxia telangiectasia. This genomic variant creates a cryptic 5'-splice site that is preferentially used instead of the normal one, causing an exonic deletion of 64 nucleotides what leads to a frameshift in the transcript. AO-mediated masking of the cryptic splice site avoids the interaction of the spliceosome with the mutated sequence and stimulates the usage of the normal downstream 5'ss (Teraoka, Telatar et al. 1999, Du, Pollard et al. 2007).

Pseudoexons are intronic sequences that carry at their extremities weak splicing signals that can be recognized by the spliceosome machinery. In general, formation of pathological pseudoexons are due to the creation of canonical splicing signals (3' or 5'ss sequence or a novel branch point site) or to mutations in splicing regulatory sequences that recruit splicing factors with enhancer or silencer functions. For instance, several patients with mutations in the MUT gene have been found to incorporate a 74 base long intronic sequence from intron 11 in the mature mRNA. This is a due of a pathological genomic variant that strengthens the 5'ss of a pseudoexon forcing the spliceosome in the choice of the cryptic splice site instead of the canonical weaker ones. In this case an AO targeting the 5'ss of the pseudoexon was used in order to impair its recognition as real exon (Perez, Rincon et al. 2009, Dhir and Buratti 2010).

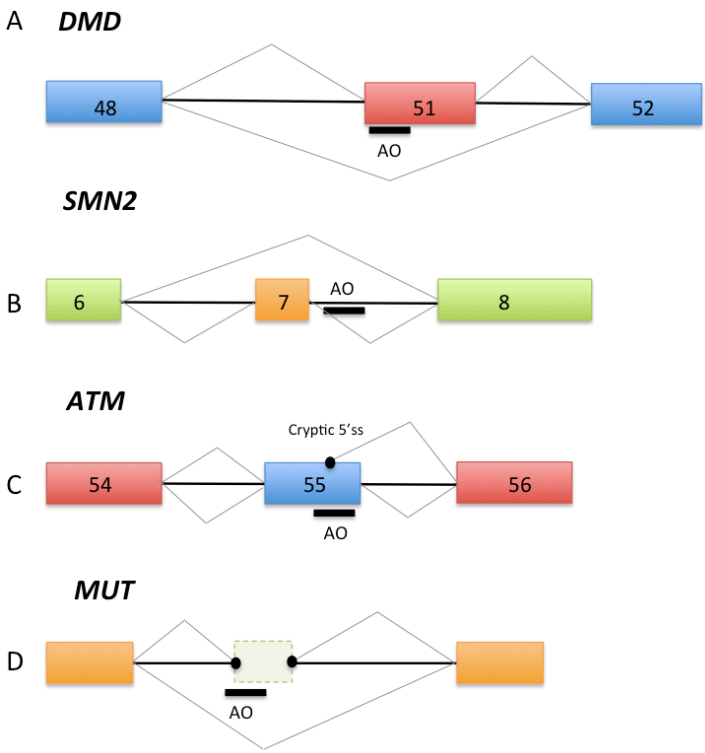


Figure 1.9. Antisense oligonucleotide mediated correction of aberrant splicing.

A. In *DMD* gene, a large scale deletion of two exons (49-50) results in a frameshift of the transcript. The use of an antisense oligonucleotide (AO) targeting an enhancer located in intron 51 leads to the skipping of this exon, restoring a correct protein reading frame without consequences on protein function.

B. In *SMN2* gene a C-to-T synonymous substitution within exon 7 impairs its inclusion in the mature mRNA. AO-mediated masking of an hnRNP A1-dependent intronic splicing silencer (ISS-N1) downstream the donor splice site rescues exon 7 inclusion.

C. In *ATM* gene an exonic mutation in exon 55 creates a strong 5'ss signal that is preferentially recognized by the splicing machinery instead of the canonical donor splice site. The resulting mRNA lacks an exonic portion of 64 nt. The use of an AO targeting the cryptic splice site region prevents the spliceosome from the selection of the cryptic site driving the recognition of the canonical one.

D. A mutation within intron 11 of *MUT* gene strengthen the 5'ss of a pseudoexon that is recognized as an exon and included in the final transcripts. AO-mediated targeting of the 5'ss of the pseudoexon avoids its recognition and restore the natural mRNA.

1.4.1.2 AON-based clinical trials for SMA and DMD.

Antisense oligonucleotide-based therapy has reached, for some splicing-related pathologies such as Duchenne muscular dystrophy (DMD) and spinal muscular atrophy (SMA), high levels of translation in clinics.

The background of mutations in DMD is high and heterogeneous, but most of the mutations happen within exon 50, causing the deletion of this exon from the final mRNA and leading to an "out of frame" transcript. Thus, for DMD, clinical trials using antisense oligonucleotide were developed with the aim to promote exon 51 skipping from dystrophyn mRNA in order to promote a shorter, but in frame, transcript lacking exons 50-51 which was previously shown to encode for a functional protein that correlates with a milder form of muscular dystrophy, named Baker muscular dystrophy (BMD). Prosensa Therapeutics (www.prosensa.eu) developed a compound, named

Drisapersen®, that is currently in clinical trial phase III, which is aimed to investigate the safety, tolerability and efficacy in the long-term in Duchenne muscular dystrophy patients (www.clinicaltrials.com). Drisapersen is intended for 13% of sub-population of DMD patients that includes those with deletions of exon 50, exon 52, exons 45-50, exons 48-50, and exons 49-50.
Sarepta Therapeutics (www.sarepta.com) developed Eteplirsen® a compound that, similarly to Drisapersen by Prosensa, is aimed to promote exon 51 skipping. Presently, Eteplirsen® is undergoing a clinical trial phase IIa (www.clinicaltrials.com) to investigate safety and long-term efficacy in DMD patients of 7-13 years of age.
For spinal muscular atrophy, a reliable therapeutic strategy is to correct SMN2 exon 7 skipping whose inclusion in the final mRNA is impaired by a C-to-T synonymous transition within exon 7. Isis Pharmaceuticals (www.isispharm.com) in collaboration with dr. Adrian Krainer (Cold Spring Harbor Laboratories) developed a compound named ISIS-SMNRx® which is potentially applicable to all the SMA patients, giving that promotes SMN2 exon 7 inclusion. ISIS-SMNRx is currently in phase II clinical trials in which the compound is tested for safety, tolerability, and pharmacokinetics of multiple doses of ISIS-SMNRx administered into the spinal fluid (www.clinicaltrials.gov).

1.4.2 U-rich small nuclear RNAs (U-snRNAs).

To modulate splicing in order to correct aberrant splicing events, U-rich small nuclear RNA, in particular U1 snRNAs and U7 snRNAs, have been engineered. U1 and U7 snRNPs are composed of 7 common Sm proteins and a set of specific proteins bound by each snRNA (Brun, Suter et al. 2003, Uguen and Murphy 2004, Martone, De Angelis et al. 2012)
In comparison to naked antisense oligonucleotides, the usage of antisense sequences embedded in small nuclear RNAs provides some advantages, the most important of which is that the snRNA prevents the antisense sequence

from the degradation prolonging its half life and allows its high rate expression in the cell.

1.4.2.1 Modified U1 snRNAs

The U1 snRNP is the first component of the spliceosome that interacts, through its 5' protruding RNA tail, with the donor splice site on the pre-mRNA molecule. It is composed of a 164 nucleotides long RNA (small nuclear RNA – snRNA) with a well defined secondary structure (see paragraph 1.2.3) that recruits 7 Sm proteins and the U1-specific 70-K, U1-A and U1-C (Hamm, Dathan et al. 1990, Nelissen, Will et al. 1994, Egloff, O'Reilly et al. 2008).

As U1 snRNA is naturally involved in the splicing mechanism and processed in the nucleus, simple modification of the 5'-tail of the U1 snRNA is possible to direct its recognition of specific regions in the pre-mRNA, avoiding problems of correct compartimentalization. It has been used both to improve the recognition of an exon and to cause exon skipping. Increasing the consensus sequence of the U1 snRNA tail according to the mutated 5'ss of a specific exon, it has been possible to redirect the aberrant splicing of exon 8 in the coagulation factor VII gene (Pinotti, Rizzotto et al. 2008). Furthermore, modified U1 snRNAs with an increased complementarity to defective donor splice sites have been proved to be effective also in combination with modified U6 snRNAs for splicing correction (Schmid, Hiller et al. 2012).

Moreover, U1 snRNAs have been engineered also as carriers of antisense sequences to target regulatory regions and modulate splicing. The advantage in comparison to antisense oligonucleotide is that the sequence, embedded in the snRNA, is less exposed to degradation (Martone, De Angelis et al. 2012). In case of the *DMD* gene, modified U1s have been successfully used to promote the skipping of exon 51 in human fibroblasts (De Angelis, Sthandier et al. 2002, Incitti, De Angelis et al. 2010). *In vivo*, engineered U1 snRNA demonstrate efficacy in rescuing pathological phenotype of the Duchenne muscular dystrophy in the mouse model. In this case, the antisense-U1 snRNA ,that induced the skipping of exon 23 of the murine *Dmd* gene, avoids

the formation of a premature termination codon (PTC) and was delivered using adeno-associated vectors (Denti, Rosa et al. 2006).

1.4.2.2 Modified U7 snRNAs

U7 small nuclear RNAs are molecules involved in the processing of histone mRNA. Thus, differently from U1 snRNAs, they are not naturally involved in the splicing mechanism and their utilization as splicing modifiers has required some molecular modification. The optimized U7 snRNA (U7-Opt) with disrupted activity of histone mRNA processing has been broadly applied for correction of splicing-related diseases (Stefanovich, 1995). The RNA structure binds only the 7 Sm proteins common to all the snRNAs and has a 5'-tail that has been engineered to specifically target pre-mRNA regions promoting inclusion or exclusion of exons from the mature transcripts. This corrective tool has been used for correction of exon 7 skipping of the SMN2 gene in primary fibroblasts from spinal muscular atrophy patients (Madocsai, Lim et al. 2005). Different anti-SMN U7 snRNAs, targeting regulatory regions both in intron 7 and exon 7 have been shown to increase the full-length SMN2 mRNA and the corresponding SMN protein level in patients' cells and *in vivo* in a SMA severe mouse model (Meyer, Marquis et al. 2009).

One of the first attempts to modulate splicing using U7-antisense-snRNAs was successfully carried out in 2004 by Goyanevalle and collegues. In this seminal work they promoted the skipping of a mutated exon associated with Duchenne muscular dystrophy *in vivo* in a mdx mouse delivering the U7 molecule through an AAV vector and demonstrating the recovery of a functional dystrophyn protein and the correction of the muscular dystrophy (Goyanvalle et al., 2004).

Moreover, U7-Opt snRNAs have been modified not only to bind exonic or intronic splicing regulatory sequences, but also to attract the binding of *trans*-acting splicing factors on the pre-mRNA. In these cases, the molecule has been engineered to mask a target sequence creating a steric blockage for binding of some splicing factor and, at the same time, the 5'-tail of the U7

carrying an ESS or ESE was able to recall splicing proteins, thus modulating splicing in a stronger manner. A bifunctional U7 carrying a splicing enhancer (U7-ESE-B) was able to promote exon 7 inclusion in the SMN2 transcript, rescuing the phenotype of severe SMA mouse model both by transgenic expression of the U7-ESE-B gene and via viral administration using lentiviral vectors (Marquis, Meyer et al. 2007).

1.4.4 Trans-splicing

Redirecting splicing through a *trans*-splicing approach involves a sequence exchange between a target pre-mRNA and a pre-*trans*-splicing molecule (PTM). *Trans*-splicing is an event that in nature occurs in different organisms, such as for example *C. elegans*. Differently from the canonical *cis*-acting reaction this splicing mechanism shows a Y shaped branched intermediate instead of the looped lariat immediately after the first transesterification reaction. For corrective purposes, this mechanism has been exploited to change the 3'-end of pre-mRNA molecules carrying mutations or deletions, providing in *trans* the correct functional PTM molecule (Puttaraju, Jamison et al. 1999). The PTM usually targets an intronic region upstream the sequence that has to be substituted and contains a functional branch point, a polypyrimidine tract and the AG dinucleotide at the 3'ss, that are fundamental elements for the recognition of the exon that has to be *trans*-spliced. The spliceosome machinery processes some transcripts in *cis* and other in *trans* providing functional mRNAs, although *in vivo* this strategy has revealed a very low efficiency (Coady, Baughan et al. 2008, Coady and Lorson 2010). However, one of the advantages of this method is that a large number of patients carrying different type of mutations could be potentially treated (Hammond and Wood 2011).

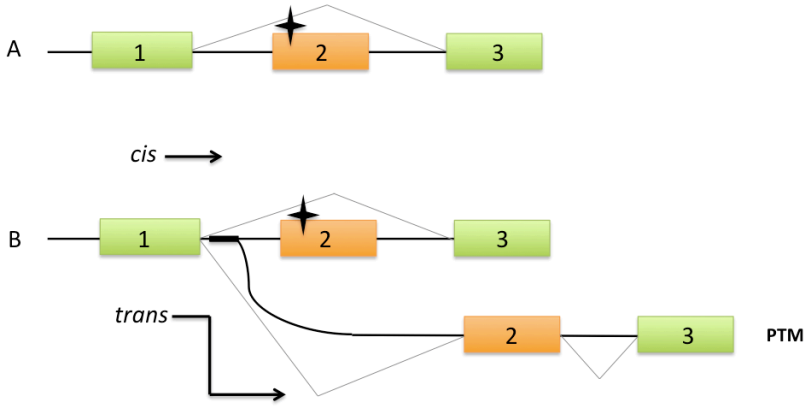


Figure 1.10. *Trans*-splicing mechanism.

A. Schematic representation of a pre-mRNA carrying a mutation in exon 2 which impair its inclusion in the final mRNA. Boxes are exons and lines are introns. The lines above and below the boxes represent the possible splicing patterns.

B. A pre-*trans*-splicing molecule (PTM) targets an intronic region upstream the exon 2 carrying the mutation. The PTM has an intact branch point sequence, polypyrimidine tract and the AG at the 3'ss of exon 2. The spliceosome processes some transcripts in *cis* with the consequent production of an aberrant mRNA and some transcripts in *trans* leading to the production of a normal mRNA that encodes for a functional protein.

GENE	PATHOLOGY	GENETIC THERAPY	MOST RECENT STUDIES
ATM	Ataxia Telengiectasia	PMO	Treatement of cell lines carrying homozygous mutations
APOB	Atherosclerosis	2'OMePS	Human liver carcinoma cells
β-globin	β-Thalassemia	PMO	Mouse model of intron 2 cryptic splice site Isolated mononuclear cells from patients
PMM2	Congenital disorder of glycosilation	PMO	Correction in splicing fibroblasts
DMD	Duchenne muscular	2'OMePS PMO Bifunctional U7 Modified U1	Clinical trials phase I/Iia Clinical trials phase I/Iia Myoblasts from patients Mouse model of DMD
DYSF	Dysferlinophaties	AON U7 antisense	Myoblasts from patients
COL7A1	Epydermolisis bullosa	2'OMePS	skin grafts on mouse injected with AON
LMNA	Hutchinson-Gilford syndrome	2'OMePS PMO	Minigene systems Patient derived fibroblasts
MUT	Methylmalonic aciduria	PMO	Patient fibroblasts
NF1	Neurofibromatosis type I	PMO	Patientderived lymphocytes
SMN2	Spinal muscular atrophy	2'OMePS Bifunctional U7 U7 antisense *Trans*-splicing	Severe SMA mouse model Patient fibroblasts Severe SMA mouse model

Figure 1.11. Examples of current genetic treatments for splicing-related diseases.

1.5 Spinal Muscular Atrophy (SMA).

Spinal Muscular Atrophy (SMA) is an inherited autosomal recessive severe neurodegenerative disorder with high incidence in newborns (1:1000).

Its clinical manifestation is characterized by progressive muscular weakness and atrophy. First symptoms appear in the early childhood. SMA have been described and clinically classified into three different types (I, II or III), based on severity and age of onset (D'Amico, Mercuri et al. 2011). The most severe is type I (OMIMs 253300) in which respiratory failure is the most common cause of death and typically occurs before 2 years of age. Some more recent classification include type 0 (severe neonatal onset) and type IV (adult onset) (Kolb and Kissel 2011).

SMA is formally caused by homozygous deletion or mutations of Survival Motor Neuron 1 (*SMN1*) gene that encodes for SMN, a protein involved in several crucial cellular functions although in many cases the details are not been yet unravelled (Lefebvre, Burglen et al. 1995).

Survival of motor neuron protein (SMN) has 294 aminoacids with a molecular weight of 34 kDa and its degree of conservation among species is very high. It is involved in the biogenesis of small nuclear ribonucleoparticles (snRNPs) and in the cytoplasmic shuttle of mRNAs from the nucleus. In the cytoplasm the SMN protein colocalizes with Gemins protein, directing the assembly of Sm proteins onto the snRNAs. After modifications in the cytoplasm, the snRNPs return to the nucleus in structures named Cajal bodies, in which SMN was found to colocalize in structure called *gems*. Recent studies discovered the implication of SMN protein in a complex network of neuronal circuitry (Gogliotti, Quinlan et al. 2012, Lotti, Imlach et al. 2012).

Although deletion of SMN1 gene totally abolishes the production of SMN protein from this gene, all SMA patients retain at least two copies of a *SMN1* homologous gene, named *SMN2*. These genes encodes the same SMN protein, and are both located on chromosome 7 and are nearly identical, for the exception of a silent C-to-T transition at position +6 of SMN2 exon 7. This synonymous substitution impairs the efficient recognition of exon 7, prejudging its correct inclusion in the mature mRNA due to an unbalance in the fine-tuned mechanism that recruits splicing factors on the exon. In particular, the mutation destroys an ASF/SF2-dependent ESE, creating a strong binding site for hnRNP A1 that, in turn, avoids the binding of ASF/SF2 to the 5' region of exon 7. Exon 7 sequence and the surrounding intronic regions have been broadly investigated in order to detect the regulatory sequences, trying to understand the network of protein interaction on SMN1 and SMN2 pre-messenger RNAs (Cartegni and Krainer 2002, Kashima, Rao et al. 2007).

The majority of SMN transcripts coming from SMN2 gene lacks exon 7, leading to a small production of the full-length functional SMN protein, which is not able to compensate the absence of functional protein from SMN1 since

the truncated Δ7 protein lacks an important C-terminal domain, required for its stability and oligomerization(Cho and Dreyfuss 2010). However, it seems that SMN Δ7 protein retains at least a small amount of functionality, as it has proven *in vivo* that transgenic SMA mice expressing the truncated SMN protein have a milder phenotype (Le, Pham et al. 2005).

This is the reason why SMN2 is an attractive target for a splicing therapy: in fact several attempts has been carried out in the past 15 years to redirect splicing of SMN2 gene for SMA treatment using AOs, modified U7 snRNAs and bifunctional oligonucleotides (Skordis, Dunckley et al. 2003, Madocsai, Lim et al. 2005, Hua, Vickers et al. 2007, Marquis, Meyer et al. 2007, Meyer, Marquis et al. 2009, Hua, Sahashi et al. 2010, Hua, Sahashi et al. 2011, Rigo, Hua et al. 2012).

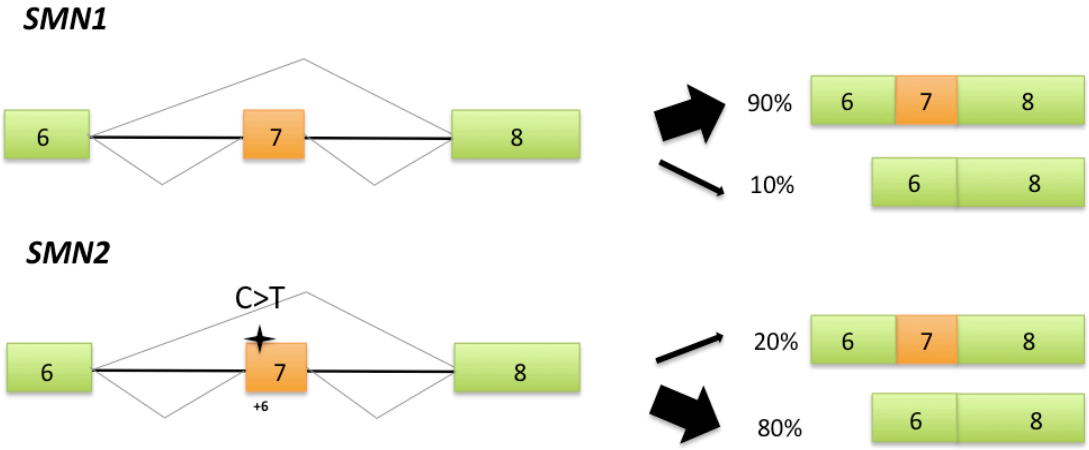


Figure 1.12. Schematic representation of exon 7 pattern of splicing in SMN genes.

Two inverted SMN genes, SMN1 and SMN2 are located on chromosome 5. Both possess 9 exon and exon 7 is alternatively spliced. In normal conditions ~90% of transcripts from SMN1 contain exon 7 and ~20% of transcripts from SMN2 contain exon 7. Spinal muscular atrophy patients have homozygous inactivation or deletion of the SMN1 gene. They retain only the SMN2 gene which is unable to compensate the loss of SMN1, as it produces only ~20% of full length transcript. The protein that is produced from SMN2 major product, the isoform lacking exon 7, is unstable and is rapidly degraded.

1.6 Netherton Syndrome (NS).

Netherton syndrome (NS, OMIM 256500) is a rare (1 in 200,000 newborns) recessive autosomal genodermatosis characterized by ichthyosiform erythroderma, a specific hair shaft defect (*trichorrhexis invaginata or bamboo hair*) and constant atopic manifestations. This skin disease begins at birth or shortly afterwards with generalized exfoliative erythroderma, which can persist throughout life in most severe cases or evolves in a milder ichthyosiform condition known as *ichthyosis linearis circumflexa (ILC)*. Hypernatraemic dehydration, infections and failure to thrive are frequent complications during the postnatal period, which is correlated with the relatively high level of mortality. NS is caused by loss of function of lymphoepithelial Kazal-type related inhibitor (LEKTI), which is an important regulator of epidermal physiology. It was observed that deficiency of LEKTI protein is related to abnormal cleavage of desmosome at the granular layer-stratum corneum transition both in human and mice. LEKTI is a 15-domain serine-protease inhibitor preceded by a signal peptide, which is strongly expressed in the granular layer of the epidermis and, in the inner root, sheets of hair follicles (figure 1.13) (Fortugno, Bresciani et al. 2011, Diociaiuti, Castiglia et al. 2013).

The LEKTI primary structure shows that each domain shares a Kazal-type related motif found in serine-protease (SP) inhibitors. However, only two of these 15 domains are typical Kazal-motif with 6 cysteine residues and precise spacing (D2 and D15), while the other domains contains only 4 cysteine patterns .

This protein is encoded by *SPINK5* (Serine protease inhibitor Kazal-type 5) gene into three different transcripts that differ through their C-terminal end by alternative splicing. All three isoforms are translated into differentiated keratinocytes as pro-protein, cleaved rapidly intracellular, transported in lamellar granules and then secreted as bioactive fragments at the granular and stratum corneum interface. LEKTI fragments efficiently inhibit epidermal proteases such as kallikrein (KLK) 5, KLK7 and KLK14, allowing the

maintenance of permeability barrier homeostasis (Fortugno, Bresciani et al. 2011).

To date, 62 different mutations are reported in *SPINK5* gene causing deficiency of LEKTI protein, that results in unopposed protease activity in the epidermis, leading to premature corneodesmosome cleavage at the granular layer-stratum corneum interface. This leads to defective adhesion of stratum corneum and results in the characteristic loss of skin barrier function observed in NS patients (Lacroix, Lacaze-Buzy et al. 2012).

The human *SPINK5* is a single copy gene composed of 33 exons, which maps to chromosome 5q32 (figure 1.14).

In physiological conditions, this gene encodes for three different isoforms by alternative processing of pre-mRNA. All transcripts display the same transcription start site, but differ from the C-terminal portion: the full-length $SPINK5_{f\text{-}l}$ mRNA is characterized by all 33 exons that encode for a 145 kDa precursor. The shorter $SPINK5_{sh}$ is created by the use of an internal polyadenylation site presented within intron 28 and encodes for a 125 kDa precursor. The last isoform $SPINK5_{l}$ is 90bp longer than full-length due to a cryptic exon (exon 28a) located within intron 28, encoding for a 148 kDa protein. Nevertheless, the reason because *SPINK5* gene produces different protein isoforms with related functions is still unknown (Fortugno, Bresciani et al. 2011).

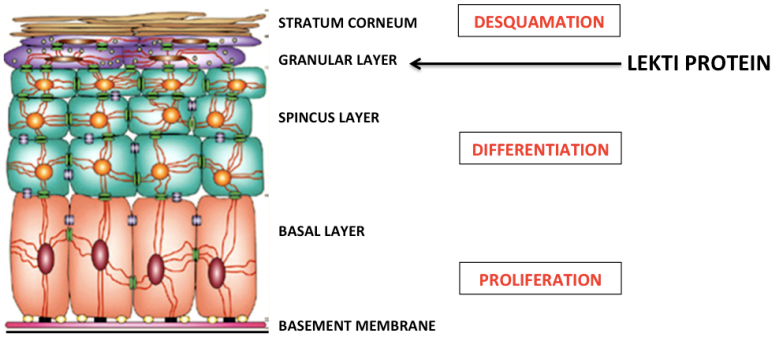


Figure 1.13. The epidermis.

The epidermis is a stratified squamous epithelium, composed of proliferating basal and differentiated suprabasal keratinocytes. LEKTI protein is strongly expressed in the granular and uppermost spinous layers of the epidermis and in the most differentiated layers of all stratified epithelia. LEKTI deficiency causes abnormal desmosome cleavage at the granular layer-stratum corneum transition.

1.6.1 The synonymous c.891C>T mutation in SPINK5 gene.

The c.891C>T substitution represents the most common *SPINK5* mutation described in the European population and induces aberrant processing of SPINK5 pre-mRNA. The c.891C>T is a synonymous mutation (p.Cys297Cys) localized 9 bp from the acceptor splice site of exon 11 and is directly responsible for skipping of the same exon, independently of other mutations of the transcript (Fortugno, Grosso et al. 2012, Lacroix, Lacaze-Buzy et al. 2012). In patients' keratinocytes, the processing of c.891C>T mutant pre-mRNA leads to skipping of entire exon 11 and consequent joining of exon 10 with the exon 12. The analysis using minigene splicing assay of this mutations confirmed the *in vivo* situation, with a total skipping of exon 11 in the mutant constructs. The molecular mechanism by which this nucleotide change affects pre-mRNA processing is still not clear.

The aberrant splicing caused by this synonymous mutation change the phase of translation, leading to a premature termination codon (TGA) within exon 12. However, *in vivo* this deleterious effect is not complete because of detected residual amount of full-length mRNA encoding for LEKTI polypeptides, correlating with the less severe phenotype in the patients. In fact, this synonymous mutation affects only one allele of compound heterozygous patients (Fortugno, Grosso et al. 2012, Lacroix, Lacaze-Buzy et al. 2012).

To date, there are no effective treatments for NS, but recent studies have shown the possibility to recover the situation using a gene therapy approach (Di, Larcher et al. 2011, Di, Mellerio et al. 2013). Both studies are based on

SPINK5 gene expression in keratinocytes stem cells through the use of lentiviral vectors that carry the gene. However, this strategy was not differentiation-specific and long-term experiments also showed a silencing phenomenon.

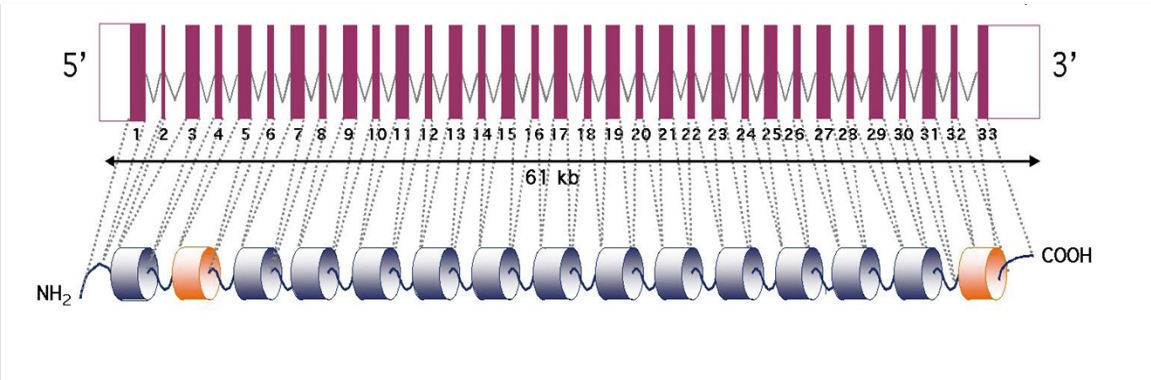


Figure 1.14. Schematic representation of SPINK5 and the predicted protein.

SPINK5 (61 kb, top) comprises 33 exons and 32 introns. Exon 1 and 33 include the start codon (ATG) and the natural stop codon (TGA) of translation, respectively. 5' and 3' UTR, coding regions and introns are denoted by white boxes, purple boxes and dashed lines, respectively. The 15 highly homologous Kazal-type modules of LEKTI (1,064 residues) are depicted at the bottom.

1.7 Aims of the thesis.

An efficient and correct splicing process relies on a multitude of interactions that involve splicesome-associated proteins and other *trans*-acting factors. As first step in exon recognition the U1 small nuclear ribonucleoparticle (U1 snRNP) binds the donor splice site and it is the key event committing the exon to splicing catalysis.

The impact of new genomic variants could have dramatic consequences on the splicing outcome of some pre-mRNA, leading to severe pathologies.

In this thesis work I have explored the possibility to revert aberrant splicing in two model genes associated with spinal muscular atrophy (SMA) and Netherton syndrome (NS), *SMN* and *SPINK5* respectively, using modified U1s (named Exon Specific U1s) that recognize intronic sequences located downstream the canonical 5' splice site of an exon.

Using different Exon Specific U1s, minigene systems, Hek293 Flip-In stable clones and primary cells from SMA and NS patients, the following aspects have been investigated:

- the identification of active ExSpe U1s able to correct mis-splicing events in SMN and SPINK5 pre-mRNAs;
- the assessment of the potential therapeutic effect of Exon Specific U1s in spinal muscular atrophy primary fibroblasts and Netherton syndrome primary keratinocytes;
- the characterization of the ExSpe U1s molecule, the investigation of the determinants involved in ExSpe U1-mediated splicing rescue and the possible mechanism of action of the ExSpe U1s;
- the identification of the players involved in SPINK5 exon 11 splicing in normal and pathological conditions.

.

2. RESULTS

2.1 Identification of modified U1 snRNAs molecules with splicing-modulation efficacy.

Previous studies showed that changing the 5'-tail of the U1 to increase its complementarity to a defective or not completely conserved 5'ss leads to a higher inclusion of the exon in the mature transcript (Zahler, Tuttle et al. 2004, Pinotti, Rizzotto et al. 2008, Sanchez-Alcudia, Perez et al. 2011, Schmid, Glaus et al. 2011, Schmid, Hiller et al. 2012). To correct aberrant exon skipping in two different gene models, SMN exon 7 and SPINK5 exon 11, I engineered the tail of the U1snRNA to target by complementarity either the 5'-splice site (5'ss) or intronic sequences located downstream of the exon that are, in comparison to the donor splice site region, less conserved.

2.1.1 Screening of exon-specific U1 molecules on SMN2 exon 7 minigene model.

I focussed my attention on SMN exon 7 splicing, an exon who plays a key role in the pathogenesis of a severe neurodegenerative disease, spinal muscular atrophy (SMA). Normal individuals possess two *Survival Motor Neuron* genes (*SMN1* and *SMN2*) that differ for a synonymous substitution within exon 7. This C>T transition in SMN2 gene causes the skipping of the exon from the mature SMN2 transcripts (figure 2.1). As SMA patients lack *SMN1* but retain the *SMN2* gene, increasing the inclusion of exon 7 in SMN2 messenger RNA (mRNA) has been broadly described as a reliable therapeutic approach for spinal muscular atrophy (Hua, Vickers et al. 2007, Hua, Vickers et al. 2008, Zhou, Zheng et al. 2012). In order to provide a new strategy to modulate exon 7 splicing, I created engineered U1 snRNAs in which I mutated their 5'-tails to allow their binding to a specific sequence of interest and I tested these new

bioengineered molecules in SMN2 minigene splicing assays. Giving that exon 7 natural donor splice site does not perfectly match the canonical consensus sequence of the 5'ss, as first approach I increased the RNA:RNA complementarity between the two sequences, modifying the 5'-tail of the U1 in position -3 and -1 (named U1 SMN 5'ss) in order to perfectly basepair the corresponding pre-mRNA sequence (figure 2.2 A). I tested the capacity of the modified U1 SMN 5'ss to affect SMN2 exon 7 splicing pattern. Transfection in Hek293 cells of pCI SMN2 plasmid along with U1 SMN 5'ss showed that increasing the complementarity of the modified U1 improved the percentage of exon 7 inclusion from 40% up to 90% (figure 2.2 B, compare lanes 1 and 8). Overexpression of U1 WT did not change the splicing pattern of SMN2 exon 7 in minigene splicing assays (data not shown).

Then, I decided to create modified U1s, named exon-specific U1s (ExSpe U1s) that target intronic regions downstream the canonical exon 7 5'-splice site. The numbering of each ExSpe U1 sm2, sm 17, sm 21, sm22 reflect their position relative to the nucleotide +1 of intron 7 (figure 2.2 A) The target sequences of ExSpe U1s EE1, EE2 and EE3 are not indicated because they are undergoing a patent registration process. ExSpe sm2 partially overlap (from position +2 to +6) with the last 4 bases of the 5'ss consensus sequence that is bound by the endogenous U1 snRNP particle. ExSpe U1s sm 17, 21, 22 and EE1 overlap totally (sm17) or in part (sm21, sm22, EE1) with the previously described hnRNPA1-dependent intronic splicing silencer N1 (ISS-N1). ExSpe U1 EE2 and EE3 bind outside of the ISS-N1. The specific sequences targeted by the ExSpe U1s are indicated in figure 2.2, panel A. Minigene splicing assay carried out with pCI SMN2 plasmid and the ExSpe U1s described above, revealed that all modified U1s modulate splicing with the same efficacy, promoting exon inclusion up to 90-95%, independently from their binding site (figure 2.2, panel B, lanes 2-7). Sequencing of the upper bands showed the usage of the correct 5'ss. These results indicate that both modified-U1 snRNA complementary to the donor splice site and to downstream intronic sequences exerts a strong splicing modulation. Moreover, their efficacy does not seem related to any specific sequence targeted.

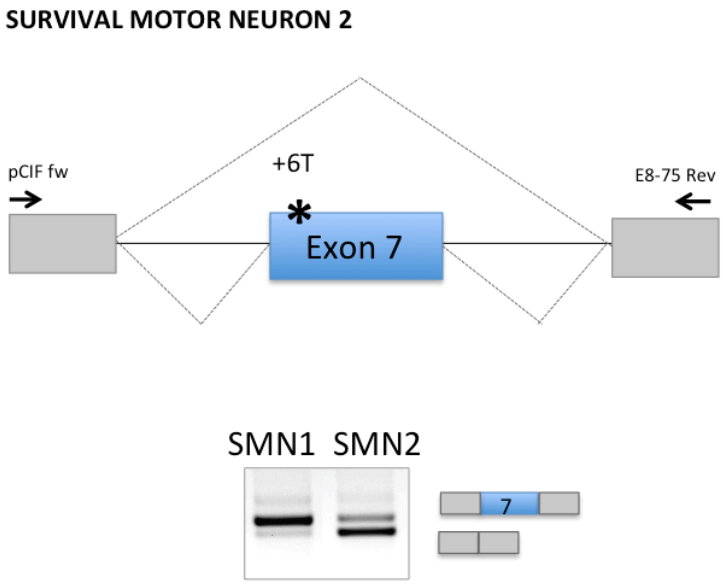


Figure 2.1. SMN1 and SMN2 exon 7 splicing patterns.

Upper panel: schematic representation of SMN2 exon 7 minigene. The boxes are the exon and the lines the introns. The asterisk showed the localization of the +6 C>T synonymous substitution that differs SMN2 from SMN1. The dotted lines indicate the two possible pattern of splicing (exon inclusion or exon skipping).

Lower panel: splicing pattern of SMN1 and SMN2 minigenes. The identity of the bands is indicated on the right of the gel. Hek293 were transfected with 0.5 µg of pCI SMN1 or pCI SMN2 plasmids. After 24 h total RNA was extracted and RT-PCR carried out using pCIF fw and E8-75 primers. The resulting amplified products were solved on 2% agarose gel.

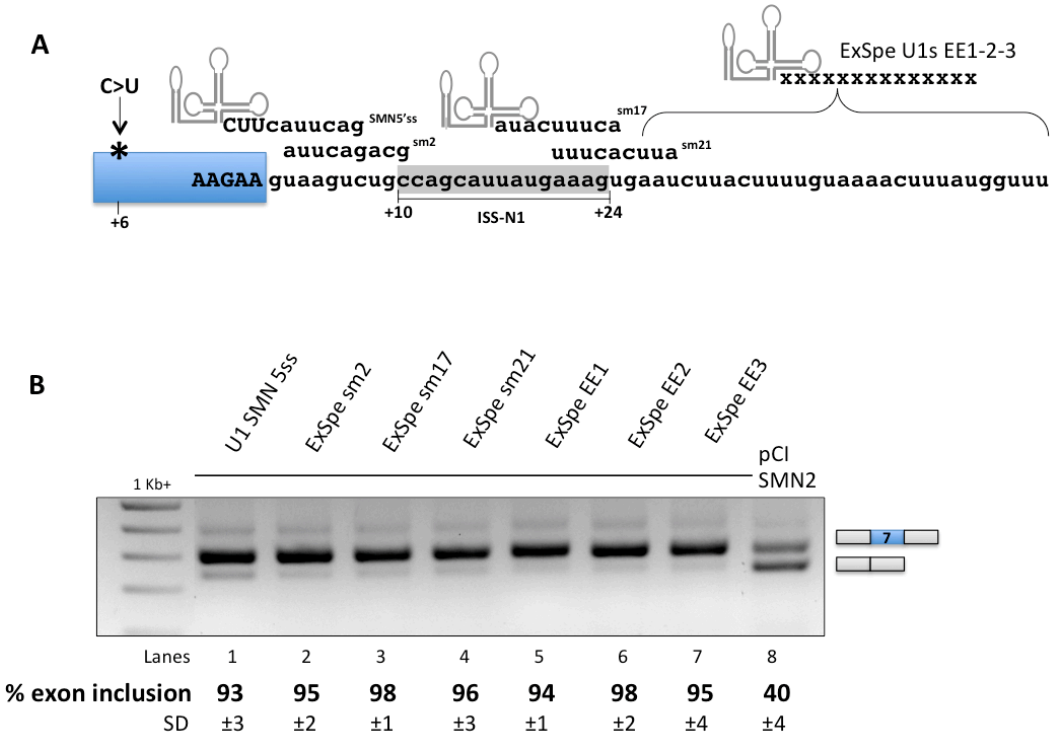


Figure 2.2. Screening of ExSpe U1s molecules in SMN2 exon 7 minigene

A. RNA sequence of exon 7 donor splice site and the downstream intronic flanking region. The exon is the blue box and its sequence is in capital letters. The intronic sequence is in bold. The gray box indicates the intronic splicing silencer N1 (ISS-N1) and its position relative to exon 7 5'ss is indicated (+10/+24). The sequences reported above the intron represent the nucleotides of the 5'-tail of modified U1 SMN 5'ss or different ExSpe U1s.

*The binding regions of ExSpe U1s EE1, EE2 and EE3 are not indicated because they are undergoing a patent registration process.

B. Screening of Exon Specific U1s on SMN2 exon 7 minigene. Hek293 cells were transfected with 0.5 µg of pCI SMN2 plasmid alone (lane 8) or in combination with 0.5 µg of plasmids encoding U1 SMN 5'ss fully complementary to exon 7 donor site (lane 1) or different ExSpe U1s (lanes 2-7). Total RNA was extracted and RT-PCR carried out as reported above. Sequencing of the upper bands showed the usage of the correct 5'ss.

2.1.2 Screening of exon-specific U1 molecules on SPINK5 exon 11 minigene model.

The analysis of keratinocytes from patients affected by netherton syndrome, a severe genodermatosis, showed a novel synonymous substitution c.891C>T in exon 11 of SPINK5 gene. The mutation causes complete skipping of exon 11 from the mature transcript in minigene splicing assays, reproducing the splicing pattern observed *in vivo* (Fortugno, Grosso et al. 2012, Lacroix, Lacaze-Buzy et al. 2012). I engineered several U1 snRNAs in order to evaluate in SPINK5 exon 11 +9C>T minigene splicing assays their capacity to promote exon inclusion. Also in this case, as first step, I decided to reinforce the capacity of the U1 snRNA to recognize the 5'ss of this exon: I created a U1 fully complementary to the non canonical exon 11 donor splice site (named U1 sk-3L), bearing compensatory mutations in position -2, -1 and +3 and I tested its splicing rescue efficacy in minigene splicing assay (figure 2.3, panel B). Cotransfection experiment in Hek 293 cells of SPINK5 exon 11 +9C>T minigene along with compensatory U1 sk-3L showed a rescue of exon 11 inclusion up to 75%, which is the level detected in SPINK5 wt minigene (figure 2.3, panel C, compare lanes 1 and 3). To explore the activity of ExSpe U1s in correcting the splicing defect caused by the synonymous substitution under analysis, I prepared 6 bioengineered exon specific U1s targeting pre-mRNA sequences downstream the 5'ss, spanning a region from position +6 to +30 relative to the exon-intron border (figure 2.3, panel B). To test their efficacy in rescuing aberrant exon 11 splicing, I cotransfected Hek293 cells with both wild type and mutant SPINK5 minigenes along with different ExSpe U1s. The analysis of the splicing pattern revealed that 5 out of 6 ExSpe U1s were able to modify, with different efficacy, the inclusion of exon 11 in the mature transcript (figure 2.3, panel C, lanes 4-9). In particular ExSpe sk12, sk15L and sk18 showed a remarkable rescue efficacy, inducing a percentage of exon inclusion of 90%, 84% and 86%, respectively (lanes 6, 8, 9); ExSpe sk9 and sk15 promoted exon inclusion up to 20% and 70% (lanes 5 and 7). ExSpe sk6, whose modified 5'-tail targets a longer pre-mRNA sequence in comparison to the other ExSpe sk U1s, did not show any effect on

the splicing pattern (lane 4). This could be related to the sequence composition of the modified tail of the ExSpe U1 RNA, which probably prevents the correct folding of the molecule due to the disruption of the secondary structure.

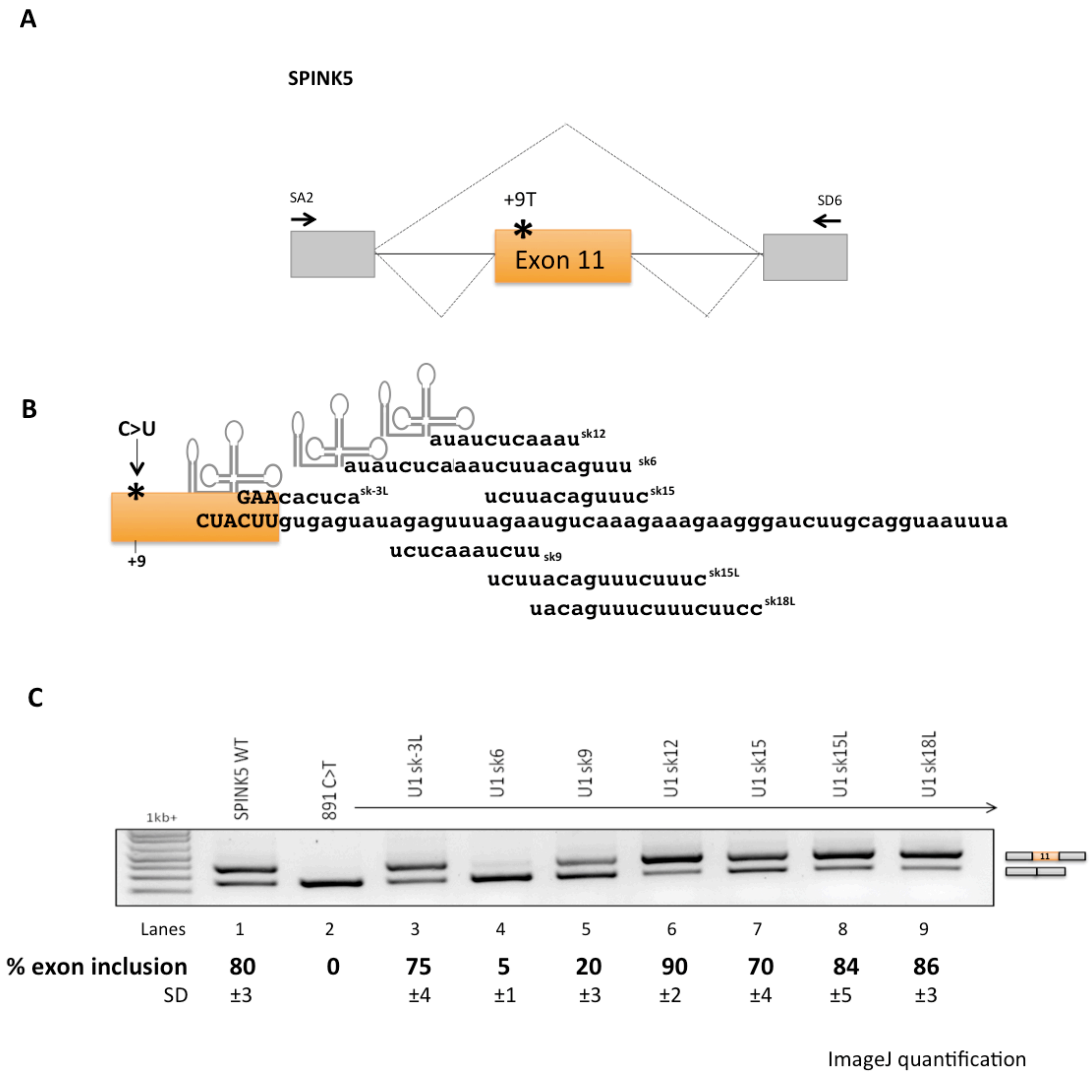


Figure 2.3. Screening of ExSpe U1s molecules on SPINK5 exon 11 minigene

A. Schematic representation of SPINK5 exon 11 minigene. The boxes are exons and the lines are introns. The asterisk shows the localization of the +9 C>T synonymous substitution. The two possible patterns of splicing (exon inclusion or exon skipping) are indicated with dotted

lines. The positions of the primers used in RT-PCR experiments are indicated.

B. Schematic representation of ExSpeU1s that bind to *SPINK5* sequences. The basepairing between the 5'-tail of the ExSpe U1s and the pre-mRNA sequence is reported.

C. Lines 1-2, Hek293 cells transfected with 0.5µg of *SPINK5* wild-type and *SPINK5* c.891C>T minigenes, respectively. Lines 3-9, Hek293 cells co-transfected with 0.5µg of mutant c.891C>T minigene and 0.5µg of ExSpeU1s. After 24 hours, total RNA was analyzed by RT-PCR, using *SD6* and *SA2* primers. The PCR products were separated on 2% agarose gel by electrophoresis. The upper band corresponds to *SPINK5* exon 11 inclusion and the lower band corresponds to skipping of *SPINK5* exon 11. ExSpeU1s sk12, sk15L and sk18L showed a high efficacy of splicing correction with a percentage of exon 11 inclusion higher than *SPINK5* wild-type (respectively of 90%, 84% and 86%). ExSpeU1 sk-3L showed a percentage of about 75%, ExSpeU1 sk15 showed a percentage of exon 11 inclusion >60%, whereas the ExSpeU1s sk6 and sk9 showed a lower efficacy of exon 11 inclusion (respectively, 4% and 20%). Sequencing of the upper bands showed the usage of the correct 5'ss.

2.2 Modulation of endogenous SMN transcripts in Hek293 cells.

To investigate the splicing rescue of the SMN-ExSpe U1s, I decided to evaluate the effect of ExSpeU1s on the endogenous pattern of splicing of SMN2 in Hek293 cells. ExSpe U1s were either transiently overexpressed or directly integrated in the same chromosomal region at a specific FRT site as a single gene copy. This last approach allowed us to evaluate if a unique ExSpe U1 gene copy is sufficient to affect SMN splicing.

2.2.1 Exon specific U1s improve endogenous SMN2 splicing pattern in Hek293 cells.

To evaluate the effect of ExSpe U1 on endogenous SMN2 exon 7 splicing transcript, I selected ExSpe sm17 and sm21 for transfection in Hek293 cells. As Hek293 cells possess both *SMN1* and *SMN2* genes, I took advantage of a single nucleotide substitution in the final part of SMN2 exon 8 that creates a restriction site for DdeI. This substitution allows specific digestion of SMN2 transcripts, while SMN1 is insensitive to the action of the enzyme. 48 hours after transfection of the ExSpeU1s, total RNA was extracted and RT-PCR carried out using a 5'-FAM-labelled Fw primer. The amplified cDNAs were digested using DdeI restriction enzyme to discriminate SMN1 or SMN2 transcripts and the resulting fragments analysed in denaturing capillary electrophoresis for the estimation of exon 7 splicing pattern. Evaluation of capillary electrophoresis output through PeakScanner software showed that ExSpe sm17 and ExSpe sm21 are able to modulate endogenous SMN2 transcripts, promoting exon 7 inclusion from 40% up to 75% in comparison to not transfected cells (figure 2.4). In addition, it is possible to notice a slight increase in the SMN1 exon 7 inclusion from 75% up to 85-90% in cells overexpressing ExSpe U1s sm17 or 21 in comparison to control cells (figure 2.4). Co-transfection of U1 WT did not affect the splicing pattern.

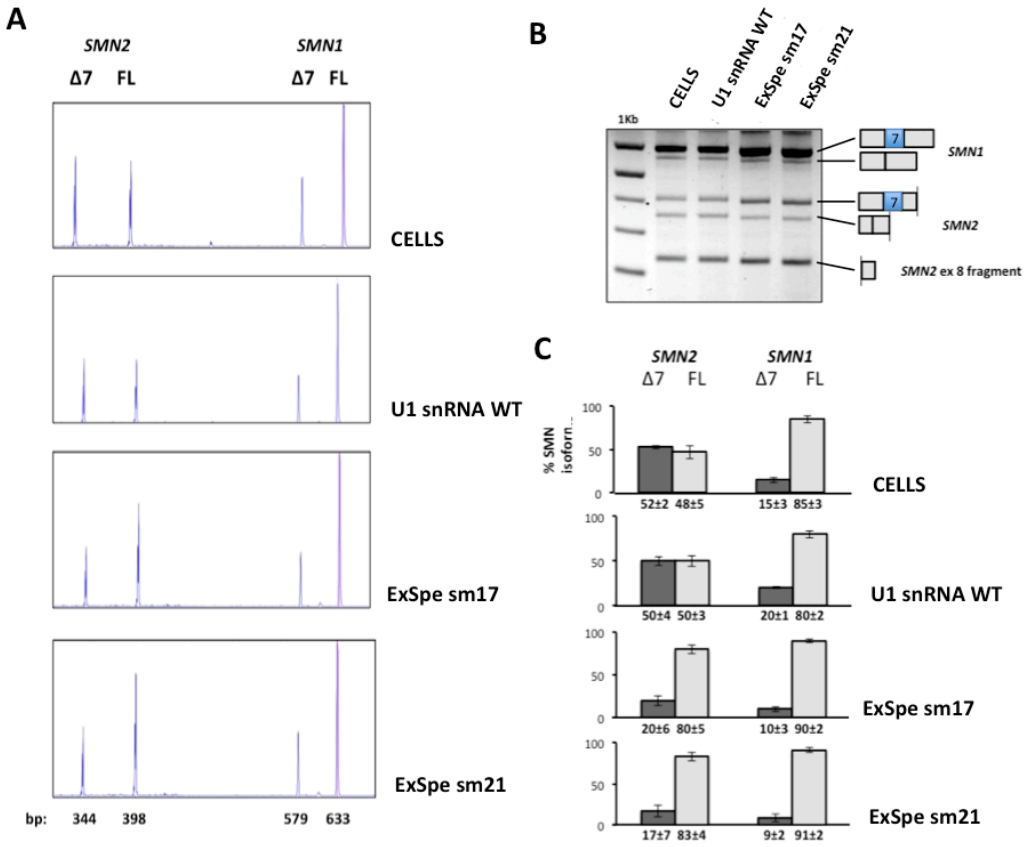


Figure 2.4. Episomal expression of ExSpe U1 sm17 and sm21 modulates SMN2 endogenous transcripts in Hek293

A. Rapresentative experiment in which the fluorescently labelled RT–PCR amplified fragments were separated on denaturing capillary electrophoresis. HEK393 cells were transfected with the indicated ExSpeU1s and amplified fragments were digested with *Dde*I to obtain SMN1 and SMN2 exon 7 inclusion (FL) and exclusion (Δ7) fragments. The size of the fragments is indicated.

B. RT-PCR amplified fragments from (A) were solved in 2% agarose gels after DdeI digestion. The identity of the bands is indicated on the right of the gel.

C. Percentage of SMN1 and SMN2 exon 7 inclusion in ExSpeU1-treated and -untreated cells. The percentage of each isoform was calculated using PeakScanner software. Data are expressed as means ± SD of three

independent experiments done in duplicate. Transfection efficiency in the different experiments was estimated through cotransfection of a GFP encoding plasmid. The percentage of GFP positive cells at confocal microscope was assessed to be 75–85%.

2.2.2 Comparison between ExSpe U1s- or AON-mediated SMN2 splicing rescue in Hek293 Flp-In cells.

To evaluate the minimum number of copies of ExSpe U1s gene necessary to mimic the corrective effect on SMN2 exon 7, I decided to create ExSpe U1s Hek293 stable clones. I took advantage of Hek293 Flp-In cell system that allows the chromosomal insertion of one copy of a gene of interest at a specific FRT site in the cell genome. Through hygromycin treatment, I selected Hek293 Flp-In positive clones expressing one copy of ExSpe sm2, sm17 and sm21, as well one additional copy of U1 snRNA wt as control. Specific RT-PCR confirmed the expression of ExSpe U1s in the corresponding Hek293 Flp-In clones.

I analysed the ExSpe U1s mediated SMN2 exon 7 splicing modulation using two different methods. The first consists of semi-quantitative RT-PCR followed by capillary electrophoresis, and with this method I evaluated the splicing ratio between SMN2 full length (FL) and the isoform lacking exon 7 (Δ7). The second is a TaqMan-based quantitative RT-PCR that was used to evaluate the total amount of each FL or Δ7 mRNAs relative to an endogenous control. In addition, I also compared the effect of ExSpe U1 with the effect of the antisense oligonucleotide 10-27 (AON 10-27). This antisense oligonucleotide 10-27 masks the hnRNPA1-dependent intronic splicing silencer N1 (ISS-N1) that is located from position +10 to +24 downstream of exon 7 donor splice site. This oligo has been extensively reported to improve SMN2 exon 7 inclusion in different SMA cellular and mouse models (Hua, Vickers et al. 2007, Hua, Vickers et al. 2008, Hua, Sahashi et al. 2010, Hua, Sahashi et al. 2011). Thus, I decided to make a comparison between the effect

of AON 10-27 and the chromosome integrated copy of ExSpe U1s on the endogenous Hek293 Flp-In SMN transcripts.

Total cellular RNA was extracted from ExSpeU1s stable clones, Hek293 Flp-In cells transfected with AON 10-27 and control cells, retrotranscribed and SMN messenger RNAs amplified using a 5'-FAM-labelled forward primer. The resulting products were digested with DdeI restriction enzyme to distinguish SMN1 or SMN2 deriving transcripts. The fragments were loaded on denaturing capillary electrophoresis to quantify the percentage of FL and Δ7 isoforms (figure 2.5, panel A). The analysis of the capillary electrophoresis output showed an increase in the percentage of SMN2 exon 7 inclusion from ~50% of Hek293 Flp-In up to ~80-85% in the stable clones expressing one copy of ExSpe sm2, sm17, sm21, as well as in Hek293 Flp-In transfected with antisense oligonucleotide 10-27 (figure 2.5, panel B). As expected, one chromosome-integrated copy of U1 snRNA wt gene did not exert any splicing modulation in SMN endogenous mRNAs. Interestingly, in our experimental conditions, the ExSpe U1s and the AON 10-27 modulated SMN2 exon 7 inclusion with a similar efficacy.

Giving that an increase in the percentage of exon 7 inclusion could potentially reflect only a decrease in Δ7 isoform without a real increase of the FL transcript, I performed a quantitative RT-PCR assay. Two different TaqMan probes and set of primers were used to detect the amount of FL or Δ7 mRNAs in control cells, in Hek293 Flp-In ExSpe U1s stable clones and in Hek293 Flp-In transfected with antisense oligonucleotide 10-27 (AON 10-27). As shown in figure 2.6, stable clones expressing one copy of ExSpe sm2, sm17, sm21, EE2, EE3 showed up to ~3 fold increase in the FL isoform in comparison to control cells. A similar increase in the amount of the FL isoform is present also in AON 10-27 treated cells. Cells expressing one additional copy of U1 snRNA wt gene did not show any change in the amount of FL or Δ7 transcripts in comparison to normal cells. Interestingly, all ExSpe U1s stable clones showed an increase also in the Δ7 isoform, although to a lesser extent in comparison to FL isoform. In fact, relative to control cells, ExSpeU1s increased to ~ 1.6 fold the amount of Δ7 isoform. On the contrary, cells

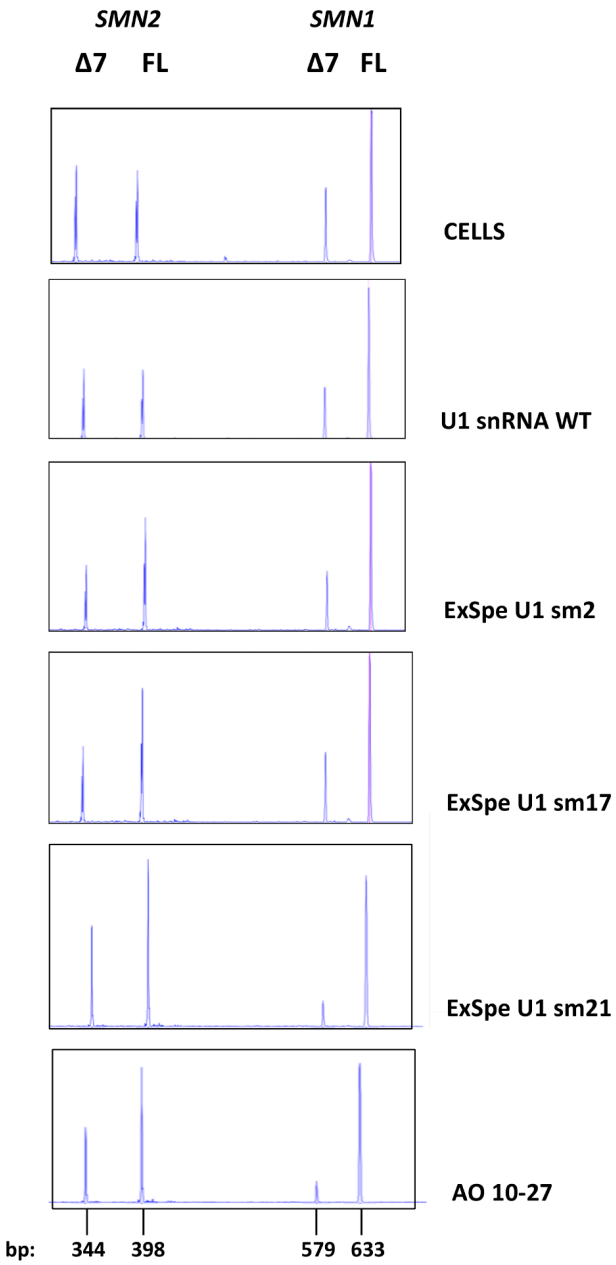
treated with AON 10-27 have a lower amount of Δ7 isoform in comparison to control cells. In this case, the Δ7 isoform was reduced ~three times in comparison to the untreated cells.

Taken together, these data demonstrate that one chromosome-integrated copy of different ExSpe U1s targeting pre-mRNA regions downstream of the exon 7 5'-splice site and AON 10-27 improve the percentage of exon 7 inclusion, increasing the total amount of full length mRNAs. However, ExSpe U1s and AON 10-27 have a different effect on the expression of the Δ7 isoform: while AON 10-27 reduces the total amount of Δ7, ExSpe U1s promote an increase of this isoform. Thus, while the AON-mediated masking of the hnRNPA1-dependent ISS mainly corrects the splicing decision, ExSpeU1s may have an additional effect on pre-mRNA processing. To follow this observation, I analysed the total amount of SMN transcripts produced using primers located upstream the alternatively spliced exon 7 (figure 2.7). Stable Hek293 Flp-In clones expressing one copy of ExSpeU1s showed a significant increase in SMN mature transcripts. ExSpeU1s induced a ~3 fold increase in SMN transcripts while AON 10-27 had no effect. Stable clones expressing U1snRNA wt or an unrelated gene did not show any change (figure 2.7). All together these data suggest that ExSpeU1s and AON 10-27 correct SMN exon 7 splicing with different molecular mechanisms.

In addition, although normal Hek293 cells produce high levels of SMN protein due to the presence of both *SMN1* and *SMN2* genes, I tried to detect changes in the SMN protein expression in Hek293 clones stably expressing ExSpeU1s or in cells trated with the oligonucleotide. Interestingly, western blot analysis revealed a small but reproducible effect (~1.3-1.5 fold increase) in the level of SMN protein in Hek293 Flp-In cells expressing ExSpe U1s and in cells treated with the oligonucleotide (figure 2.8).

Expression of the control U1 snRNA WT gene in Hek293 Flp-In cells does not affect SMN splicing (figure 2.5), SMN transcripts levels (figure 2.6 and 2.7) nor SMN protein (figure 2.8).

A


SMN2
SMN1
Δ7
FL
Δ7
FL
CELLS
U1 snRNA WT
ExSpe U1 sm2
ExSpe U1 sm17
ExSpe U1 sm21
AO 10-27
bp:
344
398
579
633

B

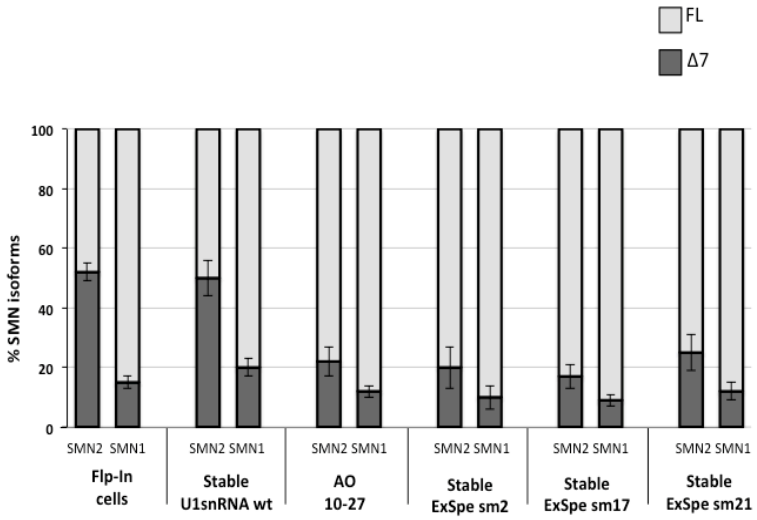


Figure 2.5. SMN1 and SMN2 splicing pattern in Hek293 stable clones.

A. Total RNA was extracted from Hek293 Flp-In cells stably expressing one copy of U1 snRNA WT or SMA-specific ExSpe U1s sm2, sm17 or sm21 or cells transfected with AO 10-27. RT-PCR was performed with FAM-labelled-E6-Fw and E8-467-Rev primers. The corresponding amplified products were then digested with DdeI restriction enzyme in order to discriminate fragments coming from *SMN1* or *SMN2* genes. The resulting products were run on capillary electrophoresis. Each peak reported in the graph represent an isoform and the area surrounding the peaks gives a relative quantification of the corresponding isoforms. The identity of each peak and its provenience (SMN2 or SMN1 is indicated on the top of the panel. The size is indicated below.

B. Graphic representation of capillary electrophoresis output after analysis with PeakScanner Software. Quantification was performed analysing the area under the peaks.

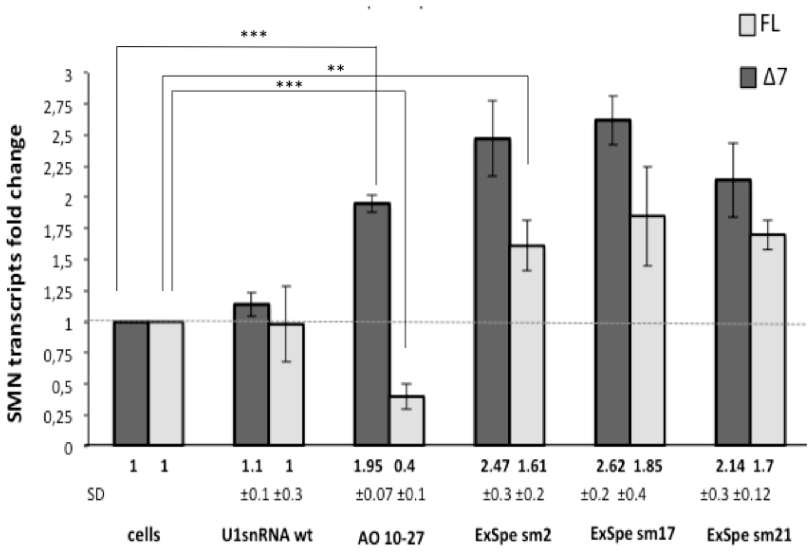


Figure 2.6. Quantitative RT-PCR on SMN transcripts in Hek293-ExSpe U1 stable clones.

Quantitative RT-PCR to detect SMN specific isoforms. The same RNA samples described in figure 2.6.A were retro-transcribed and two TaqMan probes for SMN full length isoform and lacking exon 7 (Δ7) were used in combination with FL-fw and FL-rev and Δ7-fw and Δ7-rev set of primers respectively. TaqMan probe for 18S was used as normalization control. The fold change in SMN isoforms is expressed as ratio between each isoform (FL or Δ7) level in normal cell and in cell stably expressing U1snRNA wt or ExSpe sm2, sm17 or sm21. The data are expressed as mean ±SD of three independent experiments (***p-value<0,001; T-test, paired, two tails).

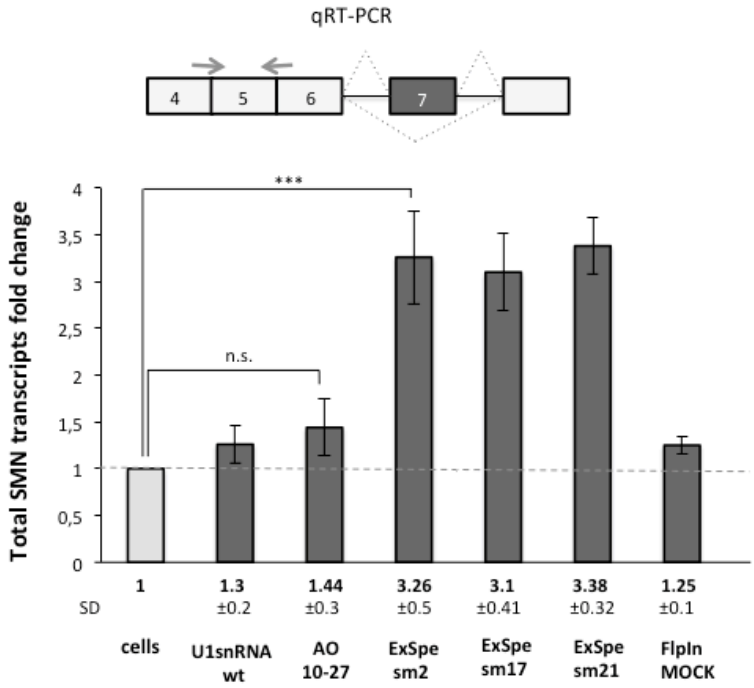


Figure 2.7. Evaluation of total SMN transcripts in Hek293-ExSpe U1 stable clones.

Quantitative RT-PCR on total SMN cellular transcripts. Total RNA was extracted and a quantitative syber green RT-PCR carried out with ex4/5Fw and ex5/6Rev primers as indicated in the top of the panel. TaqMan probe for 18S was used as normalization control. The fold change in SMN transcripts level is expressed as ratio between each transcripts levels in normal cells and cells transfected with AON 10-27 or cells stably expressing ExSpe U1s and U1snRNA wt or clones expressing an unrelated gene (mock) as control. The results are expressed as mean ±SD of three different experiments done in triplicate (***p-value<0,001; T-test, paired, two tails).

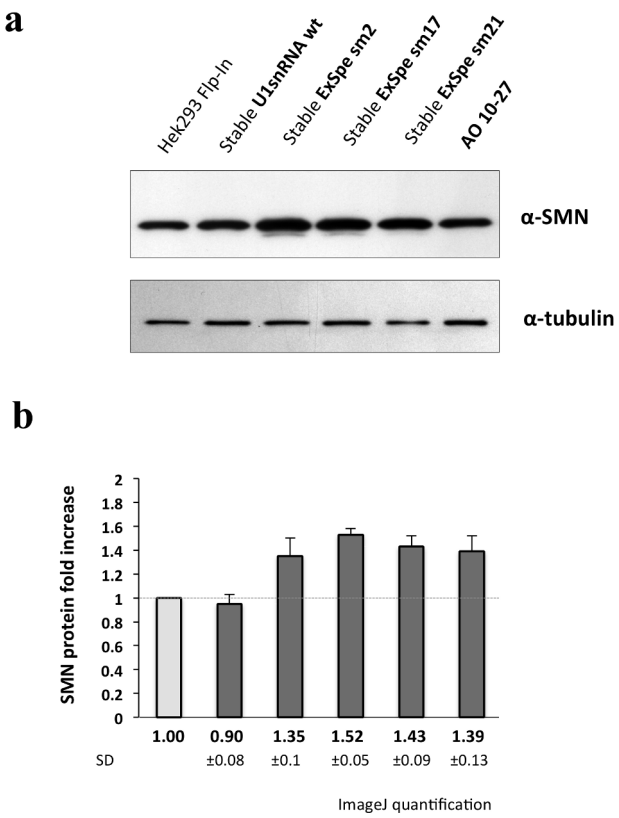


Figure 2.8. SMN protein levels in Hek293 Flp-In cells expressing ExSpe U1s.

Protein lysate from Hek293 Flp-In stable clones expressing one copy of ExSpe U1 sm 2, sm17, U1snRNA wt or cells treated with the AON 10-27 were loaded on 4-12% polyacrylamide gel.

A. Western blot of SMN protein using anti-SMN monoclonal antibody. Tubulin was used as normalization control.

B. Intensity of SMN protein bands was estimated through ImageJ Quantification program. The histograms indicate the SMN protein fold increase expressed as mean ±SD of three independent experiments.

2.3 ExSpe U1s mediated correction of SMN2 exon 7 skipping in SMA fibroblasts and SMA iPS cells.

I evaluated the therapeutic efficacy of ExSpe U1 mediated splicing correction for SMN2 exon 7 aberrant splicing in two different SMA cellular models. SMA type I fibroblasts (G3813, Coriell Cell Repositories) derived from a 3 years old caucasian male, whose clinical symptoms were marked muscular atrophy and weakness. Genetic investigations of patient genome revealed that he possessed no functional copies of SMN1 and retains two copies of the paralogue SMN2 gene. As control I used normal fibroblasts whose donor subject was the mother of the patient reported above (G3814, Coriell Cell Repositories). She did not show symptoms of spinal muscular atrophy, was clinically normal and had both SMN1 and SMN2 genes. The second cellular model I used to test the efficacy of SMN-specific ExSpe U1s were inducible pluripotent stem cells from a SMA type I patient (SMA- iPSCs) and control cells from healthy person (normal- iPSCs) (courtesy of dr. S. Corti, University of Milan, Italy). In both the models under analysis I evaluated the capacity of selected ExSpe U1s to rescue SMN full length isoform and the corresponding functional SMN protein.

2.3.1 Lentiviral transduction of ExSpe U1s rescues SMN2 exon 7 splicing in SMA type I fibroblasts and increases the cellular level of correct SMN full length mRNA.

As showed in chapters 2.1.1 and 2.2, SMN-specific ExSpe U1s efficiently promote SMN2 exon 7 inclusion in minigene splicing assays and when they are inserted in a chromosome of Hek293 Flp-In cells. To verify the therapeutic potential of ExSpe U1s strategy in affected cells, I selected three ExSpe U1s –ExSpe sm2, sm17 and sm21 that were tested on SMA type I patient fibroblasts (3813, Coriell Cell Repositories). These cells represent a useful cellular model of the disease and have been extensively used for testing novel therapeutic agents. Because of the low transfection efficiency in primary fibroblasts with conventional methods, I prepared lentiviral vectors. The lentiviral constructs were designed to express the non-coding ExSpe U1

sm2, ExSpe U1 sm17 and ExSpe U1 sm21 or control U1 snRNA WT along with the green fluorescent protein (GFP) to assess cell transduction efficiency (see material and methods). I infected SMA type I primary fibroblasts and normal control fibroblasts (Coriell Institute, G3814) with lentiviral particles ($4x10^6$ bTU/mL) with a molteplicity of infection (MOI) of 15. 72 h after infection, the cells were harvested and transduction efficiency was determined through GFP-positive cells analysis at confocal microscope. Transduction efficiency was 85-90%. I evaluated the rescue efficacy at the RNA level through the analysis of the SMN2 exon 7 splicing pattern using a semiquantitative RT-PCR assay and through the measurement of the absolute amount of the FL and Δ7 SMN mRNAs using a specific TaqMAN assay. Moreover, I checked the expression of ExSpe U1s sm17 and sm21 through a specific RT-PCR (figure 2.11). In this assay, I could not evaluate the ExSpe sm2 expression for technical limitation. In fact the sm2 sequence present in the ExSpeU1 is very similar to the sequence of endogenous U1 snRNA wt and the two molecules could not be discriminated by RT-PCR. Analysis of the expression of ExSpe U1s sm17 and sm21 in lentiviral infected fibroblasts showed expression of the corresponding RNAs (figure 2.11). Semiquantitative analysis RT-PCR in capillary electrophoresis of lentiviral infected SMA fibroblasts showed that the three ExSpeU1s increased the percentage of exon 7 inclusion from ~38% up to ~80%, while U1 snRNA WT had no effect (figure 2.9). As the increase in the percentage of exon 7 inclusion in the semiquantitative RT-PCR assay may potentially reflect only reduced exon 7 skipping, without the real increase of the correct FL transcript, I investigated if the splicing correction leads also to a significant increase of the amount of FL and Δ7 SMN mRNAs. Quantitative evaluation of SMN mRNAs showed that the three ExSpeU1s led to a significant ~2.5 fold increase in the amount of FL mRNA in comparison to untreated cells (figure 2.10). The amount of FL isoform in SMA specific ExSpe U1s-transduced fibroblasts is comparable to the level detected in SMA-unaffected control cells. Overexpression of control U1 snRNA WT did not alter the levels of SMN endogenous transcripts (figure 2.10). Interestingly, as previously reported for Hek Flp-In ExSpe U1s stable clones, SMA type I

fibroblasts treated with ExSpe U1s showed a small increase in the isoform lacking exon 7. Compared to untreated cells the Δ7 isoform increased ~ 1.7 fold. Interestingly the level of Δ7 reached in the ExSpe U1s treated cells is similar to the level found in normal fibroblasts in which Δ7 isoform derives from processing both defective SMN2 genes and, in part, from SMN1 genes.

Taken together, these results demonstrate that in SMA type I fibroblasts, ExSpeU1s sm2, sm17 and sm21 rescue SMN2 splicing pattern and increase the amount of the correct FL and Δ7 transcript to a level present in normal control fibroblasts.

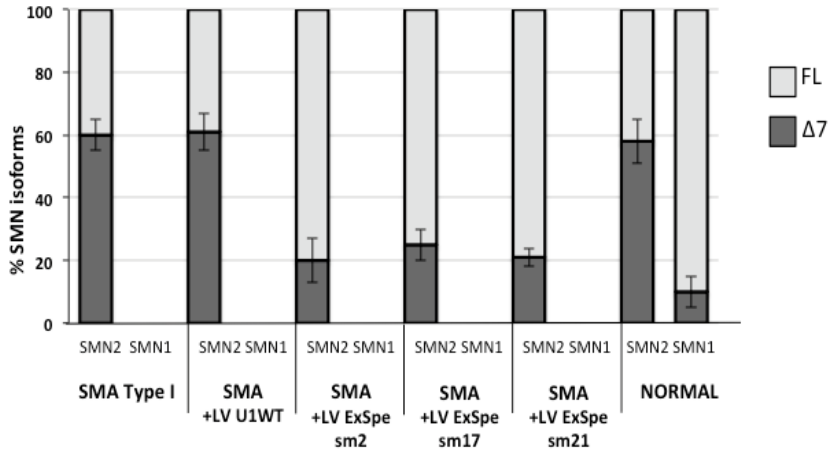


Figure2.9. SMN exon 7 splicing rescue in SMA type I fibroblasts

SMA type I fibroblasts (Coriell Institute, 3813) were transduced with 4.5 x 10^5 lentiviral particles with a molteplicity of infection (MOI) of 15. Normal fibroblasts (Coriell Institute, 3814) were used as control. Transduction efficiency was evaluated through GFP positive cells at confocal microscope. 72h after transduction total RNA was collected and RT-PCR carried out with E6-F 5'-FAM-labelled and E8-467R primers. Amplification products were then digested with DdeI restriction enzyme and ran on denaturing capillary electrophoresis. The percentage of SMN1 or SMN2 spliced isoforms was calculated using Peak Scanner software. The bars represent the quantification expressed as mean +/- SD of three different experiments.

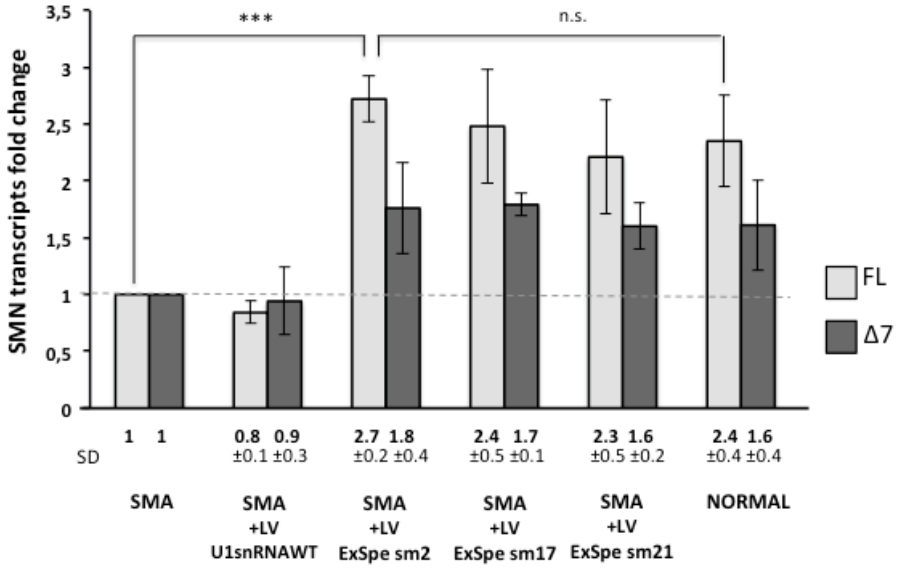


Figure 2.10. Quantitative analysis of full length and delta-7 transcripts in SMA type I fibroblasts treated with ExSpe U1s.

The same RNA samples obtained from SMA fibroblasts (treated or not) and normal control fibroblasts described above were used for quantitative RT-PCR analysis. Two different TaqMan probe for SMN full length isoform and lacking exon 7 (Δ7) were used in combination with FL-fw and FL-rev and Δ7-fw and Δ7-rev set of primers. TaqMan for 18S was used as normalization control. The fold change in SMN transcript is expressed as ratio between each isoform (FL or Δ7) level in SMA not treated fibroblasts and SMA fibroblasts treated with U1snRNA wt or ExSpe U1s sm2, sm17 or sm21 or normal control fibroblasts. The results are expressed as mean ±SD of three different experiments and *p*-values are also calculated (*** *p*-value<0.001, n.s not significant; T-test).

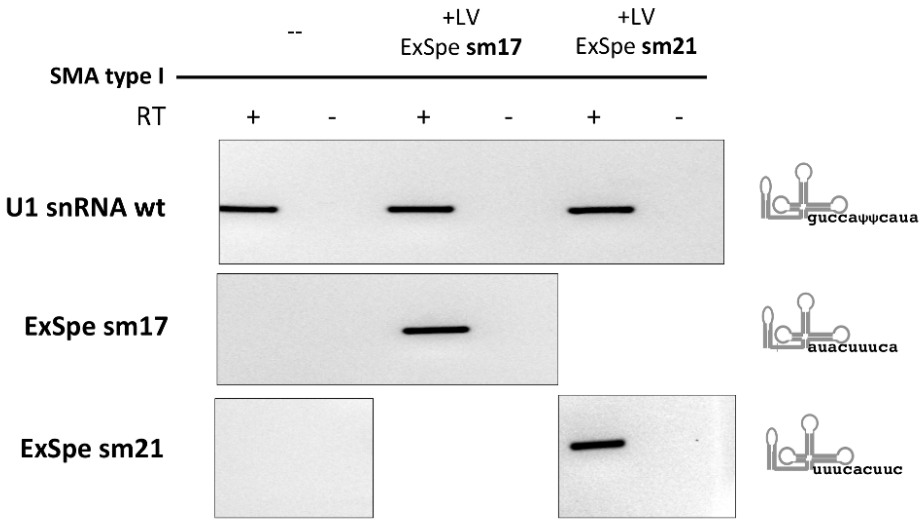


Figure 2.11. Expression of ExSpe sm17 and sm21 in SMA type I fibroblasts after lentiviral transduction.

Total RNA was extracted from SMA type I fibroblasts treated with lentiviral particles expressing ExSpe U1 sm 17 or sm 21 and treated with DNase to avoid DNA contamination. RT-PCR has been performed +/- retrotranscriptase (+/-RT) using set of primers to detect the endogenous U1 snRNA, ExSpe U1 sm17 or ExSpe U1 sm21.

2.3.2 Analysis of SMN protein abundance in SMA type I fibroblasts treated with lentiviral particles expressing ExSpe U1s.

Lentiviral transduction of ExSpe sm2, sm17 and sm21 efficiently rescue exon 7 aberrant splicing, promoting the increase of the full length SMN mRNAs. To investigate if the correction at the RNA level results in a corresponding increase in the cellular amount of SMN protein, I evaluated the SMN protein levels in SMA type I patient fibroblasts after ExSpe U1s lentiviral-mediated transduction, as well as the level of protein present in normal control fibroblasts. The SMN protein in SMA fibroblasts lentivirally transduced with ExSpe U1s sm2, sm17, sm21 or U1 snRNA WT was evaluated through western blot using a mouse anti-SMN monoclonal antibody. Quantification of SMN bands intensity showed that ExSpe U1s sm2, sm17 and sm21 promoted a ~1.65, 1.8 and 1.9 fold increase in the SMN functional full length protein in comparison to untreated SMA fibroblasts. Interestingly, the protein rescue obtained with ExSpe U1s reaches the physiological SMN level present in SMA-unaffected control fibroblasts that produce ~2 fold more protein in comparison to SMA patient cells, as detected by western blotting (figure 2.12).

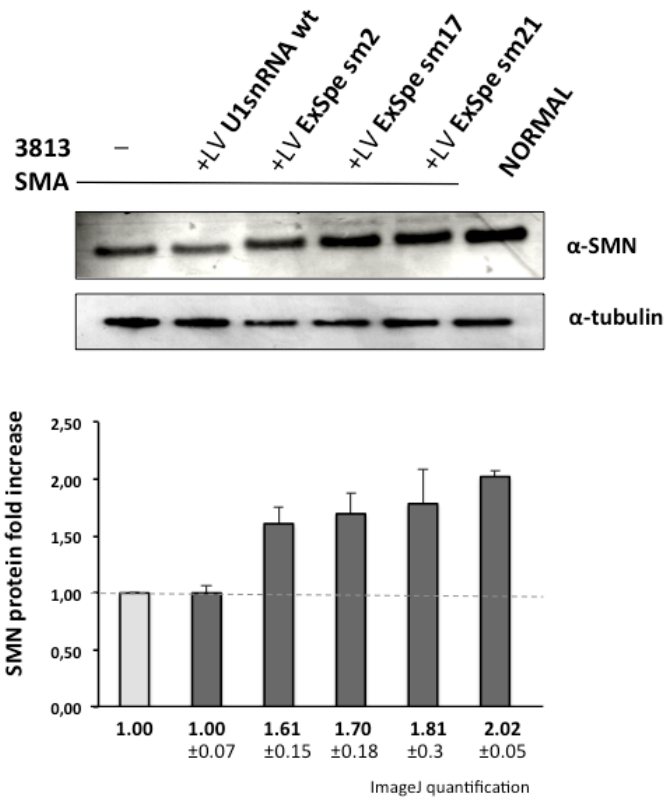


Figure 2.12. Rescue of SMN protein level in SMA type I fibroblasts treated with ExSpe U1s.

A. SMA type I fibroblasts (Coriell Institute, 3813) were transduced with 4.5 x 10^5 lentiviral particles with a molteplicity of infection (MOI) of 15. Normal fibroblasts (Coriell Institute, 3814) were used as control. Transduction efficiency was evaluated through GFP positive cells at confocal microscope. 72h after transduction cells were harvested and protein samples loaded on 4-12% polyacrylamide gel. Upper panel shows SMN protein levels in SMA patient fibroblasts (lane 1) and normal control fibroblasts (lane 6). SMN protein in SMA fibroblasts treated with U1snRNA wt or with ExSpe U1s sm2, sm17 and sm21 are shown in lanes 2 and 3-5 respectively. The lower panel show the tubulin levels, used as loading and normalization control.

B. The intensity of SMN protein bands was estimated through ImageJ Quantification program. The histograms indicate the SMN protein fold increase expressed as mean ±SD of three different experiments.

2.4 Lentiviral-mediated correction of SPINK5 exon 11 skipping in Netherton Syndrome primary cells.

The screening for ExSpe U1s active in rescuing SPINK5 exon 11 skipping due to c.891C>T mutation in minigene model showed the particular efficacy of the ExSpe U1 sk12. Thus, I created a pLVTHM vector (Trono 2001) expressing ExSpe U1 sk12 along with GFP and I produced lentiviral particles to transduce netherton syndrome (NS) primary keratinocytes. The presence of the GFP allowed the determination of transfection efficiency during the production of the virus in Hek293 cells and also the evaluation of viral transduction efficiency in primary cells using a confocal microscope. The primary NS keratinocytes used in these experiments derive from a compound heterozygote patient that carry two different mutations: one allele contains a premature termination codon (PTC) in exon 13 (c.1111C>T, p.Arg371X) and the other carries the c.891C>T genomic variant that was shown to cause total exon 11 skipping (Fortugno, Grosso et al. 2012). Exon 11 skipping leads to the change of the phase and the creation of a premature termination codon (PTC) (UGA triplet) within exon 12. The two transcripts containing PTCs coming from the two alleles are partially subjected to degradation through the nonsense-mediated pathway (NMD). I transduced primary NS keratinocytes with lentiviral particles expressing ExSpe U1 sk12 along with GFP with a molteplicity of infection (MOI) of 50. Total RNA was extracted and semi-quantitative RT-PCR carried out. The resulting amplified products were solved on 2% agarose gel and identity of bands was verified by direct sequencing. NS keratinocytes show ~45% of exon 11 inclusion (figure 2.14 A, lane 2). Direct sequencing of the upper band (*) revealed that the full length transcript entirely derives from the allele carrying the stop codon (see figure 2.15 B, upper panel). In NS keratinocytes treated with lentiviral particles expressing ExSpe sk12, the percentage of exon 11 inclusion increased from 45% up to 75-80% (figure 2.14 A, lane 1). Interestingly, in this case the sequence of the band that correspond to exon inclusion (**) showed that most of the full length mRNA derives from the mutant C>T allele

(see figure 2.14 B, lower panel). The pattern of splicing of normal keratinocytes is shown in figure 2.14 A, lane 3.

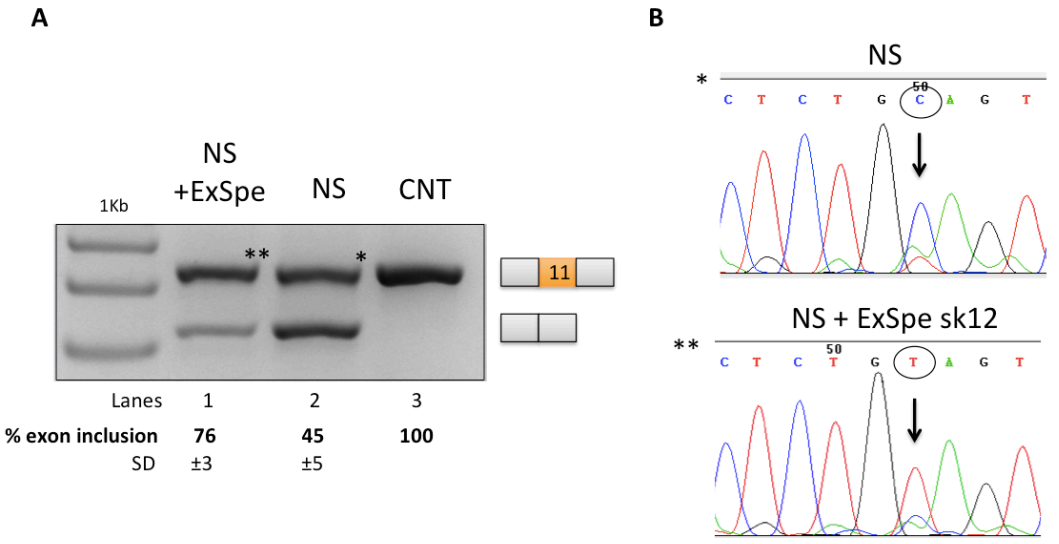


Figure 2.14. Rescue of full length correct SPINK5 mRNA with lentiviral particles expressing ExSpe U1 sk12.

A. Primary keratinocytes from a compound heterozygous netherton syndrome patient carrying the c.891C>T mutation on the first allele and a mutation creating a stop codon in the second allele (lane 2) were transduced with lentiviral particles expressing ExSpe U1 sk12 with a molteplicity of infection (MOI) of 50 (lane 1). Normal keratinocytes were used as control (lane 3). RT-PCR was performed using F1 fw and F2 rev primers and the resulting amplified products separated on 2% agarose gel. The upper band is exon 11 inclusion, the lower band exon exclusion.

B. The bands corresponding to exon 11 inclusion have been sequenced. In case of NS cells the transcripts that contain exon 11 (*) comes from the allele carrying the mutation that introduce the stop codon as evidenced by the presence of the C nucleotide in the sequence (upper panel). The sequence of the exon inclusion band in keratinocytes lentivirally transduced with ExSpe U1 sk12 (**) showed the presence of the T nucleotide, indicating that the amplified product comes from the mutated c.891C>T allele.

2.4.1 ExSpe U1 rescue correct full-length LEKTI protein in NS keratinocytes in a dose-dependent manner.

To verify if the splicing rescue mediated by the ExSpe U1 sk12 affects the corresponding LEKTI protein levels, I decided to transduce NS primary keratinocytes (NSK-2980) with lentiviral particles expressing the ExSpe U1 sk12 along with green fluorescent protein (GFP). I infected the cells with three different molteplicity of infection (MOI=25, 50, 100) in order to check a possible dose-response effect of the ExSpe U1 expression in rescuing full length LEKTI protein. As controls, I transduced netherton syndrome cells also with lentiviral particle expressing U1 snRNA WT along with GFP or GFP alone with a MOI of 25 and 50, respectively.

To verify the correlation between the different MOI used to transduce NS keratinocytes and the effective expression of the ExSpe U1, I performed RT-PCRs to amplify ExSpe U1 sk12 and GFP RNAs. GAPDH RNA was used as normalization control (Figure 2.15). Primers used for ExSpe U1 sk12 detection did not amplify the endogenous U1 snRNA and, to check for DNA contamination, the experiment was performed also in absence of retrotranscriptase (+/- RT). As expected, amplification of ExSpe U1 sk12 transcript is visible only in the samples treated with lentiviral particle expressing this ExSpe U1. The intensity of the bands, relative to GAPDH, increases with the molteplicity of infection (first panel, lanes 2-4). As supplementary control, I amplified the GFP transcript. The intensity of GFP amplified product reflects the different MOI used for cells transduction as observed for the ExSpe U1 expression (figure 2.15, third panel, lanes 2-4).

I performed a western blot using an anti-LEKTI monoclonal antibody that recognizes the functional full length LEKTI protein, whose molecular weight is 68 kDa, as well as shorter isoforms of the protein (30 and 37 kDa). These two isoforms derive from physiological proteolytic cleavages of a common precursor of 145 kDa, they are not functional and no known biological function is reported for them (Fortugno, Grosso et al. 2012, Diociaiuti, Castiglia et al. 2013). GAPDH was used as normalization control.

In normal keratinocytes immunoblot analysis detected the presence of three LEKTI isoforms (68 kDa, 37 kDa and 30 kDa), deriving from furin-mediated proteolytic cleavages of a 145 kDa precursor at specific linker regions between the functional domains (figure 2.16, lane 7). In NS keratinocytes western blot analysis detected only the 37 kDa LEKTI isoform (lane 1)., as previously reported (Lacroix, Lacaze-Buzy et al. 2012)

Cells from NS affected patients treated with lentiviral particle expressing ExSpe U1 sk12 showed an increase in the amount of FL functional LEKTI protein proportional to the molteplicity of infection used (figure 2.16, lanes 2-4). In fact, cells treated with ExSpe U1 sk12 with a MOI of 100 or 50 showed a remarkable increase in the correct full length 68 kDa protein in comparison to not treated cells (figure 2.16, compare lanes 1 and 2-3). Decreasing the dosage of lentivirus at MOI=25, it is possible to appreciate a corresponding reduced efficacy in the rescue of the correct LEKTI protein (lane 4). Lentiviral particle expressing GFP-U1 snRNA WT or GFP alone did not increase the amount of 68 kDa LEKTI protein (lane 5-6).

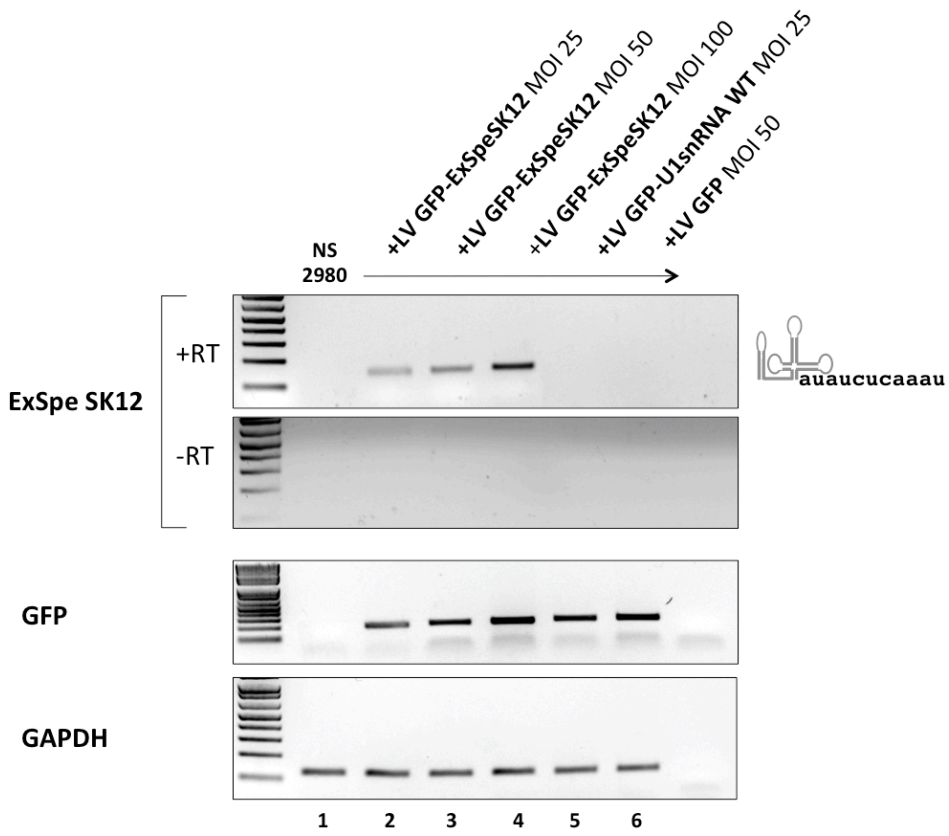


Figure 2.15. Expression of exon specific U1 sk12 in lentivirally transduced netherton syndrome keratinocytes.

Semiquantitative analysis of ExSpe U1 expression. Total RNA was extracted from the netherton syndrome primary keratinocytes treated with lentiviral particles expressing ExSpe U1 sk12 or U1snRNA along with GFP or GFP alone. RT-PCR was carried out with sk12Fw and U1160Rev primers to specifically amplify the ExSpe sk12 RNA (first panel). To avoid DNA contamination, the RT-PCR was performed also in absence of retrotranscriptase (second panel). Amplification of GFP transcript has been carried out using eGFPFw and eGFPRev primers. Amplification of GAPDH mRNA has been used as normalization control for semiquantitative comparation between different samples.

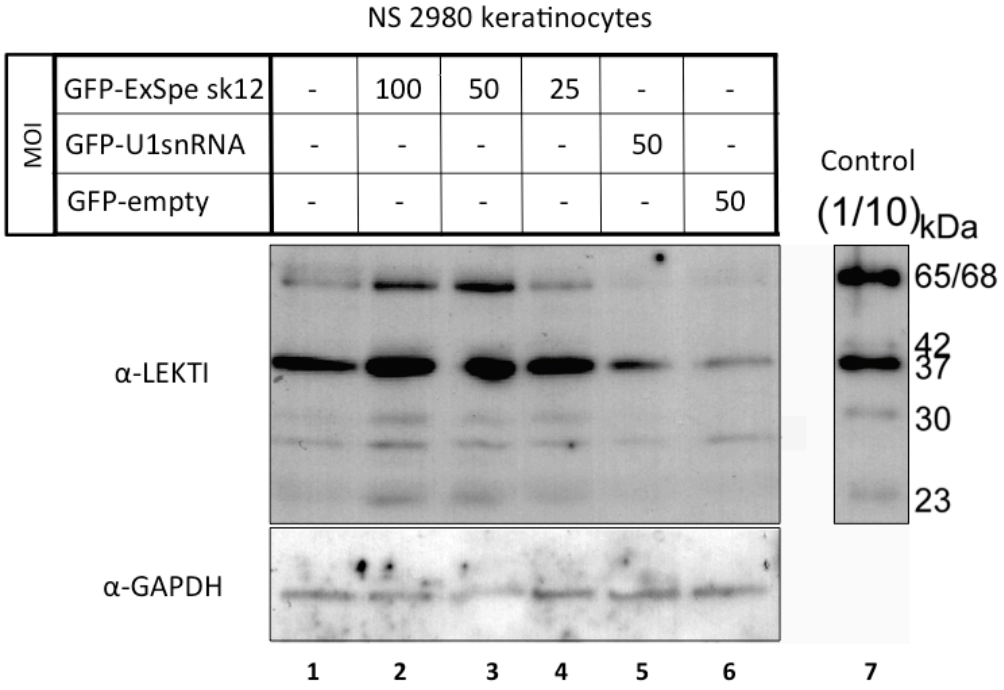


Figure 2.16. Exon Specific U1 sk12 rescue LEKTI protein level in netherton syndrome primary keratinocytes in a dose dependent manner.

Netherton syndrome primary keratinocytes (NS 2980) carrying the +9C>T synonymous substitution in SPINK5 exon 11 were transduced with lentiviral particle expressing ExSpe U1 sk12 or U1 snRNA WT along with the green fluorescent protein (GFP) or the GFP alone. Normal keratinocytes were used as control. NS 2980 were transduced with ExSpe U1 sk12 with different molteplicity of infection (MOI), as indicated in the panel above the gels. Total lysate from the cells were collected 72 hours post infection and loaded on 4-12% SDS gel. Anti-LEKTI primary monoclonal antibody was used and the signal detected with a secondary anti-mouse HRP-conjugated antibody.LEKTI functional protein corresponds to the band of 68 kDa. GAPDH protein was used as normalization control.

2.5 Characterization of the +9C>T SPINK5 exon 11 mutation

The synonymous +9 C-to-T genomic variation at the 5'-end of exon 11 in SPINK5 gene has been previously reported to affect splicing of this exon, inducing its exclusion from the final mRNA (Fortugno, Grosso et al. 2012). However, the molecular mechanism at the basis of this aberrant splicing has not been investigated. Thus, I decided to characterize the c.891C>T mutation to individuate the splicing factors acting in the mutated region of the pre-mRNA molecule that are responsible for exon 11 skipping.

2.5.1 Pull down assay for identification of nuclear protein differentially bound to SPINK5 exon 11 wt or +9C>T region

To identify specific *trans*-acting factors whose differentially binding to the wild type or mutated sequence could provide an explanation for the observed changes in the pattern of splicing, I performed a pull down assay. I analysed the binding properties of the exon 11 wt and exon 11 carrying the +9C>T mutation. The pull down experiment was carried out using RNA oligos corresponding to the 5'-end of exon 11 wt and exon 11 +9C>T sequences (Figure 2.17). These RNAs were covalently attached to adipic acid dehydrazide agarose beads through their 3'-end and incubated with Hek293 nuclear extract. Upon washing steps, the proteins bound to the target RNA were separated on SDS-PAGE gel and visualized by Coomassie Blue staining (Figure 2.17, panel A). In three independent experiments I detected one protein band that is strongly increased in exon 11 +9C>T fraction. This protein band of approximately 34 kDa was excised from the gel, sequenced using electrospray mass spectrometry analysis and eventually identified as hnRNPA1. Although the pull down analysis revealed that only this protein has a strong differential binding affinity towards the exon 11 wt sequence when compared to exon 11 mutated sequence, it is possible that there are some other proteins that due to their low amount escape from the identification during the pull down experiments (see below, figure 2.18).

A

RNA OLIGO exon 11 wt	agAAACUCUGCAGUCAAUAUCAAAAUC
RNA OLIGO exon 11 +9C>T	agAAACUCUGUAGUCAAUAUCAAAAUC

B

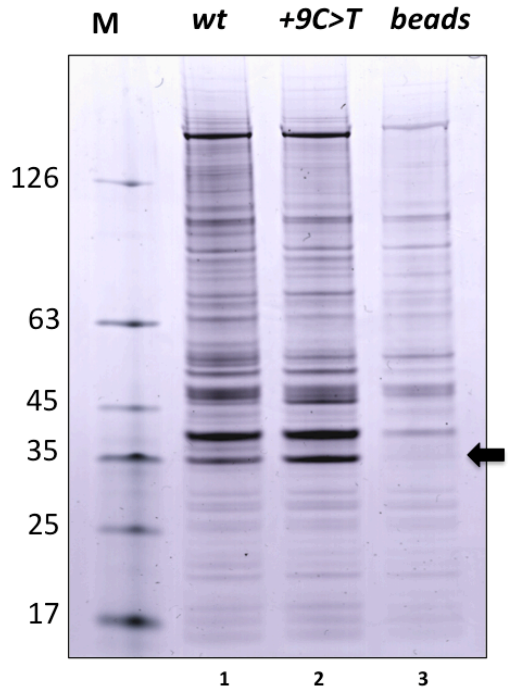


Figure 2.17: Analysis of SPINK5 exon 11 wt and +9C>T binding capacity.

A. Sequences of the exon 11 wt and +9C>T that were used for pull down analysis. The nucleotide that differs in wild type and mutated exon 11 is reported in red.

B. Coomassie staining after a pull down assay. The target RNAs were covalently bound to the agarose beads and incubated with HeLa nuclear extract. Upon rounds of washing the proteins were loaded on a 4-12% SDS-PAGE gel and stained with Coomassie. The arrow indicates a band of 34 kDa that show a differential binding between the two fractions. This band was excised from the gel and analysed by mass spectrometry.

2.5.2 Pull down analysis of exon 11 wt and mutated sequences probed with candidate proteins.

Bioinformatics analysis identified three potential candidates involved in the regulation of SPINK5 mutated exon 11 skipping. ESEfinder (Cartegni, Wang et al. 2003) and human splicing finder (Desmet, Hamroun et al. 2009) programs suggested that the C-to-T substitution in position +9 disrupts an exonic splicing enhancer recognized by ASF/SF2, as previously reported (Lacroix, Lacaze-Buzy et al. 2012). Moreover, human splicing finder identified two putative binding sites for Tra2β, 4 nucleotides downstream the site of the mutation. The mutation was also predicted to create an hnRNP A1-dependent ESS.

To understand the binding properties of these proteins on SPINK5 exon 11 wt and +9C>T, I performed immunoblot analysis. RNA oligos for exon 11 wt and with the +9C>T mutation were treated with sodium periodate to allow their binding to the agarose beads. After incubation with Hek293 nuclear extract and extensive washing to decrease the number of proteins not specifically interacting with the RNAs, the proteins bound to the target RNA were separated on a 4-12% SDS-PAGE gel, blotted onto PVDF membranes and analysed using antibodies against SF2/AF, hnRNPA1/A2, Tra2β. Western blot against hnRNPA1/A2 showed an increase in the binding of these proteins in the mutated +9C>T fraction, suggesting that the mutation reinforces a binding site for hnRNPA1 (figure 2.18 A).

Bioinformatics predictions based on Human Splicing Finder (Desmet, Hamroun et al. 2009) individuated two putative binding sites for Tra2β only 4 nucleotides downstream the +9C>T mutation. Giving that the mutation may create a new strong binding site for hnRNPA1, which may potentially interfere with the binding of Tra2beta, I checked the amount of Tra2β in exon 11 wild type and mutated fractions. Interestingly, the RNA oligo carrying the +9C>T mutation showed a remarkable decrease in the binding of this splicing factor, compatibly with a partial interference in the binding of this protein in the mutated exon 11 (figure 2.18, B). Although ESEfinder and Human Splicing Finder programs indicated a possible disruption of SF2/ASF binding site in

the mutated exon 11, immunoblot analysis did not reveal any significant change between wt and mutated fractions (figure 2.18 C).

The fact that Tra2β protein was not detected in pull down analysis after comassie staining could be explained by the small amount of the protein to a level below the detection sensitivity.

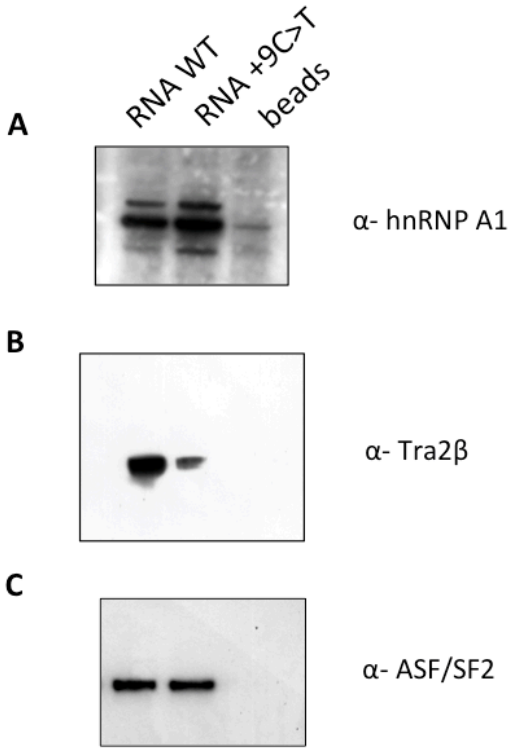


Figure 2.18. Western blot analysis of pull down assays performed on exon 11 wild type and mutated sequences.

Binding analysis of hnRNPA1, Tra2 beta ASF/SF2 proteins. The RNA oligos were covalently bound to the agarose beads and incubated with HeLa nuclear extract. After washing steps, the pulled down proteins were loaded on a 4-12% SDS-PAGE gel and analysed by western blotting using monoclonal anti-hnRNPA1 (**A**), anti-Tra2beta (**B**) and anti-ASF/SF2 (**C**) antibodies.

2.5.3 Overexpression of candidate splicing factors.

To address the functional role of the two *trans*-acting factors that showed differential binding in the mutant, I set up an *in vivo* approach. I used three different hybrid minigenes constructs: SPINK5 exon 11 wt, mutated and with a three-bases deletion in the region of the mutation (ΔTAG). After transfection in Hek293 cells along with plasmids encoding the candidate splicing factors, I evaluated through RT-PCR exon 11 pattern of splicing (figure 2.19). The splicing of exon 11 wild type was affected by overexpression of hnRNPA1, that reduces the percentage of exon inclusion from 75% to 22% (compare lanes 1 and 2). Tra2β slightly increased the exon inclusion (compare lane 1 with lane 4). In the construct that carried the +9C>T mutation, exon 11 is completely skipped and overexpression of hnRNPA1 did not change the pattern of splicing (lanes 5-6) whereas Tra2β exerts a positive effect promoting exon 11 inclusion from 0% up to 30% (lane 8).

I built up a minigene carrying a 3-bp deletion in the region +9/+11 (ΔTAG) in order to verify the role and the importance of these bases for the proper recognition and inclusion of exon 11 in the mature SPINK5 transcript. The deletion of the TAG bases reduced the percentage of exon 11 inclusion in comparison to the wild type minigene, a fact which is compatible with the disruption of a positive regulatory element. Overexpression of hnRNPA1 changed the splicing pattern, decreasing exon inclusion from 30% to 5% (lane 10) and Tra2β had a positive effect on exon 11 splicing, promoting the inclusion of the exon in the final mRNA up to 78% (lane 12). Overexpression of ASF/SF2 did not change the pattern of splicing in none of the SPINK5 exon 11 minigenes (lanes 3, 7, 11).

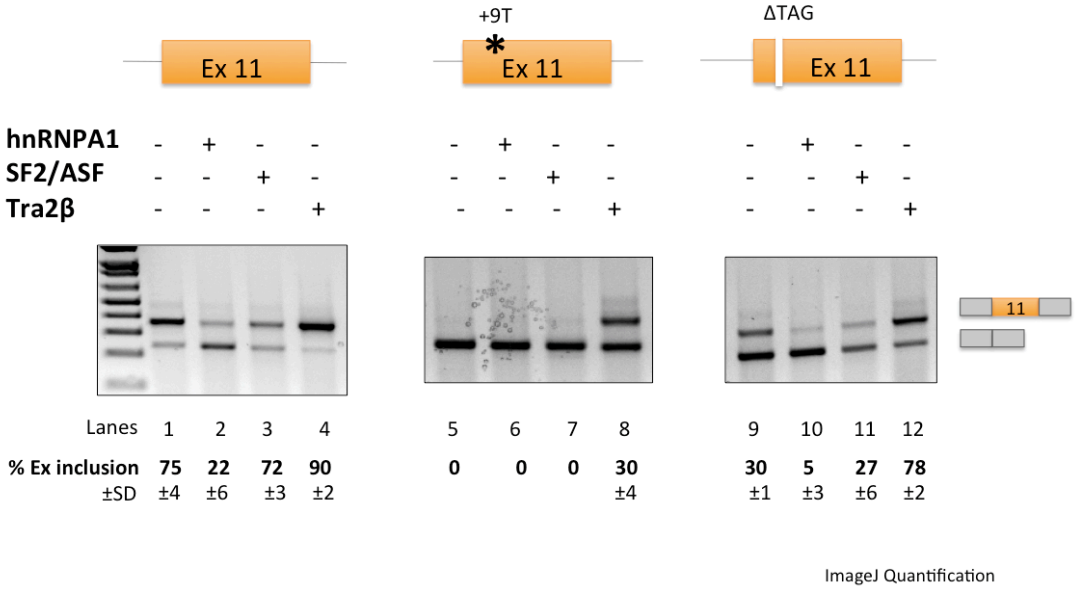


Figure 2.19: Overexpression of candidate splicing factors on SPINK5 exon 11 minigenes.

Hek293 cells were transfected with SPINK5 exon 11 wilde type, +9C>T or deltaTAG minigenes along with vectors expressing the protein of interest (hnRNPA1, SF2/ASF, Tra2β). The splicing pattern was analysed using SA2 and SD6 primers. RT-PCR fragments were solved on 2% agarose gel. The identity of the bands (exon 11 inclusion and exon 11 skipping) is indicated on the right of the gel. Quantification of bands intensity evaluated using ImageJ software is indicated, as mean +-SD, at the bottom of the panel.

2.5.4 Effect of siRNA against hnRNP A1/A2 and Tra2β on SPINK5 exon 11 splicing pattern.

To further evaluate the functional role of hnRNPA1 and Tra2β proteins on SPINK5 exon 11 pattern of splicing, I decided to deplete these proteins from cells using specific siRNAs. The *in vivo* depletion of the candidate proteins was achieved through two round of transient transfection of siRNA oligonucleotides directed against the corresponding proteins. To confirm the siRNA efficacy I performed immunoblot to check the reduction in protein expression level upon siRNA treatment (figure 2.20 B). All the siRNA treatment induced sufficient reduction in the corresponding protein levels, 80-85% for hnRNP A1, 75-80% for Tra2β and 80% for SF2/ASF, whereas the cells transfected with siRNA against an unrelated gene (luciferase) had no effects. Hek293 cells treated with siRNAs were transfected with SPINK5 exon 11 wild type, +9C>T or ΔTAG minigenes. Total RNA was extracted and the pattern of splicing evaluated. Depletion of hnRNPA1/A2 increased exon 11 inclusion in the wild type construct from 80% to 95% and in the ΔTAG minigene from 25% up to 70% (figure 2.20, lanes 2 and 12 respectively). Surprisingly, depleting hnRNPA1/A2 showed no effect on the splicing pattern in the +9C>T construct (lane 7). Silencing of Tra2β has a moderate effect on exon 11 in the wild type context, leading to a decrease of exon inclusion from 80% to 60-65% (lane 3) and a slight effect in ΔTAG context, where the percentage of exon inclusion shifted from 25% to 10-12% (lane 13). Depletion of Tra2β had no effect in the SPINK5 exon 11 +9C>T minigene which is already totally skipped (lane 8). Treatment of cells with siRNA against siRNA against SF2/ASF (lanes 4, 9, 14) or luciferase (lanes 5, 10, 15) did not affect the pattern of splicing of all the minigene constructs tested.

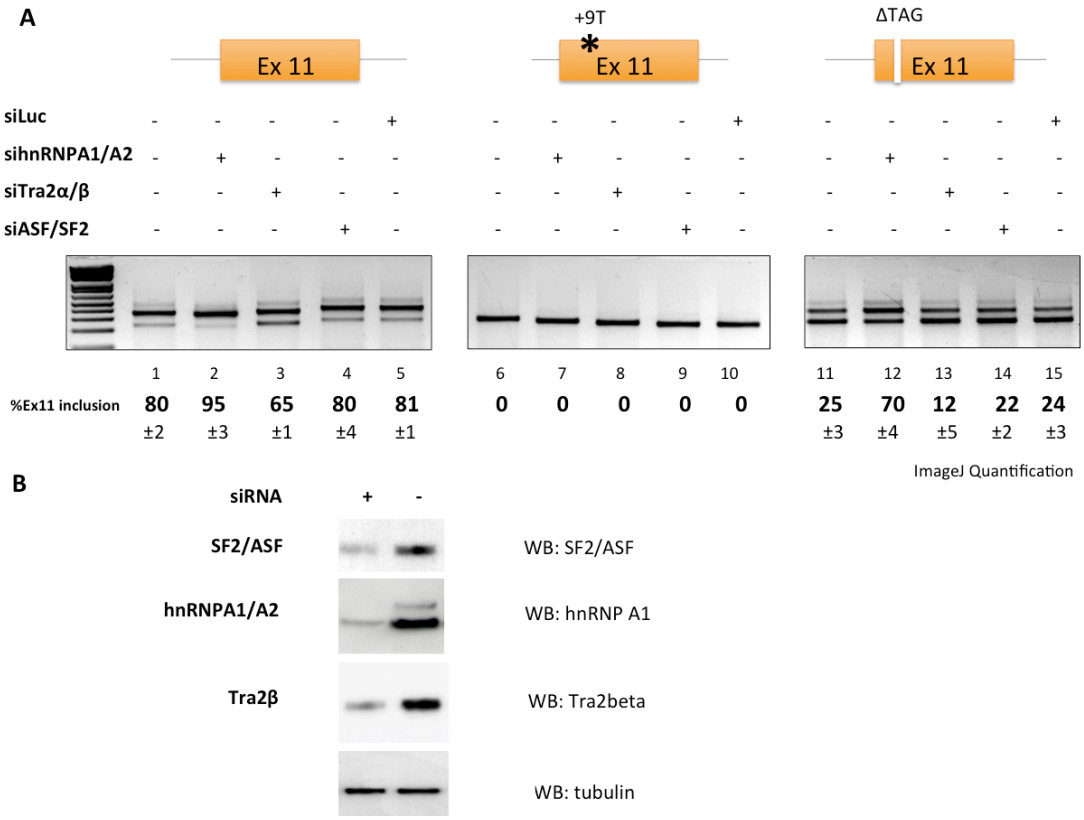


Figure 2.20. Depletion of SF2/ASF, hnRNPA1/A2 and Tra2α/β proteins by siRNA treatment.

A. SiRNA treated and untreated Hek293 cells were transfected with SPINK5 exon 11 wilde type, +9C>T or ΔTAG constructs and the splicing pattern evaluated by RT-PCR. The resulting amplified fragments were solved on 2% agarose gel. The upper band corresponds to exon 11 inclusion, the lower band to exon skipping. The intensity of the bands were evaluated through ImageJ software and the relative quantification in expressed, as mean +/- SD, on the bottom of the gel. The identity of the siRNA used to knock down the corresponding proteins is indicated.

B. Western blot against SF2/ASF, hnRNP-A1 and Tra2β to verify the efficacy of siRNA treatments on the endogenous protein levels. Tubulin was used as normalization control.

2.6 Composition of exon-specific U1 particle.

In order to clarify the mechanism at the basis of exon specific U1-mediated splicing rescue, I investigated the composition of the ExSpe U1 particle along with the role of U1-specific proteins U1-A and U1-70K.

2.6.1. ExSpe U1s are assembled into ribonucleoparticles (RNPs)

To understand if the ExSpe U1s are assembled into ribonucleoprotein particles I performed electromobility shift assays (EMSA). For this analysis, I selected the SPINK-specific ExSpe U1 sk12 and the SMN-specific ExSpe U1 EE2 because their 5'-protruding tails are longer and the corresponding sequence sufficiently different from the normal U1 RNA. These longer tails and their sequence compositions allowed an efficient discrimination from the endogenous U1 snRNA. I radiolabelled three different RNA oligos, one specific for the U1 snRNA WT and two other for the ExSpe EE2 or ExSpe sk12 and I tested them on nuclear extract of wild type cells or cells transfected with ExSpe EE2 or sk12. EMSA experiments demonstrate that the protein complex formed by the U1 snRNA has the same electromobility in non denaturing conditions of that one formed by the ExSpe U1 EE2 and ExSpe sk12 (figure 2.21, compare lanes 1-2). ExSpe EE2 or ExSpe sk12 oligo did not show any cross-reactivity with WT U1 (lane 3). Thus, the exon specific U1s EE2 and sk12 are assembled into ribonucleoparticles with the same mobility of the wild type U1 snRNP, suggesting they have a similar composition.

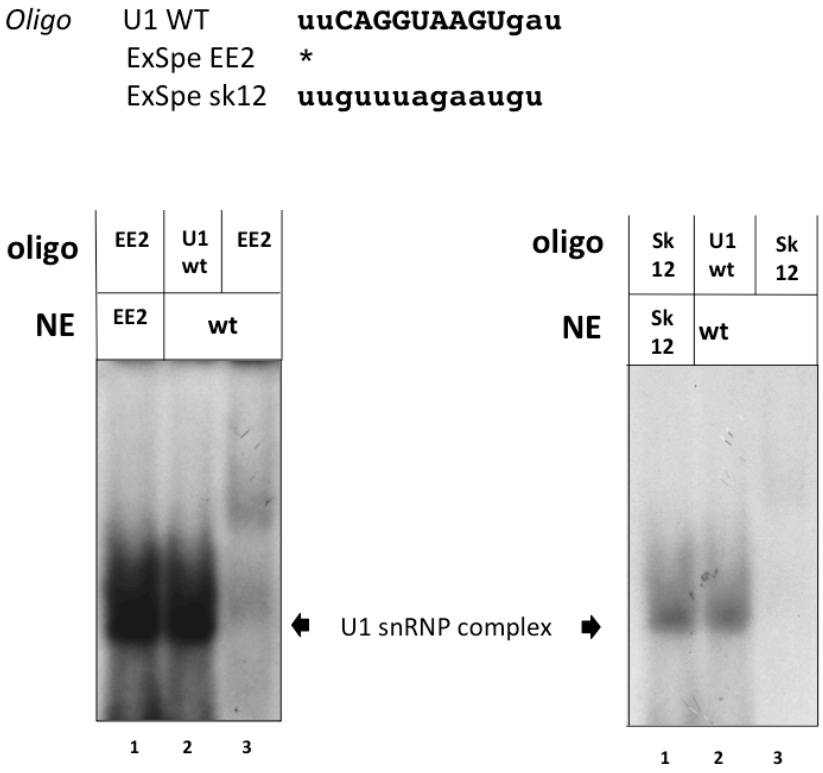


Figure 2.21. Comparison between protein complex formed onto U1 snRNA and ExSpe U1 sm25 or ExSpe sk12 RNAs.

A. Sequences of RNA oligos complementary to U1snRNA WT 5'-tail and sk12.

*The sequence complementary to ExSpe U1 EE2 is not reported because ExSpe U1 EE2 is under patent registration.

B. EMSA assay. Radiolabelled RNAs were incubated with nuclear extracts wild type or transfected with ExSpe U1s EE2 or sk12. The RNA-protein complexes were solved on non-denaturing 5% polyacrylamide gel. The position of the U1 WT complex is indicated.

2.6.2. Analysis of U1-A and U1-70K -deficient ExSpe U1s.

The normal U1 snRNP is composed by the a 164 nucleotide long RNA that forms a characteristic secondary structure on which is assembled a core of 7 Sm proteins in common to the U-rich small nuclear RNAs along with three U1-specific proteins, U1-A, U1-70K and U1-C (Egloff, O'Reilly et al. 2008). As the electromobility shift assays strongly suggest that the ExSpe U1 have a composition similar to the U1 wild type particle, I decided to further investigate the functional contribution of the U1-specific proteins in ExSpe U1-mediated splicing rescue. As the two largest U1-specific proteins, U1-70K and U1-A bind specifically to stem loops I and II, respectively (Surowy, van Santen et al. 1989, Stark, Dube et al. 2001), I carried out site-directed mutagenesis on these loops to abolish their interaction with their corresponding sites. U1-C is considered to be part of the U1 particle through a protein-protein interaction with U1-70K (Stark, Dube et al. 2001, Rosel, Hung et al. 2011). I created a U1-A ExSpe U1 sm21 mutant that carries 3 point mutations in the stem loop II that was previously shown to abolish U1-A interaction with the U1 snRNA *in vitro* (Gunderson, Polycarpou-Schwarz et al. 1998). To create the U1-70K mutant I replaced the entire stem loop I with the binding sequence for the MS2 protein, in order to ensure the total disruption of the U1-70K interaction and, possibly, in case of loss of function due to its absence, to rescue the activity with a MS2-U1-70K chimeric protein, as previously described during *in vitro* studies (Gunderson, Polycarpou-Schwarz et al. 1998).

I performed SMN2 minigene splicing assays using the U1-A and U1-70K ExSpe U1 sm 21 mutants (figure 2.22). Cotransfection of the U1-A mutant slightly affected ExSpe U1 sm25 rescue efficacy, decreasing the percentage of exon 7 inclusion from 95% to 85% (figure 2.22, compare lanes 2 and 3). ExSpe U1 sm21 70K mutant reduced exon 7 splicing correction in a more marked manner, reducing the percentage of exon inclusion up to 45% (lane 4). I tried to recover ExSpe U1 sm21 70K mutant biological activity through expression of the MS2-U1-70K fusion protein, with no success (lane 5). Thus, the U1-specific proteins U1-A and U1-70K contribute to a different extent to

the splicing rescue activity exerted by the ExSpe U1 sm25. U1A has a minimal effect on the activity of the ExSpeU1s , whereas ~ 50% of the activity is lost for the 70K mutant. However, ovexpressison of an MS2-U1-70K fusion protein, aimed at restoring U1-70K binding, had no effect.

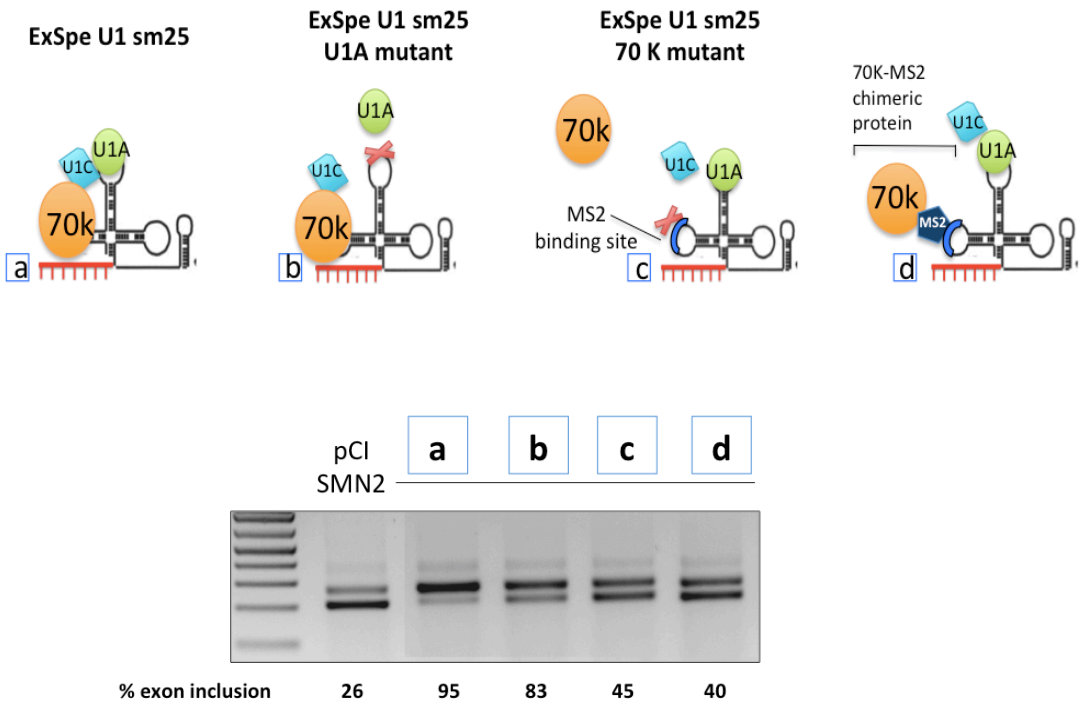


Figure 2.22: Contribution of U1-specific proteins to ExSpe U1-mediated splicing rescue.

Hek 293 cells were transfected with SMN2 minigene plasmid alone (lane 1) or along with ExSpe U1 sm21 (lane 2) or ExSpe U1 sm21 U1A mutant (lane 3) or ExSpe U1 sm21 70K mutant (lane 4) or ExSpe U1 sm21 70K mutant in cotransfection with MS2-70K chimeric protein encoding plasmid (lane 5). Total RNA was extracted and RT-PCR carried out using pCIF-Fw and E8-75 Rev primers. The amplified fragmets were loaded on 2% agarose gel. The identity of the bands is indicated on the right of the panel. Band intensity was evaluated using ImageJ software and the percentage of exon 7 inclusion, expressed as mean of three different expriments, is reported below the gel.

2.7 Effect of U7 RNAs targeting the regions bound by the ExSpe U1s.

In order to better elucidate the mechanism by which the ExSpe U1s exert their splicing correction activity, I investigate the activity of the target sequences inserted in a different particle.. To perform this analysis I took advantage of the U7smOpt molecule (see chapter 1.4), which is an optimized RNA that has broadly been used as antisense carrier for masking pre-mRNA regions (Madocsai, Lim et al. 2005, Marquis, Meyer et al. 2007, Meyer, Marquis et al. 2009). I replaced the RNA binding sequence of the U7smOPT antisense molecule with the different RNA binding sequence of ExSpe U1 EE2 for SMN2 exon 7 and ExSpe U1 sk12 and sk15L for SPINK5 exon 11. I tested the engineered U7 molecules that overlap the same intronic regions downstream the 5' splice sites targeted by the ExSpe U1s. Minigene splicing assays revealed that none of the engineered U7 targeting the intronic region downstream SMN2 exon 7 donor splice site was able to exert an effect on splicing (figure 2.23, panel A, compare lanes 1-7 and panel B compare lanes 2, 7, 12). To prove that the U7s are really produced and target the intronic regions, I performed competition experiments. A fixed amount of ExSpeU1 was contrasfected with increasing amount or the corresponding U7s. This analysis showed that each U7 molecule tested is able to compete with the corresponding ExSpe U1 in a dose dependent manner (panel A, lanes 3-6 and panel B lanes 4-6, 9-11), antagonizing the positive effect on exon 7 inclusion exerted by the ExSpe U1. The competition is more pronounced for the Sk12 and Sk15L and less pronounced for EE2. This indicates that the U7 antisense molecules are correctly produced and are able to bind to the same regions targeted by their corresponding ExSpe U1s. Thus, masking of the same target intronic sequences by RNAs that differ from ExSpe U1 is not sufficient to induce exon inclusion. This suggests that the specific composition of the ExSpe U1 particle is directly responsible for their splicing rescue activity.

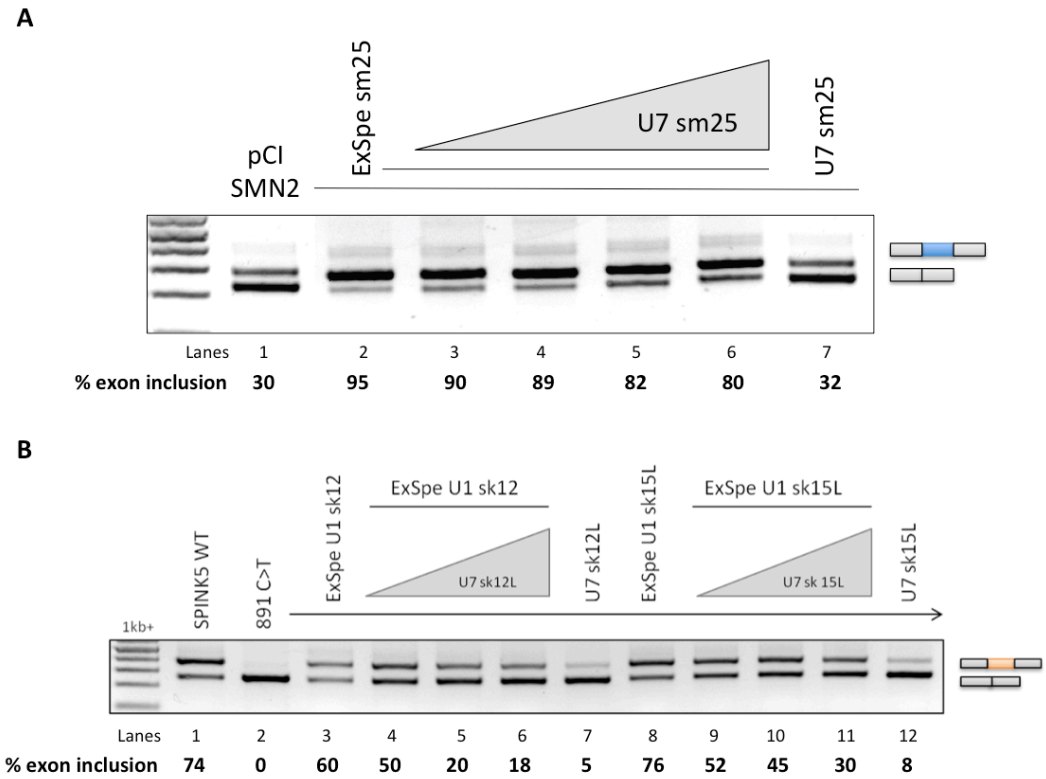


Figure 2.23. Competition between ExSpe U1 and the corresponding U7 snRNA

A. SMN2 exon 7 minigene plasmid was transfected alone (lane 1) or along with ExSpe U1 EE2 (lane 2) or U7 EE2 (lane 7) encoding plasmids. The competition for the target region on the SMN2 intron 7 pre-mRNA was achieved cotransfecting 0.25 µg of ExSpe U1 EE2 along with increasing doses of U7 EE2 (0.25 – 0.5 -1 µg). Total RNA was extracted and RT-PCR carried out using pCIFfw and E8-75Rev primers. The resulting amplified products were solved on 2% agarose gel. The identity of the bands is indicated on the right of the gel. The percentage of exon 7 inclusion was calculated using ImageJ software and the quantification reported below the gel are the mean of three independent experiments.

B. SPINK5 exon 11 +9C>T minigene plasmid was transfected in Hek293 cells alone (lane 2) or in combination with ExSpe U1 sk12 or sk15L (lanes 3 and 8). The competition between each ExSpe U1 and its corresponding U7 snRNA for the target region on the SPINK5 intron 11

was performed as described in (A). Total RNA was extracted and RT-PCR carried out using SA2 and SD6 primers. The resulting amplified products were solved on 2% agarose gel. The identity of the bands is indicated on the right of the gel. The percentage of exon 11 inclusion was calculated using ImageJ software and the quantification reported below the gel are the mean of three independent experiments.

2.8 Role of endogenous U1 snRNP and hnRNPA1 in SMN2 exon 7 splicing rescue induced by EXSpeU1s and AON 10-27

To further investigate how ExSpeU1 rescues exon skipping, I considered in the SMN2 case the roles of the endogenous U1snRNP and of the inhibitory splicing factor hnRNPA1 that binds to the ISS-N1. Both overexpression of hnRNPA1 and functional suppression of U1snRNP have been shown to inhibit SMN2 exon 7 inclusion (Hua, Vickers et al. 2008, Roca and Krainer 2009). In order to determine if the ExSpe U1 molecules require the presence of the endogenous U1 particle at the canonical 5' splice site to exert its splicing correction activity, I decided to monitor the effect of the ExSpe U1 and of the antisense oligonucleotide 10-27 on SMN2 minigene when the endogenous U1 particle is sequestered in the cells. I performed SMN2 minigene splicing assays using U1 snRNA specific decoy (D1), a plasmid encoding an RNA decoy previously described to interact by complementarity with the endogenous U1 RNA (figure 2.24 A). The resulting functional inactivation of U1 snRNP causes SMN1/2 exon 7 exclusion from the mature transcripts (Roca and Krainer 2009). As control I used a U1 decoy that contains a mis-match in the basepairing with the U1 (D3), which has no effect on SMN1/2 splicing pattern (Roca and Krainer 2009) I cotransfected the decoy plasmids and the SMN2 minigene along with ExSpe U1 sm21 or AON 10-27 to test if their SMN2 splicing correction on the minigene model was functionally related to the endogenous U1 snRNP (figure 2.24). As previously reported, the U1snRNA decoy (D1) affects SMN2 splicing, with a decrease in the percentage of exon 7 inclusion from 31% to 5% (figure 2.24, panel B,

lanes 1, 2). Control U1 decoy (D3) did not significantly alter exon 7 inclusion (panel B, lanes 1, 3). Cotransfection of the ExSpeU1 sm21 completely prevented the negative effect produced by the decoy D1 (figure 2.25 B, compare lanes 2 and 4), On the contrary, AON 10-27 is not able to improve exon 7 inclusion in the presence of the decoy D1 (panel B, lane 6), compatibly with its well described action of steric blocking of the hnRNPA1-dependent ISS-N1.

In addition, I tested in SMN2 exon 7 minigene the dose-dependend effect of hnRNPA1 overexpression on the rescue efficiency mediated by ExSpeU1s or by AON 10-27. According to the proposed antisense mechanism of AON 10-27, cotransfection of hnRNPA1 did not affect the splicing rescue induced by the oligonucleotide, indicating that it directly blocks the binding of the splicing factor to the ISS N1 (Fig. 2.25, lanes 5-8). On the contrary, increasing the amount of co-transfected splicing factor progressively reduced the splicing rescue induced by ExSpeU1 sm 21 (figure 2.25, lanes 9-12).

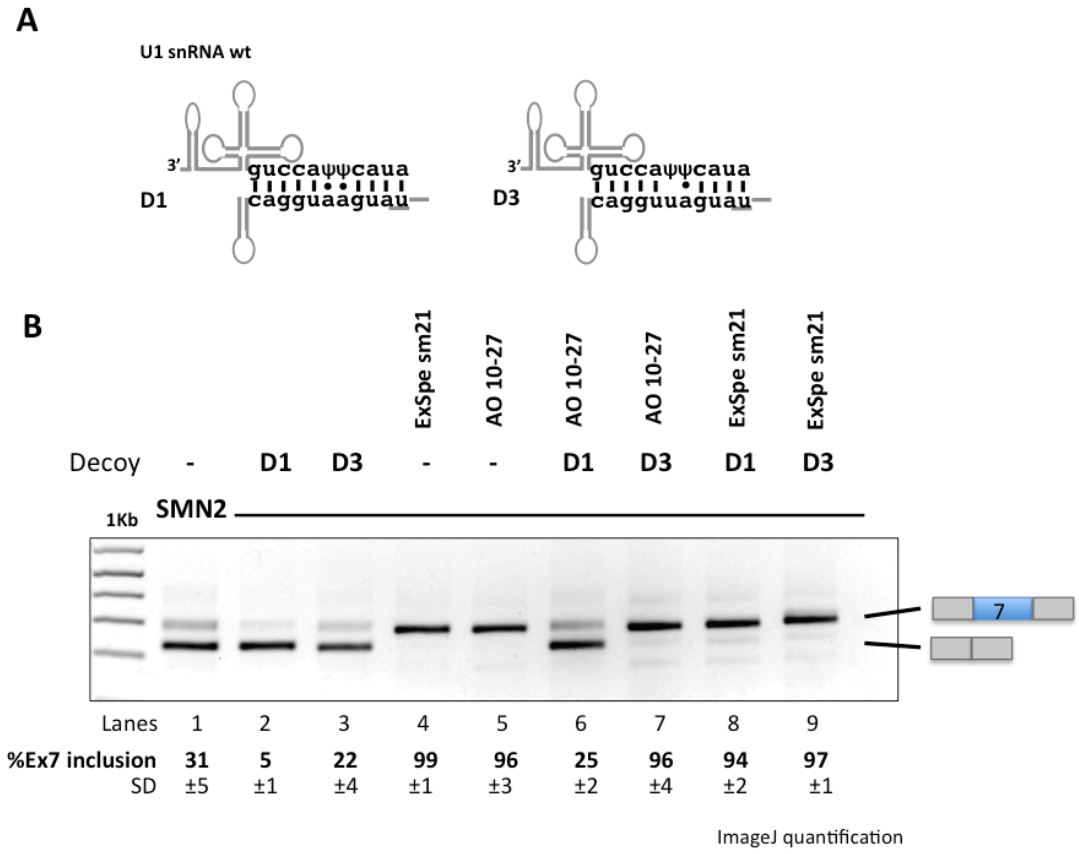


Figure 2.24. ExSpe U1 –but not the AON- does not require the endogenous U1 snRNP at the 5'ss.

A. Schematic representation of the basepairing between the 5'-tail of the U1 snRNA wt and the tail of the RNA U1-decoy.

Hek 293 cells were transfected with pCI SMN2 plasmid alone or in combination with U1snRNA decoy (D1), U1snRNA decoy with a mis-match in position +3 (D3), ExSpe U1 sm21 or antisense oligonucleotide 10-27 (AO 10-27).

B. Total RNA was extracted and retrotranscribed using random primers. RT-PCR was performed using pCIFdir and E8-75Rev primers. The resulting amplified products were solved on 2% agarose gel. The upper band is exon 7 inclusion, the lower band exon 7 skipping. Bands intensity was evaluated using ImageJ software and the percentage of exon inclusion has been reported for each lane at the bottom of the gel, as mean ±SD of three independent experiments.

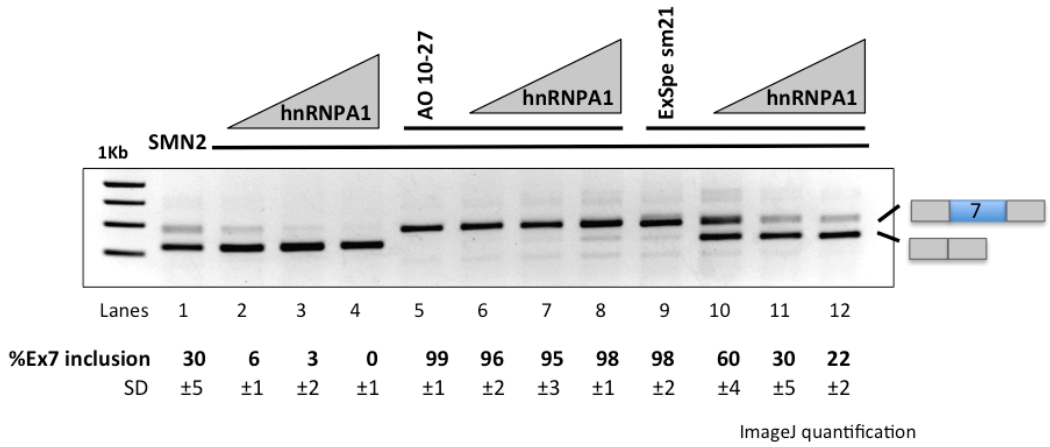


Figure 2.25. Overexpression of hnRNP A1 impair exon 7 splicing correction mediated by ExSpe U1, but not the one promoted by AO 10-27.

SMN2 minigene splicing assay has been performed in Hek293 cells with ExSpe U1 sm 21 or AO 10-27 along with increasing amount of hnRNP A1 splicing factor (50-150-250 ng). Total RNA was extracted and RT-PCR carried out using pCIF-fw and E8-75 Rev primers. The amplified products were solved on 2% agarose gel. The identity of the bands is indicated on the right of the panel and their intensity has been evaluated through ImageJ software. The percentage of exon 7 inclusion, expressed as mean ±SD of three different experiments, is reported below the gel.

3. DISCUSSION

In this thesis work, I have studied two synonymous exonic substitutions that cause skipping of the exon from the final mRNA and I have evaluated a novel strategy for splicing correction based on ExSpeU1s. Exon-Specific U1 snRNAs (ExSpeU1s) are modified U1 snRNAs that interact by complementarity to intronic sequences downstream of the 5'-splice sites and rescue exon skipping caused by different types of mutations. To investigate their therapeutic activity and molecular mechanism of action *in vivo,* I focussed on spinal muscular atrophy (SMA) and netherton syndrome (NS) where a silent exonic substitution induces SMN2 exon 7 and SPINK5 exon 11 skipping. In SMA-derived primary fibroblasts, lentiviral-mediated transduction of different ExSpeU1s improved SMN2 exon 7 splicing and SMN protein levels. In normal Hek293 Flip-In cell system, a significant correction of endogenous SMN2 splicing was obtained with a single integrated copy of ExSpeU1 gene. In comparison to antisense oligonucleotide (AON)-mediated SMN2 splicing correction, ExSpeU1s exploit a completely different mechanism. Even if AONs and ExSpe U1 interact with partially overlapping sequences in the intron, ExSpeU1 -but not AON- is sensitive to hnRNPA1 overexpression, but does not require functional endogenous U1 snRNP. In addition, ExSpeU1 induces higher levels of both spliced and unspliced mature SMN transcripts. Thus, ExSpeU1s have an overall positive effect on pre-mRNA processing, which results in significant SMN2 exon 7 splicing correction activity *in vivo*.

3.1 Bioengineered U1 snRNAs molecules rescue defective splicing due to exonic mutations in SPINK5 exon 11 and SMN2 exon 7.

Binding of U1 snRNP at the donor splice site represents the first step in the spliceosome assembly. U1 snRNP interacts with the 5'ss through its protruding 5'-tail and this promotes the recruitment of the splicing machinery and the beginning of the splicing reaction (Sperling, Azubel et al. 2008). Recognition of the 5'ss depends on the complementarity between the 5'-splice site and the 5'-tail of U1 snRNA. However, as the 5'ss consensus is degenerated, additional splicing factors can assist the U1 snRNP in the recognition of the donor site. At the end a network of complex interactions is formed on the exon, which is identified and processed by the spliceosome.

The two defective exons I have evaluted in this thesis, the SMN2 exon 7 and SPINK5 exon 11, are characterized by the presence of exonic mutations that cause their skipping and by the fact that the sequence surrounding their 5'ss is not completely conserved. Both these exons are excluded from the mature transcript due to an exonic C>T synonymous substitution in position +6 and +9, respectively (Lorson, Rindt et al. 2010, Fortugno, Grosso et al. 2012). SMN2 exon 7 and SPINK5 exon 11 donor splice sites do not perfectly match the canonical 5'ss consensus sequence. In fact, exon 7 carries variations in position -3 and -1 and exon 11 has variations in position -2, -1 and +3 in the donor splice sites, which however do not affect the invariant GU dinucleotide. As a result of these variations the corresponding 5'ss are weak.

To identify a way for exon skipping correction, I firstly considered the possibility to directly load the U1 snRNA on the weak 5'splice sites. To this aim, I prepared U1 snRNAs that carry compensatory substitutions to perfectly match the consensus sequence of SMN2 exon 7 and SPINK5 exon 11 donor splice sites. Using minigene splicing assays, in both the systems, I demonstrated that these U1snRNP fully complementary to the corresponding 5'-splice sites promotes complete inclusion of defective exons (figures 2.2 B and 2.3 C, lanes 1, 8 and 2, 3 respectively).

These results are similar to previous studies that analysed exon skipping mutations that directly affect the 5'ss. In these cases, exon skipping is the

result of a mutation that affects the conserved sequences of the 5'ss: in several cases U1 snRNAs complementary to the mutated donor sites were shown to be able to rescue the splicing defect (Pinotti, Rizzotto et al. 2008, Schmid, Glaus et al. 2011, Schmid, Hiller et al. 2012).

However, as the tail of these modified U1 snRNA are very similar to the endogenous U1 sRNA (they differ just for few bases and retain the invariant GU "core") they can interact by complementarity with other similar 5'ss. As a result of this interaction they can significantly affect the splicing pattern of others genes with unpredictable consequences, such as alterations of alternatively spliced exons, activation of cryptic splice sites and inclusion of pseudoexons in the final transcripts. For example, a recent paper showed that a modified U1 snRNA loaded to the defective donor splice site of coagulation factor VII exon X leads to hepato-toxicity *in vivo* in a mouse model of haemophilia A (Cavallari, Balestra et al. 2012). To avoid these possibilites, I prepared and tested modified U1s, named exon specific U1s (ExSpe U1s), which target intronic sequences located at different distances downstream the canonical donor sites. Specifically, the exon-specific U1 snRNAs I used, have the 5'-tail engineered to allow their binding to intronic regions downstream SMN2 exon 7 and SPINK5 exon 11 5'-splice sites. For SMN2 exon 7, I tested in minigene assay six different ExSpe U1s and all of them were able to promote exon inclusion with the same high efficacy (from 30% in untreated cells to up to 95%), compensating the deleterious effect of the exonic +6 C-to-T transition. Similarly, in SPINK5 exon 11, I tested six ExSpe U1s. Five out of six ExSpe U1s tested were able to enhance exon inclusion, even if with different efficacy. Three of them correct the total exon 11 skipping caused by the c.891C>T mutation promoting its inclusion up to 85-90% and two ExSpe U1s enhance exon inclusion up to 20% and 70%. Only one, ExSpe U1 sk6, did not show any significant splicing rescue activity: this could be due to an alteration of the secondary structure of the modified U1 snRNA, due to its 5'-tail sequence composition that could, in turn, impairs the correct biosynthesis of the particle.

3.2 ExSpe U1s as novel therapeutic molecules for mis-splicing events.

To understand the potential therapeutic activity of ExSpe U1s molecules, I tested them in SMA-derived primary fibroblasts and inducible pluripotent stem cells (iPSCs) and NS primary keratinocytes.

SMN2 exon 7 splicing correction is a reliable therapeutic approach for spinal muscular atrophy. In fact, in the past years several attempts have been carried out in order to promote exon 7 inclusion in endogenous SMN trascripts in different SMA cellular models using antisense oligonucleotides, bifunctional oligonucleotides, engineered U7 snRNAs and *trans*-splicing (Madocsai, Lim et al. 2005, Hua, Vickers et al. 2008, Rigo, Hua et al. 2012, Zhou, Zheng et al. 2012)

As a first attempt to test the efficacy of ExSpe U1s in a physiological context, I tested their activity on Hek293 SMN endogenous transcripts. Episomal expression of ExSpe U1 molecules was able to modulate with high efficacy endogenous exon 7 splicing rescuing its inclusion in the final mRNA (figure 2.4). Moreover, for a future application of ExSpeU1s in gene therapy, it would be important to assess how many copies of ExSpe U1s are necessary to ameliorate the SMN2 exon 7 pattern of splicing. Using the well-controlled Hek293 Flp-In cell system that allows the insertion of only one copy of a gene of interest in a specific FRT site, I was able to perform a reliable comparison of the different ExSpeU1s activities. The resulting stable clones demonstrated that a single copy of ExSpe U1 resulted in a significant increase in the percentage of SMN2 exon 7 inclusion and in the SMN FL isoforms (figures 2.5 and 2.6). For a therapeutic application *in vivo*, the finding that just one copy of ExSpe U1 is able to modulate splicing is advantageous in terms of the lower amount of viral-mediated delivery of the ExSpe U1s.

Interestingly, none of the Hek293 stable clones showed impairment in viability, suggesting that, at least in my experimental conditions, ExSpe U1s do not lead to cellular toxicity .

In addition, the finding that ExSpe U1s modify SMN2 splicing increasing not only the FL isoform but also the Δ7 may have interesting consequences for SMA treatment. In fact, ExSpe U1s not only correct splicing, but also increase

the delta7 isoforms, even if to a lesser extent in comparison to the FL ones. Even if the delta7 isoform is unstable due to the presence of a strong degradation signal at its C-terminus, it retains some partial activity. In vivo, the amount of the delta7 isoform positively correlates with a less severe phenotype of the disease and it is associated with a mild SMA phenotype in mice (Le, Pham et al. 2005). For example, the first SMA mouse model developed has a deletion of the murine *Smn* and contains two human SMN2 gene copies, leading to the most severe clinical phenotype of SMA. However, it has been shown that the transgenic addition of SMN cDNA lacking exon 7 (SMNΔ7, the major product of SMN2) extend the lifespan of the SMA mice from ~6 days up to ~13 ((Le, Pham et al. 2005)). Thus, ExSpe U1 therapy could represent a valid option to have stronger phenotypic improvement for those patients with type I SMA, that retain only two copies of SMN2 genes, in which the splicing correction could be reinforced by an increase in the delta7 isoform. In this context, the concomitant usage of AON and ExSpe U1 may have a synergic effect.

In netherton syndrome primary keratinocytes, I demonstrated that lentiviral transduction of ExSpe U1 sk12 rescues SPINK5 full length transcript and the corresponding protein levels (figures 2.14 and 2.16). Netherton syndrome is a pathology for which there is currently no cure and ExSpe U1s represent the first approach for an RNA-based therapy in this disease. The recent success of lentiviral mediated gene therapy in Wiscott-Aldrich pathology indicates the safety and the efficacy of these viral vectors for clinical applications (Aiuti, Biasco et al. 2013, Biffi, Montini et al. 2013). Thus, these findings suggest a possible approach using the ExSpe U1-mediated correction for treatment of NS with an *ex vivo* therapy, as previously demonstrated by De Luca and collegues in their seminal works (Mavilio, Pellegrini et al. 2006).

In this work I have specifically evaluated ExSpe U1s in two systems where exon skipping is caused by synonymous mutations relying on the exons. Moreover, complementary data from my lab has shown that these engineered molecules with strong splicing modulation activity are active also on other type of mutations. In fact we recently showed that these molecules

are able to exert a splicing correction activity also in case of mutations that affect the 5'ss consensus sequence or the 3'ss. In coagulation factor IX and CFTR gene models, we showed that a single ExSpe U1 loaded downstream of the corresponding exon was able to rescue exon skipping caused by different mutations in the exon, at the 5'ss (but not at the invariant GU dinucleotide) and in the polypyrimidine tract sequence (figure 3.1, Fernandez Alanis et al, 2012).

The possibility to revert aberrant splicing due to mutations located far away from the canonical splice sites using an ExSpe U1 represents a novel approach for splicing modulation. In fact, in comparison to modified U1 that reinforce defective donor splice sites, targeting less conserved intronic regions downstream the 5'ss decreases the possibility of undesired off-target events, opening the opportunity of new efficient molecular therapies for mis-splicing events. Moreover, for instance in SMA, ExSpe U1s could represent a novel option to treat the disease. As ExSpe U1s seem to exploit different molecular mechanism to redirect exon 7 splicing in comparison to the AO-based therapy, they can be used in a synergic way with the AONs to ameliorate the most severe SMA type I phenotype.

In addition, in comparison to classical gene-replacement approaches, the introduction of engineered U1 molecules has the advantage of leaving the mutated gene in its natural genetic context, avoiding possible disruption of epigenetic regulation processes. In case of NS, for instance, it is of crucial importance to leave SPINK5 gene in its own genetic context under specific epigenetic conditions, as it has to be expressed in a tissue-specific manner at the spinous layer in the skin. Thus, a genetic therapy using ExSpe U1s could represent a valid option for the treatment of those diseases in which the correction has to occur only in specific tissues or during specific developmental stages.

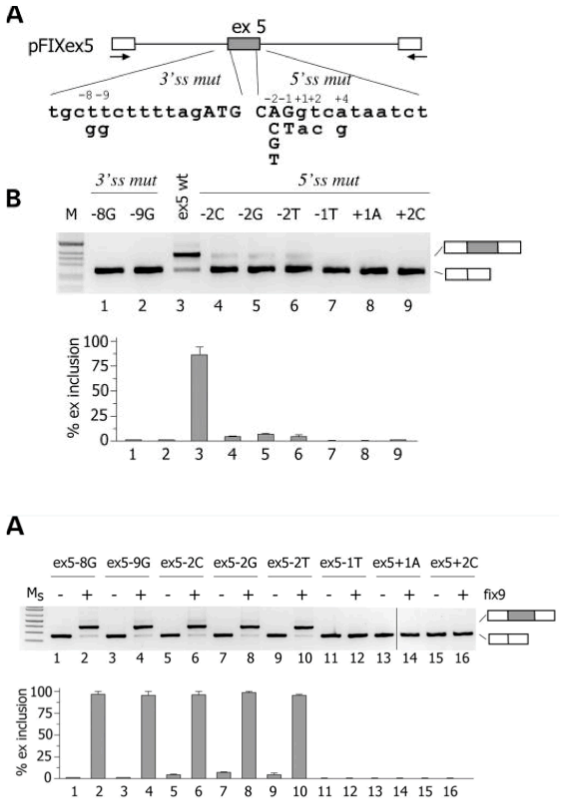


Figure 3.1 ExSpe U1s correct exon skipping due to different type of mutations in other genes (Fernandez Alanis et al., 2012).

A. Schematic representation of the central region of the pFIX exon 5 minigene. Exonic sequences are boxed, introns are lines and the arrows represent primers for PCR amplification. The position of genomic variants relative to the splice sites is indicated.

B. Analysis of pFIX exon 5-spliced transcripts. HeLa cells were transfected with pFIXex5 wt or mutant minigenes and splicing pattern evaluated by RT–PCR. Amplified products were resolved on a 2% agarose gel. The identity of the bands is indicated on the right of the panel. M is the molecular 1 Kb marker. Lower panel shows the quantification of the percentage exon 5 inclusion (mean+SD of three independent experiments done in duplicate).

A. Splicing correction mediated by fix9 and cf11 U1 ExSpeU1s on natural mutations. HeLa cells were transfected with the indicated FIX exon 5 mutant minigenes alone (even lanes) or along with fix9 U1 ExSpeU1 (odd lanes). Splicing pattern was evaluated by RT–PCR and amplified products

were resolved on a 2% agarose gel. Ms is the molecular 1 Kb plus marker. Lower panel shows the quantification of exon 5 splicing pattern. Percentage of exon inclusion is expressed as means+SD of three experiments done in duplicate. [From Fernandez-Alanis, Pinotti, Dal Mas et al., 2012 Human Molecular Genetics]

3.3 c.891 C>T mutation creates an hnRNP A1-dependent exonic splicing silencer in SPINK5 exon 11.

The development of fast sequencing methods has allowed in the past 20 years the discovery of genomic variants in the human DNA during the analysis of genes associated to diseases. Nevertheless, genomic variants are not always ascribable to a clear molecular impairment of some biological functions (change of aminoacids, creation of premature termination codons, shift of the reading frame). Variations that seem to have no effect on the phenotype, may exert unpredictable and deleterious effect on pre-mRNA molecules causing severe diseases (Pagani and Baralle 2004, Cooper 2005, Cooper, Wan et al. 2009). It has been estimated that 15% of pathological human mutations cause the disease through the defect they introduce in the splicing mechanism (Baralle, Lucassen et al. 2009). The study of the correlation between genomic variants that could potentially affect splicing and the occurrence of human diseases has become a central issue in medical practice (Baralle, Lucassen et al. 2009). In fact, although it is quite easy to predict that a mutation affecting one of the invariant splicing signals will cause aberrant splicing, it is not so obvious in case of synonymous mutations located in exons or variations in introns far away from splice sites, that could be erroneously classified as benign substitutions. Frequently, those mutations destroy or modify splicing regulatory elements, leading to the mis-recognition of the exon by the spliceosome machinery (Pagani, Stuani et al. 2003, Pagani and Baralle 2004).

The synonymous c.891C>T in SPINK5 gene is a recently described genomic variant that, inducing exon 11 skipping in patients primary keratinocytes, causes netherton syndrome (Fortugno, Grosso et al. 2012). This represents a relative frequent mutation and 12 patients have been identified. The molecular mechanism at the basis of exon 11 skipping is unknown but previously bioinformatics analysis proposed that the mutation disrupts an SF2/ASF-dependent enhancer (Cartegni and Krainer 2002)

To better understand how the mutation affects splicing I performed pull down analysis of candidates splicing factors followed by the analysis of their functional role in overexpression and silencing experiments.

Bioinformatics analysis of the mutated sequence with ESEfinder (Cartegni, Wang et al. 2003) and Human Splicing Finder predicted a disruption of an exonic splicing enhancer (ESE) recognized by SF2/AF. However, in *vitro* studies of RNA-protein interactions clearly demonstrate that there are no differential binding of this regulatory protein in wild type and mutated exon. Overexpression and silencing of SF2/AF do not change exon 11 pattern of splicing. It is well known that not always bioinformatics predictions correlate with functional analysis.

The +9C>T mutation creates, according to bioinformatics analysis, a TAG motif recognized by hnRNP A1 suggesting the creation of an exonic splicing silencer (ESS). Indeed, this triplet was previously found in silencer elements described also in other gene models (Disset, Bourgeois et al. 2006, Kashima, Rao et al. 2007), in the hnRNP A1 SELEX sequence (Burd and Dreyfuss 1994) and in a screening of a large number of putative silencer elements (Wang, Rolish et al. 2004). To confirm this hypothesis and to characterize additional proteins assembled on the wild type and +9C>T sequence, I performed a pull down analysis with specific RNA oligos. Pull down followed by mass spectrometry analysis identified hnRNP A1 as the splicing factor that increases its binding with the exonic silencer element created by the mutation. I confirmed hnRNP A1 increased binding also by western blot analysis (figure 2.18). However, siRNA-mediated depletion of hnRNP A1/A2 does not rescue at all the exon inclusion in the mutated minigene. (figure 2.20).

Moreover, two consecutive binding sites for Tra2β just 4 nucleotides downstream the site of the mutation were predicted by Human Splicing Finder program. To investigate the role of Tra2β I performed a pull down analysis followed by western blot, and the result showed that the amount of Tra2beta is strongly decreased in the mutant fraction in comparison to the

wild type one (figure 2.18), suggesting that the mutation disrupts in some manner the activity of an ESE.

All together these results indicate that the primary determinant of the T9 SPINK5 exon 11 exclusion from the final mRNA is not the disruption of an ASF/SF2-dependent exonic splicing enhnacer (as bioinformatics prediction suggested), but probably the combination between the reinforcement of a silencer and a concomitant disruption of an enhancer. The C/T AG deletion of three bases supports the presence of a bipartite exonic regulatory region with enhancer and silencer function.

Based on the data, I propose a model for wild type and T9 SPINK5 exon 11 splicing regulation (figure 3.2). SPINK5 exon 11 wild type it is not well defined by itself due to its weak 5'ss and contains two weak ESE regions located at position +12 and +17 that recruit the splicing factor Tra2β and a weak ESS at position +9. The C-to-T mutation in position +9 induces exon skipping through the strengthening of a silencer regulatory element located at the 5'-end of the exon, that binds to inhibitory splicing factor hnRNP A1, interfering with the nearby ESE sequence recognized by Tra2β. Several mechanisms have been proposed to explain hnRNP A1-mediated splicing repression (Mayeda and Krainer 1992, Burd and Dreyfuss 1994, Cartegni, Maconi et al. 1996, Chabot, Blanchette et al. 1997, Guil and Caceres 2007, Kashima, Rao et al. 2007, Okunola and Krainer 2009). In this case the most probable scenario is that hnRNP A1, bound to the newly created ESS, antagonizes by steric hindrance the binding of Tra2β at two adjacent weak ESEs located 4 nucleotides downstream the site of the mutation. Moreover, the hnRNP A1 binding site is very close to the 3'-splice site and this could interfere with the early spliceosome assembly by blocking the access to general splicing factors such as U2AF65, U2AF35 and U2 snRNP to the acceptor site. However, it is also possible that the initial binding of hnRNP A1 at the T9 sequence could nucleate the formation of a chain of inhibitory proteins along the exon. The creation of this "zone of silencing" has been previously described in HIV-1 *tat* exon 3 system (Amendt, 1995).

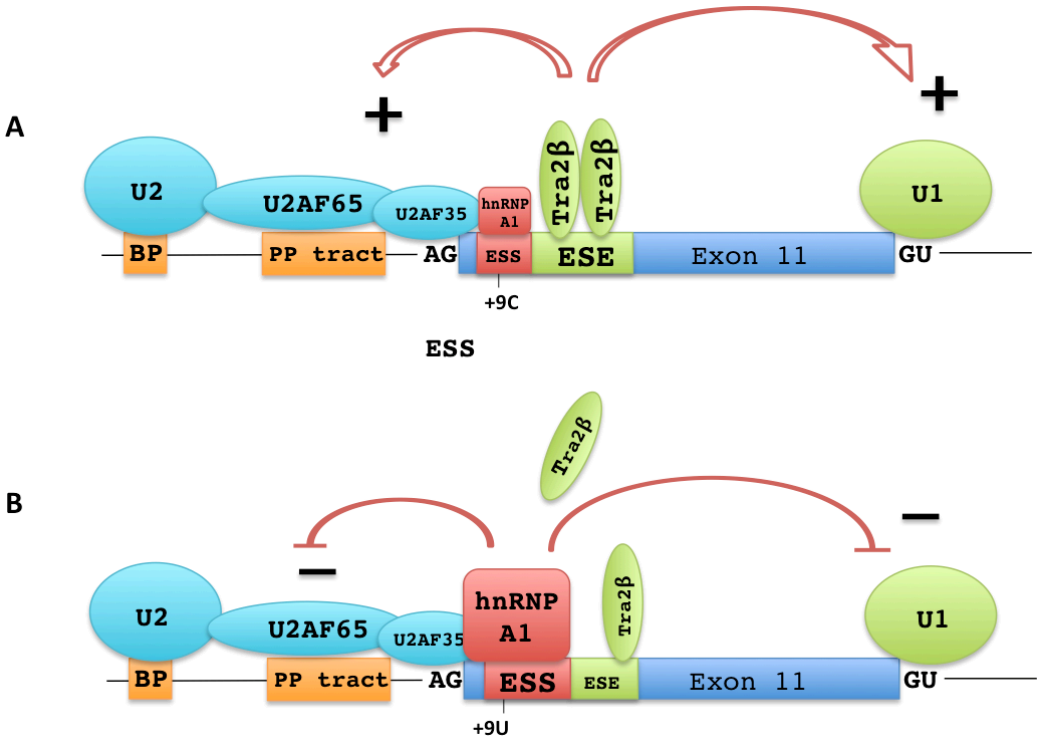


Figure 3.2. SPINK5 exon 11 +9C>T mutation reinforces an hnRNP A1-dependent ESS that impairs exon recognition.

A. Wild type SPINK5 exon 11. In wild type exon 11, two weak exonic regulatory sequences (ESEs) located nearby the 5'-end of the exon cooperate recruiting Tra2β, which probably contributes in the bridging of SR-protein over the exon and promotes its inclusion in the final SPINK5 mRNA.

B. +9C>T SPINK5 exon 11. The mutation strongly reinforces a silencer element (ESS) that efficiently recruits the inhibitory factor hnRNP A1 at the end of exon 11. This probably acts blocking the binding of Tra2β at the ESE located 3 nts downstream and interfering with the components of the spliceosome machinery that recognize the 3'-splice site, causing exon 11 skipping.

3.4 Molecular mechanism at the basis of ExSpe U1 mediated splicing rescue.

U1 small nuclear ribonucleoparticle (U1 snRNP) is the first element that binds to donor sites in the pre-messenger RNA molecule, promoting spliceosome assembly and, in turn, the beginning of the splicing reaction. The complementarity between its 5'-tail and the consensus sequence of the 5'ss is thought to be crucial for exon recognition, as mutation of the GU dinucleotide at the exon-intron border completely disrupts the splice site causing the skipping of the exon (Roca, Akerman et al. 2012, Roca, Krainer et al. 2013).

Our exon-specific U1s have a 5'-tail engineered to specifically target an intronic region downstream of the canonical donor site of an exon, promoting exon inclusion with the same efficacy of modified U1s fully complementary to the corresponding defective donor splice site. To understand the mechanism of ExSpeU1-mediated splicing rescue I performed a series of complementary experiments.

The first point I tried to address is the relationships between the ExSpe U1s and the endogenous U1 snRNP. EMSA experiments demonstrated that ExSpe U1s complex has the same electromobility of the normal U1 snRNP, strongly suggesting that ExSpe U1 and U1 snRNP recruit the same proteins and follow the same biogenesis pathway (figure 2.21). Starting from this consideration I was wondering if the ExSpe U1s could be able to functionally substitute the endogenous U1 particle, promoting the recognition of the exon and the spliceosome assembly. To verify this hypothesis, I focussed on SMN2 where sequestration of the endogenous U1 snRNP using a specific U1 decoy impairs SMN exon 7 inclusion (Roca, Akerman et al. 2012). In the absence of U1 snRNP at the 5' splice site, the ExSpe U1s were able *per se* to promote the recognition of the exon, the correct recruitment of the spliceosomal machinery and, in turn, the inclusion of SMN2 exon 7 in the final mRNA (figure 2.24). Thus, ExSpe U1s are able to functionally replace the endogenous U1 particle, suggesting that U1 snRNP does not necessarily have to bind at the 5' splice site, but its loading to a nearby region is sufficient to promote the spliceosome assembly and to promote the proper recognition of

the exon-intron borders. Some evidences that the base-pairing of the 5'-tail of the U1 snRNA and the donor site is not an insurmountable dogma was proposed at least for some exons, in which the recognition of the donor site sequence could happen also with one nucleotide-shifted basepairing with the tail of the U1 snRNA, creating a shift register (Roca and Krainer 2009, Roca, Akerman et al. 2012). Nevertheless, as for the proper identification of the splice junction ExSpe U1 does not have to directly interact with the 5'ss, probably other mechanisms or other components of the spliceosome machinery are fundamental to define the exon borders, whose recognition may occur independently from physical presence of the U1 snRNP exactly at the donor splice site.

To investigate the contribution of U1-specific proteins in ExSpe U1-mediated splicing correction, I prepared ExSpeU1 mutants. ExSpe U1s carrying mutations that abolished their interaction with the U1-A factor showed only a slight decrease in their capacity to rescue splicing, while mutations that avoid the recruitment of U1-70K on the ExSpe U1 RNA partially impaired its functionality (figure 2.22). The third U1-specific protein, U1-C, is thought to be part of the U1 complex through protein-protein interaction with U1-70K. It is possible that the ExSpe U1 70-K mutant does not bind also U1-C and, in turn, the strong decrease of splicing rescue efficacy may be due to the absence of both U1-70K and U1-C. Thus, ExSpeU1s rescue activity is probably due to the associated proteins that constitute the particle: U1-70K, in comparison to U1A seems to have the most important contribution. However, even if these mutations have been previously used in other systems to study the effect *in vivo* of U1 snRNPs, their binding specificity have been addressed only in vitro (Hamm, Dathan et al. 1990, Heinrichs, Bach et al. 1990, Nelissen, Will et al. 1994, Gunderson, Polycarpou-Schwarz et al. 1998, Du and Rosbash 2002, Hernandez, Makarova et al. 2009) . Accordingly, it is not clear if the mutant ExSpeU1s really lack, when transfected in cells, some of the U1-specific proteins. A better characterization of the structural composition of the mutants is required to establish their associated proteins.

Exon-specific U1s binding regions are located at different distances downstream the 5'ss of SMN2 exon 7 and SPINK5 exon 11. In the SMN context, some ExSpe U1s partially interact with the well-described hnRNP A1-dependent intronic splicing silencer (ISS) (Hua, Sahashi et al. 2011). Thus, for SMN2, the splicing rescue activity could be in part due to a masking effect of the ExSpeU1s on this hnRNP A1-dependent inhibitory regulatory region. Most of evidences presented in this thesis do not support a major effect of ExSpe U1s on the ISS.

The position of some ExSpe U1 partially overlaps with the final part of the ISS, while some others are located outside. The disruption of the 3'end of the ISS-N1 (motif 2) does not completely rescue exon inclusion and ExSpe sm17 and sm21 specifically interact with this motif (AG dinucleotide in position +23/+24)(Hua, Vickers et al. 2008). However, these two exon-specific U1s completely rescue exon 7 splicing. Moreover, the ExSpe U1s EE2 and EE3 target intronic regions outside of the ISS and completely rescue exon inclusion (figure 2.2) .

Modified U7 snRNAs with the same target sequences of ExSpeU1s do not rescue exon skipping. These molecules have been optimized and broadly used as antisense carriers to cover regulatory sequences and to modulate splicing in different gene models (Brun, Suter et al. 2003, Madocsai, Lim et al. 2005, Marquis, Meyer et al. 2007, Goyenvalle 2012). Using U7 snRNAs I showed that targeting the same intronic sequences downstream SMN2 exon 7 donor splice site does not recover exon inclusion (figure 2.23). Engineered U7s were correctly synthetized and effectively target the same region bound by the corresponding ExSpe U1, as in competition experiments U7s antagonize the positive effect of their corresponding ExSpe U1s in a dose dependent manner (figure 2.23).

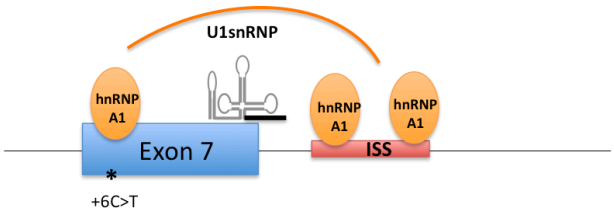
In SMN2, the AON and ExSpeU1 have a different mechanism to prevent exon skipping. The antisense oligonucleotide 10-27 that has been described as an efficient tool to redirect aberrant splicing in SMN2 exon 7 (Hua, Vickers et al. 2008, Hua, Sahashi et al. 2010, Hua and Krainer 2012). To investigate the effect of ExSpeU1s and AON 10-27 on SMN pre-mRNA processing, I analysed

the relative percentage of the two FL or Δ7 SMN isoforms, their total amount and the abundance of transcripts derived from the SMN genes. In the well-controlled HEK293 Flp-In system and in lentiviral-infected SMA type I fibroblasts, ExSpeU1s increased both the full-length SMN isoform and, to a lower extent, the isoform that lacks exon 7, suggesting that ExSpeU1s not only modulate splicing but also increase the total amount of mature SMN (figure 2.6 and 2.7). Notably, quantitative RT-PCR analysis showed that the stable clones expressing a single copy of ExSpeU1s have three times more SMN total transcripts in comparison with normal cells (figure 2.6). Interestingly, cells treated with AON 10-27 showed a similar increase in the percentage of SMN2 and SMN1 exon 7 inclusion (figure 2.6) but lacked of the specific effect on SMN isoforms and on transcripts: the oligonucleotide improves the FL isoform at the expense of Δ7 isoform (figure 2.6) and does not increase the amount of total SMN transcripts (figure 2.7). A possible explanation of this effect is that the oligonucleotide-mediated displacement of hnRNPA1 would simply affect the splicing decision and accordingly increase the exon inclusion at the expense of the skipped isoforms. On the contrary, ExSpeU1s could also affect other pre-mRNA processing steps or the stability of the final SMN transcripts. In addition, ExSpeU1s and AON have a completely different sensitivity to functional depletion of endogenous U1 snRNP and to overexpression of hnRNPA1. In fact, whereas inactivation of endogenous U1 snRNP has no effect on ExSpeU1- mediated splicing rescue, this inactivation completely abolishes the effect of the oligo, indicating that this U1 snRNP is required for the activity of the oligo, On the other hand, consistent with a "pure" antisense mechanism of the AON on the hnRNPA1-dependent ISS, the overexpression of hnRNPA1 does not affect the SMN2 splicing rescue of the oligonucleotide (figure 2.24). On the contrary, the splicing rescue mediated by ExSpeU1 was significantly reduced by hnRNPA1 overexpression.

The fact that ExSpeU1s and AON have different effects on isoforms, on SMN transcripts and are differently affected by U1 snRNP sequestration and by hnRNPA1 overexpression, strongly suggest that they act on the pre-mRNA in

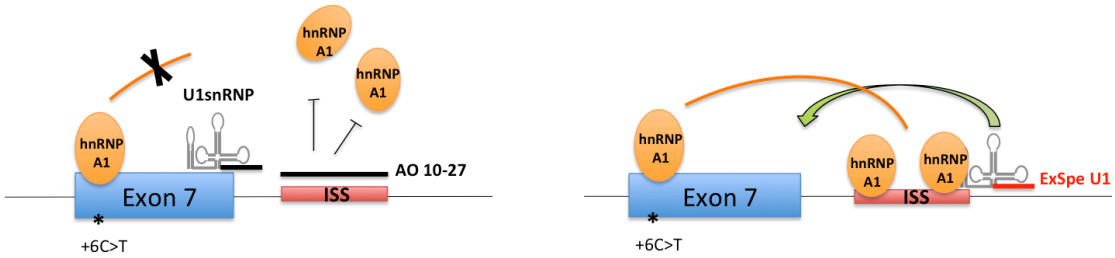
different manners. The AON 10-27 acts as a powerful antisense molecules binding to a bipartite hnRNPA1-dependent ISS downstream SMN2 exon 7 (Rigo, Hua et al. 2012),. On the contrary ExSpeU1s does not act as a mere antisense molecule but functionally replaces the endogenous U1 snRNP ameliorating exon definition.

ExSpe U1s may have a positive effect not only on splicing but, more in general, also on mRNA processing. In the last years several lines of evidence showed that the U1 particle is involved in different cellular processes, such as the protection of transcripts from premature polyadenylation signals (Kaida, Berg et al. 2010, Merkhofer and Johnson 2012) or promotion of transcription when bound to the first exon (Damgaard, Kahns et al. 2008). Thus, it does not seem unlikely that ExSpe U1s like normal U1 snRNPs, could promote transcription, stabilize the mRNAs or affect some cryptic polyadenylation sites. These events might be responsible of the peculiar positive effect of ExSpe U1s on the total amount of SMN transcripts and on the D7 transcript. However, more deep investigation will be required to better understand how ExSpeU1s affect processing of pre mRNA transcrips.



A Interruption of a silencing network

B Improvement of exon definition



3.3 ExSpe U1-mediated exon 7 definition in SMN2 pre-mRNA.

Upper panel: SMN2 exon 7 pre-mRNA. The C>T synonymous substitution at the 5'-end of the exon disrupts an SF2/ASF-dependent ESE and creates a novel hnRNPA1-dependent ESS that interacts with the ISS-N1 creating a silencing network along the exon that causes, in turn, its skipping from the final mRNA.

A. Antisense oligonucleotide 10-27 masks the ISS-N1 avoiding the interaction of inhibitory hnRNP A1 protein with the regulatory region. The silencing network is interrupted and the exon is included in the mature transcript.

B. ExSpe U1 targets an intronic region outside the ISS-N1, thus it does not interfere with the binding of hnRNP A1 inhibitory protein to the ISS. The binding of ExSpe U1 in a region downstream of the canonical 5'ss is able to promote exon definition also in presence of a silencing network, improving exon inclusion.

4. CONCLUSIONS & FUTURE PLANS

I studied two model genes, SMN2 exon 7 and SPINK5 exon 11 in which synonymous C-to-T substitutions relying on the 5'-end of the exons, concomitantly with the presence of a weak 5'ss, lead to their skipping from the mature transcripts. These mutations are associated to spinal muscular atrophy (SMA) and netherton syndrome (NS), two inherited pathology for which there are currently no cure. In this thesis work, I investigated the efficacy of a novel approach to modulate splicing and to correct exons skipping defects.

In minigene splicing assays, I demonstrated that increasing the complementarity of the 5'-tail of the U1 snRNAs to a weak donor splice site improves exon inclusion in the final mRNA for two different model exons carrying exonic mutations. Moreover, I bioengineered U1 snRNAs –named exon specific U1s (ExSpe U1s)- that target non conserved intronic regions downstream the canonical 5'ss of SMN2 exon 7 and SPINK5 exon 11. Using minigene models, I performed a screening of ExSpe U1s in SMN2 exon 7 and SPINK5 +9C>T exon 11 to select the most promising molecules and I tested them in primary fibroblasts and keratinocytes from patients with NS. Lentiviral mediated delivery of ExSpe U1s in those cells recover the correct full length mRNAs and rescue the physiological levels of functional protein to a level present in normal cells, providing the first proof of concept that ExSpe U1s could represent a reliable therapeutic approach for these pathologies. Moreover, I showed that a single copy of ExSpe U1 gene is sufficient to modulate SMN2 exon 7 pre-mRNA processing not only directing the splicing decision, but also increasing the levels of the spliced isoforms.

Contrarily to previously reported bioinformatics predictions (Lacroix, Lacaze-Buzy et al. 2012). I do not found SF2/ASF as the key splicing factor involved in SPINK5 exon 11 splicing. Through pull down, overexpression and

silencing experiments I individuated a bipartite splicing regulatory element in the 5'-end of exon 11 that interact with hnRNPA1 and Tra2beta, whose balance is altered by the c.891C>T mutation that cause, in turn, exon exclusion from the final mRNA. Moreover, additional studies on SPINK5 pre-mRNA processing in physiological and pathological conditions should be performed, in particular to unravel the interesting concern about a possible decrease in NMD degradation upon ExSpe U1 treatment. Quantitative RT-PCR experiments should be done on total SPINK5 mRNA in order to define the contribution of ExSpe U1s in rescuing mRNA stability.

I investigated both the ExSpe U1 composition and the molecular mechanism underlying ExSpe U1-mediated splicing modulation activity. Experimental data demonstrate that ExSpe U1s are assembled into ribonucleoparticles that are similar to the endogenous U1 snRNP. ExSpe U1 mutants for U1-specific proteins showed a reduced efficacy when the interaction with U1-A and U1-70K is abolished, indicating that these proteins have a role in ExSpe U1-mediated splicing rescue. Furthermore, the data suggest that, at least in the case I have analysed, ExSpe U1s do not act as antisense carriers, but possess the same functionality of the U1snRNP, although they bind downstream of the canonical 5'-splice sites to less conserved regions, and are able to promote splicing *per se*, also when the endogenous U1 particle is not available at the canonical splice site.

Additional studies will be required to understand the composition of the ExSpe U1 particle and the role of U1-specific associated proteins. These concerns could be achieved through affinity purification experiments of wild type and mutated ExSpe U1s and western blot to detect specific U1-associated proteins. The role of each protein that interacts with ExSpe U1 RNA have to be evaluated through appropriate assays. For instance new different mutant ExSpe U1s could be prepared to test the role of U1-associated proteins in ExSpe U1 biogenesis, splicing rescue and pre-mRNA processing.

From a therapeutic poin of view, ExSpe U1s could represent a novel approach to correct splicing-associated diseases. SMA is an extensively studied

pathology for which several mouse models have been developed, accordingly to different degrees of severity of the disease (Corti, Nizzardo et al. 2008, Williams, Schray et al. 2009, Corti, Nizzardo et al. 2010, Gogliotti, Hammond et al. 2010, Passini, Bu et al. 2010). To test the efficacy of ExSpe U1s *in vivo* it will be useful the creation of an ExSpe U1 transgenic mice to breed, for instance, with the most severe SMA mouse model (Smn-/-, hSMN2+/+, lifespan ~6 days) in order to evaluate the phenotypic rescue and the possible lifespan extension of the animal. The effects of ExSpe U1s in SMA-ExSpe transgenic mice should be evaluated also in all the organs both by semiquantitative and quantitative evaluation of SMN2 FL and D7 isoforms. In addition, transgenic ExSpe U1 viability will provide immediate evidences of ExSpe U1 toxicity *in vivo.*

As it has not been clearly established the organs and the specific cells that have to be targeted in order to correct SMA phenotype, it could be also interesting to use adenoviral vectors with different serotypes to specifically target only some organs. For instance, scAAV serotype 9 has been recently used for gene replacement to deliver SMN1 gene in mice through intra-muscular administration and it has been demonstrated to efficiently bypass the brain-blood barrier, targeting the brain (Benkhelifa-Ziyyat, Besse et al. 2013).

In NS pathology the most promising preliminary approach *in vivo* could be represented by creation of NS skin-grafts in mice. Correction, in this case, could be achieved both by the usage of AAV or lentiviral vectors, whose safety in clinical trials has been recently pointed out (Aiuti, Biasco et al. 2013).

5. MATERIALS and METHODS

Chemical reagents.

General chemicals and kits were purchased from Sigma, Merck, Gibco, Boehringer-Mannheim, Invitrogen, Fluka, Riedel de Haen, Qiagen, Stratagene.

General solutions.

All solutions are identified in the text except for the following:

TE: 10 mM Tris-HCl (pH 7.4), 1 mM EDTA (pH 7.4)

10xTBE: 108 g/l Tris-HCl, 55 g/l Boric acid, 9.5 g/l EDTA

6X DNA sample buffer: 0.25 % w/v bromophenol blue, 0.25 % w/v xylene cyanol FF, 30 % v/v glycerol in H_20.

5X Running Buffer: 30 g Tris-HCl, 144 g glycine, 5 g SDS in 1L

1X Blotting buffer: 3.03 g Tris-HCl, 14.4 g glycine, 20% methanol in 1L

10X Protein sample buffer: 20 % w/v SDS, 1 M DTT, 0.63 M Tris-HCl (pH 7), 0.2 % w/v bromophenol blue, 20 % v/v glycerol, 10 mM EDTA (pH 7).

PBS 1X: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH7.4

Synthetic oligonucleotides.

Synthetic DNA oligonucleotides were purchased from Sigma.

Radioactive isotopes.

Radioactive γ-^{32}P-ATP was supplied by PerkinElmer.

Bacterial cell culture.

The *E. coli* K12 strain DH5α was transformed with the plasmids described in this study and used for their amplification. Plasmids were maintained in the short term as single colonies on agar plates at 4 °C but for long term storage

they were kept on glycerol stocks made by adding sterile glycerol to a final 30% v/v concentration to liquid bacterial cultures. Glycerol stocks were stored at –80°C. When necessary, from the glycerol stocks an overnight culture of bacteria was grown in Luria-Bertani medium [LB medium: per litre: 10 g Difco Bactotryptone, 5 g Oxoid yeast extract, 10 g NaCl, (pH 7.5)]. Bacterial growth media were sterilized before use by autoclaving. When needed, ampicillin was added to the media at a final concentration of 100 μg /ml.

Preparation of bacterial competent cells.

Bacterial competent cells were prepared following the method described by Chung and Niemela (Chung, 1989). *E. Coli* strains were grown overnight in 3 ml of LB at 37°C. The following day, 300 ml of fresh LB were added and the cells were grown at room temperature for 4-5 h until the OD_{600} was 0.3-0.38. The cells were then put in ice and centrifuged at 4 °C and 1000g for 15 min. The pellet was resuspended in 30 ml of cold TSS solution (10% w/v PEG molecular weight 4000, 5% v/v DMSO, 35mM Mg Cl2, pH 6.5 in LB medium). The cells were aliquoted, rapidly fronzen in liquid nitrogen and stored at –80°C. Competence was determined by transformation with 0.1 ng of pUC19 and was deemed satisfactory if this procedure resulted in more than 100 colonies.

Transformation of bacteria.

Transformations of ligation reactions were performed using 1/2 of the reaction volume. Transformation of clones was carried out using 1 ng of the plasmid DNA. The DNA was incubated with 60 μl of competent cells for 20 min on ice and at 42°C for 2 minutes. After the step of the heat shock, 60 μl of LB were added and the bacteria allowed to recover for 10 min at 37 °C. The cells were then spread onto agarose plates containing the appropriate antibiotic. The plates were then incubated for 12-15 hours at 37 °C.

When DNA inserts were cloned into β-galactosidase-based virgin plasmids, 30 μl of IPTG 100 mM and 20 μl of X-Gal (4 % w/v in dimethylformamide) were spread onto the surface of the agarose before plating to facilitate screening of positive clones (white colonies) through identification of β-galactosidase activity (blue colonies) which indicates the negative clones.

Small scale preparation of plasmid DNA from bacterial cultures.

Small amounts of bacterial plasmid DNA was obtained as described in "Molecular cloning: A laboratory manual" (Maniatis, Cold Spring Harbor Laboratory Press). Briefly, alkaline lysis of recombinant bacteria was performed by resuspending the bacterial pellet in 200 μl of ddH2O; 200 μl of solution II (0.2 M NaOH, 1 % w/v SDS) were then added and the contents mixed by inversion. 300μl of solution III (3 M potassium acetate pH 5.2) were then added and the contents mixed by inversion. The bacterial lysate was then centrifuged in an Eppendorf microcentrifuge at maximum speed and the supernatant transferred to a new tube. An equal volume of 1:1 v/v phenol:chloroform solution was added to the supernatant. The tube was then vortexed and centrifuged as above. The aqueous phase containing the DNA was transferred to a new tube. An equal volume of chloroform was added to the supernatant. The tube was then vortexed and centrifuged as above. The aqueous phase containing the DNA was then recovered and the DNA pelleted by ethanol precipitation. The final pellet was resuspended in 50 μl of ddH2O +3μl of RNAse A (10mg/ml) for at least 1h at 37^{o}C and 5 μl of such preparation were routinely taken for analysis by restriction enzyme digests.

Large scale preparations of plasmid DNA from bacterial cultures.

For large-scale preparations of plasmid DNA that was necessary for the transfection experiments, JetStar purification kit (Genomed) was used according to the manufacturer's instructions. In order to get a good amount of plasmid, we used a 50 ml of overnight bacterial culture using LB medium.

RNA preparation from cultured cells.

Cultured cells were washed with PBS and then RNA Trizol (Invitrogen) was added, vortexed and centrifuged at max speed for 20' at 4°C. Aqueous phase was recovered and then chloroform extraction was performed. Supernatant was precipitated with isopropanol. The pellet was resuspended in 50 μl of ddH2O and digested with 1U of DNase RNase free. The mix was incubated at RT for 30 minutes, and then the RNA was purified by acid phenol extraction. The final pellet was resuspended in 35 μl of ddH2O and frozen at –80 °C. The RNA quality was checked by electrophoresis on 1% agarose gels and quantified using Nanodrop.

Enzymatic modification of DNA.

Restriction enzymes

Restriction enzymes were purchased from New England Biolabs, Inc. (USA) or Pharmacia Biotech (Sweden). All buffers were also supplied by the same company and were used according with the manufacturer's instructions. For many analytical digests we also used 0.5X, 1X or 2X concentration of 10x OPA buffer (100 mM Tris-acetate (pH 7.5), 100 mM magnesium acetate and 500 mM potassium acetate). For analytical digests 100-500 ng DNA were digested in a volume of 20 μl containing the appropriate U of the restriction enzyme per μg DNA. The digest was incubated for 2-4 hours at the optimal temperature required by the enzyme used. Preparative digestions of vectors and inserts were made of 5-10 μg DNA using the appropriate condition needed by the restriction enzyme in 100-200 μl reaction volume. Enzymatic activity was then removed either by heat inactivation or by phenol-chloroform extraction.

DNA Polymerase I, Large (Klenow) fragment.

Klenow enzyme was used to treat PCR products for blunt-end creation by 3' overhang removal and 3' recessed end fill-in. Briefly the DNA was dissolved in any 1X restriction enzyme NEBuffer or 1X EcoPol Reaction Buffer and

supplemented with 33 µM each dNTP if "fill-in" was required (DNA fragments with protruding 3'ends). Then, 1U of Klenow per µg DNA was added and the mixture was incubated 15 minutes at RT. The reaction was inactivated by by heating at 75°C for 20 minutes.

T4 Polynucleotide Kinase.

T4 polynucleotide kinase (New England Biolabs Inc) catalyzes the transfer and exchange of phosphate from ATP to the 5´ -hydroxyl terminus of double- and single-stranded DNA and RNA to allow subsequent ligation or end-label for probes production. The proper units of T4 polynucleotide kinase, 1X of its reaction buffer (70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol) and ATP were added to the DNA and incubated at 37°C for 30 minutes. The enzyme was inactivated by incubation at 65°C for 20 min.

Calf intestinal phosphatase (CIP).

Calf intestal phosphatase (CIP), provided from New England Biolabs Inc, catalyzes the removal of 5´ phosphate groups from DNA and RNA. Since CIP-treated fragments lack the 5´ phosphoryl termini required by ligases, they cannot self-ligate. This property was used to decrease the vector background in cloning strategies. The standard reaction was carried out in a final volume of 50-100 µl using 1U of enzyme per 0.5 µg DNA at 37°C for 1 hour.

T4 DNA ligase

T4 DNA ligase catalyzes the formation of phosphodiester bonds between neighbouring 3´-hydroxyl- and 5´-phosphate ends in double-stranded DNA, requiring ATP as a cofactor in this reaction. This enzyme was purchased from Roche Diagnostic and was used to join double stranded DNA fragments with compatible sticky or blunt ends. A standard reaction comprises 20 ng of linearised vector and 5-10 fold molar excess of insert in a total volume of 20 µl containing 1X ligase buffer and 1U of enzyme. Reactions were carried out at room temperature for at least 2-4 hours. In some ligations synthetic oligonucleotides were used as inserts for the reactions. In these cases,

amounts added to each reaction to obtain ligation of the oligonucleotides in the resulting plasmids were about 50-100 fold molar excess over the DNA vector.

Agarose DNA gel electrophoresis.

DNA samples were size fractionated by electrophoresis in agarose gels ranging in concentrations from 0.8 % w/v (large fragments) to 2 % w/v (small fragments). The gels contained ethidium bromide (0.5 μg /ml) and 1X TBE. Horizontal gels were routinely used for fast analysis of DNA restriction enzyme digests, estimation of DNA concentration, or DNA fragment separation prior to elution from the gel. Samples of 20 μl containing 1X DNA loading buffer were loaded into submerged wells. The gels were electrophoresed at 50-80 mA in 1X TBE running buffer for a time depending on the fragment length expected and gel concentration. DNA was visualized by UV transillumination and the result recorded by digital photography.

Elution and purification of DNA fragments from agarose gels.

This protocol was used to purify small amounts (less than 1 μg) of DNA for sub-cloning. The DNA samples were electrophoresed onto an agarose gel as described previously. The DNA was visualized with UV light and the required DNA fragment band was excised from the gel. This slab was cut into pieces, and the JETquick Spin Column Technique (Genomed) was used according to the manufacturer's instructions. Briefly, 300 μl of gel solubilisation solution L1 ($NaClO_4$, Na acetate and TBE) were added for each 100 mg of the gel slice pieces and incubated at 55 ˚C for 15 min vortexing every 5 min. The mixture was loaded into a prepared JETquick column and it was centrifuged at maximum speed for 1 min. The flowthrough was discarded. 700 μl of washing and reconstituted solution L2 (ethanol, NaCl, EDTA and Tris-HCl) were added into the spin column and after 2 min, the column was centrifuged in the same conditions twice. The flowthrough was again discarded both times. To elute the bound DNA, 30-50 μl of pre-warmed sterile water were

added onto the centre of the silica matrix of the spin column and the system was centrifuged for 2 min. The amount of DNA recovered was approximately calculated by UV fluorescence of intercalated ethidium bromide in an agarose gel electrophoresis.

Sequence analysis for cloning purposes.

Sequence analysis of plasmid DNA were performed in Macrogen Inc. The sequences were analysed using FinchTV and Strider programs.

The primers used for sequencing are listed below.

For the pUC19 plasmid:

M13-Fw	5'-GTTTCCCAGTCACGAC-3'
M13-Rev	5'-GGAAACAGCTATGACCATG-3'

For the pcDNA5 FRT/TO plasmid:

CMV-Fw	5'- CGCAAATGGGCGGTAGGCGTG -3'
BGH Rev	5'- TAGAAGGCACAGTCGAGG-3'

Hybrid minigene constructs.

Transient transfection of minigenes is commonly used *in vivo* technique to identify the different features of exon regulation during the splicing process. It is a good method to study specific *cis*-acting elements that control constitutive and alternative exons, the cell specific splicing pattern and also to identify *trans*-acting factors that recognize these elements and regulates splicing (Cooper, 2005). Any region of interest is introduced in a basic minigene, thus creating a new construct, a hybrid minigene. The genomic segment was generated by PCR amplification and the oligonucleotides are designed to have restriction sites at their ends to match those of the recipient plasmid. This system, after transient transfection and RNA analysis, allows us to study the splicing outcome of the minigene.

The minigenes used in this thesis work are pCI SMN2 and pSPINK5 wt, mutant and deltaTAG. The pCI SMN2 palsmid has been developed in dr. Adrian Krainer Lab (Hua et al., 2006) and pSPINK5 wt, mutant and deltaTAG

in the Molecular biology laboratory at "La Sapienza" University, Rome (Fortugno et al., 2012).

Plasmid for modified-U1 expression.

The vector used for the production of U1 snRNAs was the pGEM3, a standard cloning vector. The *wt* U1 snRNA gene was cloned in *BamH*I restriction site and the region between *Bgl*II and *Bcl*I sites encoding for the 5' tail complementary to 5' ss was replaced with specific annealed oligos. The identity of all modified U1 snRNAs expression vectors was ultimately confirmed through sequencing analysis.

Denaturing polyacrylamide gel electrophoresis (SDS-PAGE).

Protein samples were added to protein sample buffer (2X final). Conventional slab gel SDS PAGE (Laemmli, 1970) was performed in vertical gels with the required percentage of polyacrylamide (37,5:1 acrylamide:bis-acrylamide, ProtoGel, National Diagnostics). The gels were run at 40 mA in 1X SDS-PAGE running buffer (50 mM Tris, 0.38 M glycine, 0.1 % w/v SDS). Gels were stained with coomassie Blue R250 in methanol-water-acetic acid 45:45:10 (v/v/v).

Western blotting

After the proteins were separated on SDS-PAGE gels, they were electroblotted onto PVDF membrane (Amersham Biosciences) according to standard protocols and the membrane was blocked by incubation in 5% milk in 1X PBS buffer. Membranes were probed with different primary antibodies, then incubated with secondary antibodies conjugated with horseradish peroxidase enzyme and the resulting immune-reactive bands were detected with enhanced chemiluminescence reagent (ECL; Amersham Biosciences).

Antibodies

The antibodies used in western blot analysis and their dilutions are:

Anti-SMN from Santa Cruz Biotechnologies	1:2000
Anti-Lekti (courtesy of dr. Fortugno, University La Sapienza, Rome)	1:1000
Anti-SF2 (ab38813) from Abcam	1:1000
Anti-Tra2beta (ab31353) from Abcam	1:1000
Anti-hnRNPA1/A2 kindly provided by dr. Muro, ICGEB Trieste	1:1000
Anti-Tubulin kindly provided by dr. Muro, ICGEB Trieste	1:5000

EMSA

RNA synthetic oligonucleotides (200 ng) were labelled by phosphorylation with gamma-32P ATP and T4 polynucleotide kinase for 1h at 37C, precipitated and resuspended in 200 uL of water. Each binding reaction was made mixing the Hek293 nuclear extract from cells transfected with appropriate plasmid encoding ExSpe U1s with the labelled RNA oligo in 1X binding buffer to a final volume of 20 uL. The reaction was left at room temperature for 15 minutes before loading the sample on a native polyacrylamide gel (5%) in 0.5X TBE buffer which was run at 100V at +4C for 3h. The protein-nucleic acid complexes were visualized using biomax MS films (Kodak).

Eukaryotic cell lines.

The cell line used for siRNA, transfection and cotransfection experiments was Hek293 cells. SMA type I fibroblasts (G3813) and SMA control fibroblasts (G3814) were purchased from Coriell Institute Repository. Primary NS keratinocytes from a patient carrying the c.891C>T mutation were obtained from Molecular Biology Laboratory, University “La Sapienza”, Rome.

Maintenance and analysis of cells in culture.

Hek293 and SMA primary fibroblasts cells were grown in Dulbecco’s Modiefied Eagle Medium (with glutamine, sodium pyruvate, pyriodoxine and 4.5 g/l glucose) supplemented with 10% fetal calf serum (Euro Clone) and

antibiotic Antimycotic (Sigma) according to the manufacturer's instruction. A standard 100mm dish containing a confluent monolayer of cells was washed with 1XPBS solution, treated with 1-2 ml Trypsin (PBS containing 0.045 mM EDTA and 0.1% trypsin) and incubated at 37°C for 2 minutes or until cells were dislodged. After adding 5 ml of media, cells were precipitated by centrifugation and resuspended in pre-warmed medium. A subcultivation ratio of 1:4 to 1:6 was used for SMA fibroblasts and a ratio of 1:6 to 1:8 was performed for Hek293. 1 ml of this cell dilution was added to 10 ml fresh medium and plated in a new 100 mm dish.

Transfection of recombinant DNA.

Hek293 cells were transfected or, when indicated, co-transfected in 6-well plates (about 5x10^5 cells/well) by standard calcium phosphate technique (Sambroock et al., 1989).

Lentiviral particle production.

Lentiviral particles encoding eGFP and U1snRNA wt or ExSpe sm2 or sm17 or sm21 were prepared through cotransfection in Hek293T cells of psPAX vector encoding for gag-pol, tet and rev proteins, pVSV-G encoding packaging proteins and modified pLVTHM plasmid encoding the green fluorescent protein (GFP) under the cytomegalovirus (CMV) promoter and U1snRNA WT or SMA-/NS-specific ExSpe U1s gene under their own promoter. Transfection efficiency was monitored through GFP evaluation at FACS. The viral supernatant was collected from transfected cells after 48h and concentrated by ultracentrifugation at 25000 rpm for 90 minutes. The viral pellet was resuspended in PBS 1X, flash frozen in liquid nitrogen and aliquots were stored at -80° until use. SMA fibroblasts were grown on six well plates and transduced with $4x10^6$ lentiviral particles with a molteplicity of infection (MOI) of 15 for 12 h. The titer of the virus was $4x10^8$ bTU/mL. The biological titer was estimated through transduction of Hek293 cells with serial dilution of the lentivirus and evaluation of the percentage of GFP-positive cells at confocal microscope.

cDNA synthesis.

In order to synthesize the first-strand cDNA, 3-5 μg of total RNA extracted from the cells were mixed with the following components: 200 ng of Random Primers (Pharmacia), 1 mM dNTPs mix and strile water to reach the final volume of 12 μl. The mixture was then denaturated at 65°C for 5 minutes and quick chilled on ice. After denaturation specific solutions, purchased from Invitrogen, were added to the reaction: 1X first strand buffer (10 mM Tris-HCl ph8.4, 50 mM KCL, 2.5 mM $MgCl_2$) 10 mM DTT, 200 U of Moloney Murine Leukemia Virus reverse transcriptase and 20 U of RNAse inhibitor (Ambion). The final mixture was then incubated at least 1 hour at 37°C. 10% of the first strand reaction (2 μl of the cDNA) was used as PCR template.

PCR analysis

The polymerase chain reaction was performed on genomic or plasmid DNA following the basic protocols of the Roche Diagnostic *Taq* DNA Polymerases. The volume of the reaction was 50 μL. The reaction buffer was: 1X *Taq* buffer, dNTP mix 200 μM each, oligonucleotide primers 20 μM each, *Taq* DNA Polymerase 2.5 U. As DNA template, 0.1 ng of plasmid or 100-500 ng of genomic DNA were used for amplification. When a DNA fragment longer than 2000 bp was amplified, DMSO 3% was also added to the mixture. The amplification conditions are described for each particular PCR. The amplifications were performed on a Cetus DNA Thermal Cycler (Perkin Elmer) or on a Gene Amp PCR System (Applied Biosystems).

Oligonucleotides used fro PCR analysis were those specific for the minigene system used or for the endogenous RNA of interest.

For pCI-SMN2 minigene:

pCI-Fw 5’-CATGACGATCCGTAGGCATA -3’

E8-75 Rev 5’-TAAAGCGATGAAAGCGAGGCAT-3’

For pSPINK minigene:

SA2-Fw 5’-ATCTCAGTGGTATTTGTGAGC-3’

SD6-Rev 5’-TCTGAGTCACCTGGACAACC-3’

For endogenous SMN1 and/or SMN2 mRNAs in Hek293 and SMA type I fibroblasts (G3813; G3814 Coriell Institute):

FAM-E6-Fw 5'-FAM-ATAATTCCCCCACCACCTCCC -3'

E8-467-Rev 5'-TTGCCACATACGCCTCACATA -3'

Quantitative RT-PCR

TaqMan-based and Sybr green-based qRT-PCR were carried out as previously described (Sumner et al., 2006; Pastor et al., 2011).

Primers and probe for amplification of SMN delta7 isoform:

Delta7 Fw 5'-CATGGTACATGAGTGGCTATCATACTG-3'

Delta7 Rev 5'-TGGTGTCATTTAGTGCTGCTCTATG-3'

Delta7 probe 6-FAM-CCAGCATTTCCCATATAATAGC-MGB

Primers and probe for amplification of SMN full length (FL) isoform:

FL fw 5'-CAAAAAGAAGGAAGGTGCTCACATT-3'

FL rev 5'-GTGTCATTTAGTGCTGCTCTATGC-3'

FL probe 6-FAM- CAGCATTTCTCCTTATTTA-MGB

Primers sequence for Sybr green-based qRT-PCR for evaluation of total SMN transcripts:

Ex4-5 fw 5'-GGACAAGGAAAGCCAGGTCT-3'

Ex5-6 rev 5'-TGGGGAATTATTGGTGGTCCA-3'

GAPDH fw 5'-GACAGTCAGCCGCATCTTCT-3'

GAPDH rev 5'-TTAAAAGCAGCCCTGGTGAC-3'

Preparation of samples for capillary electrophoresis

RNA extracted from cells was amplified using E8-467 and a labelled fluorescent E6-Fw (FAM) primers. The PCR products were diluted 1:100 and then 1 μL of DNA was dehydrated at the temperature of 60°C for 20 min.

In silico predictions.

In silico analysis was performed using the free access ESEfinder, RESCUE-ESE and Human Splicing finder.
(http://rulai.cshl.edu/tools/ESE/)

RNA binding protein analysis.

Protein samples were added to protein sample buffer (2X final). Conventional slab gel SDS PAGE (Laemmli, 1970) was performed in vertical gels with the required percentage of polyacrylamide (37.5:1 acrylamide:bis-acrylamide, ProtoGel, National Diagnostics), depending on each case. The gels were run at 40 mA in 1X SDS-PAGE running buffer. After running, gels were either stained with coomassie Blue R250 in methanol-water-acetic acid 45:45:10 (v/v/v) or electroblotted onto PVDF membrane (Amersham Biosciences) for Western Blot analysis.

Affinity purification of RNA binding proteins.

Two synthetic RNA oligos were generated by Integrated DNA Technologies and used as targets for pull down assays.

- SPINK5 Exon 11 WT

5'-agAAACUCUGCAGUCAAUAUCAAAAUC -3'

- SPINK5 exon 11 +9C>T

5'- agAAACUCUGUAGUCAAUAUCAAAAUC -3'

Briefly, 12 µg of target SPINK5 RNA oligos were placed in 400 µl of reaction mixture (100 mM NaOAC, pH 5.0 and 5mM sodium m-periodate), incubated for 1 h in the dark at room temperature, ethanol-precipitated and finally resuspended in 100 µl of 100 mM NaOAC (pH 5.0). Approx. 400 µl of adipic acid dehydrazide agarose beads 50% slurry (Sigma) previously equilibrated with 100 mM NaOAC (pH 5.2) were added to each periodate-treated RNA and the mix was incubated for 12 h at 4°C on a rotator. The beads with the bound

RNA were then washed two times with 1 ml of 2 M NaCl, and equilibrated in 1X washing buffer (5.2 mM HEPES at pH 7.5, 1 mM $MgCl_2$, 0.8 mM MgAcetate). Then the beads were incubated, in a final volume of 500 µl, with 0.5 mg of HeLa cell nuclear extract (C4, Biotech), 1X binding buffer (5.2 mM HEPES, pH 7.9, 1 mM $MgCl_2$, 0.8 mM MgAcetate, 0.52 mM DTT, 3.8% glycerol, 0.75 mM ATP, 1mM GTP) and Heparin (final concentration 1 µg/µl), for 30 min on a rotator, at room temperature. The beads were then washed four times with 1.5 ml of washing buffer before addition of SDS sample buffer and loading on SDS-10% polyacrylamide gels. Proteins were visualized by Coomassie brilliant blue staining as described above. Protein sequence analysis of the bands excised from the gel was performed using an electrospray ionization mass spectrometer (LCQ DECA XP-ThermoFinnigam). Protein bands were digested with trypsin and the resulting peptides were extracted with water and 60% acetonitrile/1% trifluoroacetic acid. Fragments were then analyzed by mass spectrometry and proteins were identified by analysis of the peptide MS/MS data with Turbo SEQUEST (ThermoFinnigam) and MASCOT (Matrix Science).

Small interfering RNA (siRNA) transfection.

siRNA transfections were performed in Hek293 cells using Oligofectamine Reagent (Invitrogen). The sense strands of RNAi oligos (Dharmacon) used for silencing the different target proteins were the following:

human hnRNP A1, 5'-CAGCUGAGGAAGCUCUUCA-3', (Kashima, 2003);
human hnRNP A2, 5'-GGAACAGUUCCGUAAGCUC-3' (Kashima, 2003);
luciferase #2 gene control, 5'-GCCAUUCUAUCCUCUAGAGGAUG-3'.
Human SF2/ASF 5'-ACGAUUGCCGCAUCUACGU-3' (Cartegni, Hastings et al., 2006).

Hek293 were plated at 2.5 x 10^5 cells per well in 60 mm plates to achieve 40–50% confluency. The next day, 6 µl Oligofectamine was combined with 24 µl of Opti-MEM medium (Invitrogen) and 5-10 µl of 40µM siRNA duplex oligos

were diluted in a final volume of 400 µl of Opti-MEM medium. The two mixtures were combined and left for 20 min at RT. Finally, this mix was added to the cells, which were maintained in 1.6 ml of Opti-MEM. After 24h a second round of siRNA transfection was performed as described above. Six to eight hours later Opti-MEM was exchanged with DMEM medium without antibiotic and the cells were transfected with the minigene of interest using Qiagen Effectene transfection reagents. One µg of DNA was mixed with 8 µl of Enhancer for each transfection and the mixture was incubated at room temperature for 5 minutes to allow the condensation of the DNA. Then, 10 µl of Effectene were added to the mixture and an incubation of 10 minutes has been performed. After the addition of 500 µl of complete culture DMEM medium, the mixture was added to the cells in 4 ml of the same medium and incubated at 37°C. After 12 h, HeLa cells were harvested and divided in two parts for protein and RNA extractions. RT-PCR from total RNA was performed as described above in the transfection protocol. Whole protein extracts were obtained by cell sonication in lysis buffer (15 mM HEPES, pH 7.5, 250 mM NaCl, 0.5% NP40, 10% glycerol and 1 mM PMSF) and analyzed for hnRNP A1, A2 and ASF/SF2 endogenous protein expression by immunoblot analysis using specific monoclonal antibodies. Tubulin was used as protein loading control. Each siRNA treatment experiment was repeated at least three times.

6. REFERENCES

Aartsma-Rus, A., I. Fokkema, J. Verschuuren, I. Ginjaar, J. van Deutekom, G. J. van Ommen and J. T. den Dunnen (2009). "Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations." Hum Mutat **30**(3): 293-299.

Aebi, M., H. Hornig and C. Weissmann (1987). "5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU." Cell **50**(2): 237-246.

Aiuti, A., L. Biasco, S. Scaramuzza, F. Ferrua, M. P. Cicalese, C. Baricordi, F. Dionisio, A. Calabria, S. Giannelli, M. C. Castiello, M. Bosticardo, C. Evangelio, A. Assanelli, M. Casiraghi, S. Di Nunzio, L. Callegaro, C. Benati, P. Rizzardi, D. Pellin, C. Di Serio, M. Schmidt, C. Von Kalle, J. Gardner, N. Mehta, V. Neduva, D. J. Dow, A. Galy, R. Miniero, A. Finocchi, A. Metin, P. P. Banerjee, J. S. Orange, S. Galimberti, M. G. Valsecchi, A. Biffi, E. Montini, A. Villa, F. Ciceri, M. G. Roncarolo and L. Naldini (2013). "Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome." Science **341**(6148): 1233151.

Amendt, B. A., D. Hesslein, L. J. Chang and C. M. Stoltzfus (1994). "Presence of negative and positive cis-acting RNA splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1." Mol Cell Biol **14**(6): 3960-3970.

Ashkenas, J. (1997). "Gene regulation by mRNA editing." Am J Hum Genet **60**(2): 278-283.

Baillat, D., M. A. Hakimi, A. M. Naar, A. Shilatifard, N. Cooch and R. Shiekhattar (2005). "Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II." Cell **123**(2): 265-276.

Baralle, D., A. Lucassen and E. Buratti (2009). "Missed threads. The impact of pre-mRNA splicing defects on clinical practice." EMBO reports **10**(8): 810-816.

Barash, Y., J. A. Calarco, W. Gao, Q. Pan, X. Wang, O. Shai, B. J. Blencowe and B. J. Frey (2010). "Deciphering the splicing code." Nature **465**(7294): 53-59.

Beaudoing, E., S. Freier, J. R. Wyatt, J. M. Claverie and D. Gautheret (2000). "Patterns of variant polyadenylation signal usage in human genes." Genome Res **10**(7): 1001-1010.

Beelman, C. A. and R. Parker (1995). "Degradation of mRNA in eukaryotes." Cell **81**(2): 179-183.

Benkhelifa-Ziyyat, S., A. Besse, M. Roda, S. Duque, S. Astord, R. Carcenac, T. Marais and M. Barkats (2013). "Intramuscular scAAV9-SMN Injection Mediates Widespread Gene Delivery to the Spinal Cord and Decreases Disease Severity in SMA Mice." Molecular therapy : the journal of the American Society of Gene Therapy.

Benne, R., J. Van den Burg, J. P. Brakenhoff, P. Sloof, J. H. Van Boom and M. C. Tromp (1986). "Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA." Cell **46**(6): 819-826.

Bentley, D. L. (2005). "Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors." Curr Opin Cell Biol **17**(3): 251-256.

Berg, M. G., L. N. Singh, I. Younis, Q. Liu, A. M. Pinto, D. Kaida, Z. Zhang, S. Cho, S. Sherrill-Mix, L. Wan and G. Dreyfuss (2012). "U1 snRNP determines mRNA length and regulates isoform expression." Cell **150**(1): 53-64.

Berglund, J. A., K. Chua, N. Abovich, R. Reed and M. Rosbash (1997). "The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC." Cell **89**(5): 781-787.

Bernstein, L. B., T. Manser and A. M. Weiner (1985). "Human U1 small nuclear RNA genes: extensive conservation of flanking sequences suggests cycles of gene amplification and transposition." Mol Cell Biol **5**(9): 2159-2171.

Biffi, A., E. Montini, L. Lorioli, M. Cesani, F. Fumagalli, T. Plati, C. Baldoli, S. Martino, A. Calabria, S. Canale, F. Benedicenti, G. Vallanti, L. Biasco, S. Leo, N. Kabbara, G. Zanetti, W. B. Rizzo, N. A. Mehta, M. P. Cicalese, M. Casiraghi, J. J. Boelens, U. Del Carro, D. J. Dow, M. Schmidt, A. Assanelli, V. Neduva, C. Di Serio, E. Stupka, J. Gardner, C. von Kalle, C. Bordignon, F. Ciceri, A. Rovelli, M. G. Roncarolo, A. Aiuti, M. Sessa and L. Naldini (2013). "Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy." Science **341**(6148): 1233158.

Birney, E., S. Kumar and A. R. Krainer (1993). "Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors." Nucleic Acids Res **21**(25): 5803-5816.

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

Blaustein, M., F. Pelisch, T. Tanos, M. J. Munoz, D. Wengier, L. Quadrana, J. R. Sanford, J. P. Muschietti, A. R. Kornblihtt, J. F. Caceres, O. A. Coso and A. Srebrow (2005). "Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT." Nat Struct Mol Biol **12**(12): 1037-1044.

Boehringer, D., E. M. Makarov, B. Sander, O. V. Makarova, B. Kastner, R. Luhrmann and H. Stark (2004). "Three-dimensional structure of a pre-catalytic human spliceosomal complex B." Nat Struct Mol Biol **11**(5): 463-468.

Boucher, L., C. A. Ouzounis, A. J. Enright and B. J. Blencowe (2001). "A genome-wide survey of RS domain proteins." RNA **7**(12): 1693-1701.

Boukis, L. A. and J. P. Bruzik (2001). "Functional selection of splicing enhancers that stimulate trans-splicing in vitro." RNA **7**(6): 793-805.

Brun, C., D. Suter, C. Pauli, P. Dunant, H. Lochmuller, J. M. Burgunder, D. Schumperli and J. Weis (2003). "U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping." Cell Mol Life Sci **60**(3): 557-566.

Buratti, E. and F. E. Baralle (2004). "Influence of RNA secondary structure on the pre-mRNA splicing process." Molecular and cellular biology **24**(24): 10505-10514.

Buratti, E., M. Baralle and F. E. Baralle (2006). "Defective splicing, disease and therapy: searching for master checkpoints in exon definition." Nucleic acids research **34**(12): 3494-3510.

Burd, C. G. and G. Dreyfuss (1994). "RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing." EMBO J **13**(5): 1197-1204.

Caceres, J. F., G. R. Screaton and A. R. Krainer (1998). "A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm." Genes Dev **12**(1): 55-66.

Caceres, J. F., S. Stamm, D. M. Helfman and A. R. Krainer (1994). "Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors." Science **265**(5179): 1706-1709.

Caputi, M. and A. M. Zahler (2002). "SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D." EMBO J **21**(4): 845-855.

Cartegni, L., S. L. Chew and A. R. Krainer (2002). "Listening to silence and understanding nonsense: exonic mutations that affect splicing." Nat Rev Genet **3**(4): 285-298.

Cartegni, L. and A. R. Krainer (2002). "Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1." Nat Genet **30**(4): 377-384.

Cartegni, L., M. Maconi, E. Morandi, F. Cobianchi, S. Riva and G. Biamonti (1996). "hnRNP A1 selectively interacts through its Gly-rich domain with different RNA-binding proteins." J Mol Biol **259**(3): 337-348.

Cartegni, L., J. Wang, Z. Zhu, M. Q. Zhang and A. R. Krainer (2003). "ESEfinder: A web resource to identify exonic splicing enhancers." Nucleic Acids Res **31**(13): 3568-3571.

Cavallari, N., D. Balestra, A. Branchini, I. Maestri, A. Chuamsunrit, W. Sasanakul, G. Mariani, F. Pagani, F. Bernardi and M. Pinotti (2012). "Activation of a cryptic splice site in a potentially lethal coagulation defect accounts for a functional protein variant." Biochim Biophys Acta **1822**(7): 1109-1113.

Celotto, A. M. and B. R. Graveley (2001). "Alternative splicing of the Drosophila Dscam pre-mRNA is both temporally and spatially regulated." Genetics **159**(2): 599-608.

Chabot, B., M. Blanchette, I. Lapierre and H. La Branche (1997). "An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1." Mol Cell Biol **17**(4): 1776-1786.

Cho, S. and G. Dreyfuss (2010). "A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity." Genes Dev **24**(5): 438-442.

Coady, T. H., T. D. Baughan, M. Shababi, M. A. Passini and C. L. Lorson (2008). "Development of a single vector system that enhances trans-splicing of SMN2 transcripts." PLoS One **3**(10): e3468.

Coady, T. H. and C. L. Lorson (2010). "Trans-splicing-mediated improvement in a severe mouse model of spinal muscular atrophy." J Neurosci **30**(1): 126-130.

Cooper, T. A. (2005). "Use of minigene systems to dissect alternative splicing elements." Methods **37**(4): 331-340.

Cooper, T. A., L. Wan and G. Dreyfuss (2009). "RNA and disease." Cell **136**(4): 777-793.

Corti, S., M. Nizzardo, M. Nardini, C. Donadoni, S. Salani, D. Ronchi, F. Saladino, A. Bordoni, F. Fortunato, R. Del Bo, D. Papadimitriou, F. Locatelli, G. Menozzi, S. Strazzer, N. Bresolin and G. P. Comi (2008). "Neural stem cell transplantation can ameliorate the phenotype of a mouse model of spinal muscular atrophy." The Journal of clinical investigation **118**(10): 3316-3330.

Corti, S., M. Nizzardo, M. Nardini, C. Donadoni, S. Salani, D. Ronchi, C. Simone, M. Falcone, D. Papadimitriou, F. Locatelli, N. Mezzina, F. Gianni, N. Bresolin and G. P. Comi (2010). "Embryonic stem cell-derived neural stem cells improve spinal muscular atrophy phenotype in mice." Brain : a journal of neurology **133**(Pt 2): 465-481.

Covaciu, C., F. Grosso, E. Pisaneschi, G. Zambruno, P. A. Gregersen, M. Sommerlund, J. M. Hertz and D. Castiglia (2011). "A founder synonymous COL7A1 mutation in three Danish families with dominant dystrophic epidermolysis bullosa pruriginosa identifies exonic regulatory sequences required for exon 87 splicing." Br J Dermatol **165**(3): 678-682.

D'Amico, A., E. Mercuri, F. D. Tiziano and E. Bertini (2011). "Spinal muscular atrophy." Orphanet J Rare Dis **6**: 71.

Damgaard, C. K., S. Kahns, S. Lykke-Andersen, A. L. Nielsen, T. H. Jensen and J. Kjems (2008). "A 5' splice site enhances the recruitment of basal transcription initiation factors in vivo." Mol Cell **29**(2): 271-278.

Das, R. and R. Reed (1999). "Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels." RNA **5**(11): 1504-1508.

De Angelis, F. G., O. Sthandier, B. Berarducci, S. Toso, G. Galluzzi, E. Ricci, G. Cossu and I. Bozzoni (2002). "Chimeric snRNA molecules carrying antisense sequences against the splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in Delta 48-50 DMD cells." Proc Natl Acad Sci U S A **99**(14): 9456-9461.

Denti, M. A., A. Rosa, G. D'Antona, O. Sthandier, F. G. De Angelis, C. Nicoletti, M. Allocca, O. Pansarasa, V. Parente, A. Musaro, A. Auricchio, R. Bottinelli and I. Bozzoni (2006). "Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model." Proc Natl Acad Sci U S A **103**(10): 3758-3763.

Desmet, F. O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres and C. Beroud (2009). "Human Splicing Finder: an online bioinformatics tool to predict splicing signals." Nucleic Acids Res **37**(9): e67.

Dhir, A. and E. Buratti (2010). "Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies." FEBS J **277**(4): 841-855.

Di, W. L., F. Larcher, E. Semenova, G. E. Talbot, J. I. Harper, M. Del Rio, A. J. Thrasher and W. Qasim (2011). "Ex-vivo gene therapy restores LEKTI activity and corrects the architecture of Netherton syndrome-derived skin grafts." Mol Ther **19**(2): 408-416.

Di, W. L., J. E. Mellerio, C. Bernadis, J. Harper, A. Abdul-Wahab, S. Ghani, L. Chan, M. Martinez-Queipo, H. Hara, A. M. McNicol, F. Farzaneh, J. McGrath, A. Thrasher and W. Qasim (2013). "Phase I study protocol for ex vivo lentiviral

gene therapy for the inherited skin disease, Netherton syndrome." Hum Gene Ther Clin Dev **24**(4): 182-190.

Diociaiuti, A., D. Castiglia, P. Fortugno, A. Bartuli, M. Pascucci, G. Zambruno and M. El Hachem (2013). "Lethal Netherton syndrome due to homozygous p.Arg371X mutation in SPINK5." Pediatr Dermatol **30**(4): e65-67.

Disset, A., C. F. Bourgeois, N. Benmalek, M. Claustres, J. Stevenin and S. Tuffery-Giraud (2006). "An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements." Hum Mol Genet **15**(6): 999-1013.

Dominguez, C. and F. H. Allain (2006). "NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition." Nucleic Acids Res **34**(13): 3634-3645.

Dominski, Z. and R. Kole (1993). "Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides." Proc Natl Acad Sci U S A **90**(18): 8673-8677.

Dreumont, N., S. Hardy, I. Behm-Ansmant, L. Kister, C. Branlant, J. Stevenin and C. F. Bourgeois (2010). "Antagonistic factors control the unproductive splicing of SC35 terminal intron." Nucleic Acids Res **38**(4): 1353-1366.

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

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

Du, H. and M. Rosbash (2002). "The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing." Nature **419**(6902): 86-90.

Du, L., J. M. Pollard and R. A. Gatti (2007). "Correction of prototypic ATM splicing mutations and aberrant ATM function with antisense morpholino oligonucleotides." Proc Natl Acad Sci U S A **104**(14): 6007-6012.

Egloff, S., D. O'Reilly, R. D. Chapman, A. Taylor, K. Tanzhaus, L. Pitts, D. Eick and S. Murphy (2007). "Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression." Science **318**(5857): 1777-1779.

Egloff, S., D. O'Reilly and S. Murphy (2008). "Expression of human snRNA genes from beginning to end." Biochemical Society transactions **36**(Pt 4): 590-594.

Emery, A. E. (2002). "The muscular dystrophies." Lancet **359**(9307): 687-695.

Fairbrother, W. G. and L. A. Chasin (2000). "Human genomic sequences that inhibit splicing." Mol Cell Biol **20**(18): 6816-6825.

Fairbrother, W. G., G. W. Yeo, R. Yeh, P. Goldstein, M. Mawson, P. A. Sharp and C. B. Burge (2004). "RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons." Nucleic Acids Res **32**(Web Server issue): W187-190.

Faustino, N. A. and T. A. Cooper (2003). "Pre-mRNA splicing and human disease." Genes Dev **17**(4): 419-437.

Fernandez Alanis, E., M. Pinotti, A. Dal Mas, D. Balestra, N. Cavallari, M. E. Rogalska, F. Bernardi and F. Pagani (2012). "An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects." Human molecular genetics **21**(11): 2389-2398.

Fortugno, P., A. Bresciani, C. Paolini, C. Pazzagli, M. El Hachem, M. D'Alessio and G. Zambruno (2011). "Proteolytic activation cascade of the Netherton syndrome-defective protein, LEKTI, in the epidermis: implications for skin homeostasis." J Invest Dermatol **131**(11): 2223-2232.

Fortugno, P., F. Grosso, G. Zambruno, S. Pastore, F. Faletra and D. Castiglia (2012). "A synonymous mutation in SPINK5 exon 11 causes Netherton syndrome by altering exonic splicing regulatory elements." J Hum Genet **57**(5): 311-315.

Garcia-Blanco, M. A., A. P. Baraniak and E. L. Lasda (2004). "Alternative splicing in disease and therapy." Nat Biotechnol **22**(5): 535-546.

Gogliotti, R. G., S. M. Hammond, C. Lutz and C. J. Didonato (2010). "Molecular and phenotypic reassessment of an infrequently used mouse model for spinal muscular atrophy." Biochemical and biophysical research communications **391**(1): 517-522.

Gogliotti, R. G., K. A. Quinlan, C. B. Barlow, C. R. Heier, C. J. Heckman and C. J. Didonato (2012). "Motor neuron rescue in spinal muscular atrophy mice demonstrates that sensory-motor defects are a consequence, not a cause, of motor neuron dysfunction." The Journal of neuroscience : the official journal of the Society for Neuroscience **32**(11): 3818-3829.

Goyenvalle, A. (2012). "Engineering U7snRNA gene to reframe transcripts." Methods in molecular biology **867**: 259-271.

Gozani, O., R. Feld and R. Reed (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**(2): 233-243.

Graveley, B. R. (2000). "Sorting out the complexity of SR protein functions." RNA **6**(9): 1197-1211.

Gross, J. D., N. J. Moerke, T. von der Haar, A. A. Lugovskoy, A. B. Sachs, J. E. McCarthy and G. Wagner (2003). "Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E." Cell **115**(6): 739-750.

Grover, A., H. Houlden, M. Baker, J. Adamson, J. Lewis, G. Prihar, S. Pickering-Brown, K. Duff and M. Hutton (1999). "5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10." J Biol Chem **274**(21): 15134-15143.

Guil, S. and J. F. Caceres (2007). "The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a." Nat Struct Mol Biol **14**(7): 591-596.

Gunderson, S. I., M. Polycarpou-Schwarz and I. W. Mattaj (1998). "U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase." Molecular cell **1**(2): 255-264.

Hamilton, B. J., C. M. Burns, R. C. Nichols and W. F. Rigby (1997). "Modulation of AUUUA response element binding by heterogeneous nuclear ribonucleoprotein A1 in human T lymphocytes. The roles of cytoplasmic location, transcription, and phosphorylation." J Biol Chem **272**(45): 28732-28741.

Hamm, J., N. A. Dathan, D. Scherly and I. W. Mattaj (1990). "Multiple domains of U1 snRNA, including U1 specific protein binding sites, are required for splicing." EMBO J **9**(4): 1237-1244.

Hammond, S. M. and M. J. Wood (2011). "Genetic therapies for RNA mis-splicing diseases." Trends in genetics : TIG **27**(5): 196-205.

Heinrichs, V., M. Bach, G. Winkelmann and R. Luhrmann (1990). "U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site." Science **247**(4938): 69-72.

Hernandez, H., O. V. Makarova, E. M. Makarov, N. Morgner, Y. Muto, D. P. Krummel and C. V. Robinson (2009). "Isoforms of U1-70k control subunit dynamics in the human spliceosomal U1 snRNP." PloS one **4**(9): e7202.

Hernandez, N. (2001). "Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription." J Biol Chem **276**(29): 26733-26736.

Hiller, M., Z. Zhang, R. Backofen and S. Stamm (2007). "Pre-mRNA secondary structures influence exon recognition." PLoS Genet **3**(11): e204.

Hong, W., M. Bennett, Y. Xiao, R. Feld Kramer, C. Wang and R. Reed (1997). "Association of U2 snRNP with the spliceosomal complex E." Nucleic Acids Res **25**(2): 354-361.

Honore, B., H. H. Rasmussen, H. Vorum, K. Dejgaard, X. Liu, P. Gromov, P. Madsen, B. Gesser, N. Tommerup and J. E. Celis (1995). "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromosomes." J Biol Chem **270**(48): 28780-28789.

Horowitz, D. S. and A. R. Krainer (1994). "Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing." Trends Genet **10**(3): 100-106.

Hua, Y. and A. R. Krainer (2012). "Antisense-mediated exon inclusion." Methods in molecular biology **867**: 307-323.

Hua, Y., K. Sahashi, G. Hung, F. Rigo, M. A. Passini, C. F. Bennett and A. R. Krainer (2010). "Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model." Genes & development **24**(15): 1634-1644.

Hua, Y., K. Sahashi, F. Rigo, G. Hung, G. Horev, C. F. Bennett and A. R. Krainer (2011). "Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model." Nature **478**(7367): 123-126.

Hua, Y., T. A. Vickers, B. F. Baker, C. F. Bennett and A. R. Krainer (2007). "Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon." PLoS biology **5**(4): e73.

Hua, Y., T. A. Vickers, H. L. Okunola, C. F. Bennett and A. R. Krainer (2008). "Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice." American journal of human genetics **82**(4): 834-848.

Huang, Y. and J. A. Steitz (2005). "SRprises along a messenger's journey." Mol Cell **17**(5): 613-615.

Iervolino, A., G. Santilli, R. Trotta, C. Guerzoni, V. Cesi, A. Bergamaschi, C. Gambacorti-Passerini, B. Calabretta and D. Perrotti (2002). "hnRNP A1 nucleocytoplasmic shuttling activity is required for normal myelopoiesis and BCR/ABL leukemogenesis." Mol Cell Biol **22**(7): 2255-2266.

Incitti, T., F. G. De Angelis, V. Cazzella, O. Sthandier, C. Pinnaro, I. Legnini and I. Bozzoni (2010). "Exon skipping and duchenne muscular dystrophy therapy: selection of the most active U1 snRNA antisense able to induce dystrophin exon 51 skipping." Mol Ther **18**(9): 1675-1682.

Izaurralde, E. and I. W. Mattaj (1995). "RNA export." Cell **81**(2): 153-159.

Kaida, D., M. G. Berg, I. Younis, M. Kasim, L. N. Singh, L. Wan and G. Dreyfuss (2010). "U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation." Nature **468**(7324): 664-668.

Kamma, H., D. S. Portman and G. Dreyfuss (1995). "Cell type-specific expression of hnRNP proteins." Exp Cell Res **221**(1): 187-196.

Kan, J. L. and M. R. Green (1999). "Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor." Genes Dev **13**(4): 462-471.

Kashima, T., N. Rao, C. J. David and J. L. Manley (2007). "hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing." Hum Mol Genet **16**(24): 3149-3159.

Kashima, T., N. Rao and J. L. Manley (2007). "An intronic element contributes to splicing repression in spinal muscular atrophy." Proc Natl Acad Sci U S A **104**(9): 3426-3431.

Kent, O. A., D. B. Ritchie and A. M. Macmillan (2005). "Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway." Mol Cell Biol **25**(1): 233-240.

Khoo, B., X. Roca, S. L. Chew and A. R. Krainer (2007). "Antisense oligonucleotide-induced alternative splicing of the APOB mRNA generates a novel isoform of APOB." BMC Mol Biol **8**: 3.

Kielkopf, C. L., S. Lucke and M. R. Green (2004). "U2AF homology motifs: protein recognition in the RRM world." Genes Dev **18**(13): 1513-1526.

Kim, S., B. M. Merrill, R. Rajpurohit, A. Kumar, K. L. Stone, V. V. Papov, J. M. Schneiders, W. Szer, S. H. Wilson, W. K. Paik and K. R. Williams (1997). "Identification of N(G)-methylarginine residues in human heterogeneous RNP protein A1: Phe/Gly-Gly-Gly-Arg-Gly-Gly-Gly/Phe is a preferred recognition motif." Biochemistry **36**(17): 5185-5192.

Kiss, T. (2004). "Biogenesis of small nuclear RNPs." J Cell Sci **117**(Pt 25): 5949-5951.

Kolb, S. J. and J. T. Kissel (2011). "Spinal muscular atrophy: a timely review." Archives of neurology **68**(8): 979-984.

Konarska, M. M., P. J. Grabowski, R. A. Padgett and P. A. Sharp (1985). "Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors." Nature **313**(6003): 552-557.

Krawczak, M., N. S. Thomas, B. Hundrieser, M. Mort, M. Wittig, J. Hampe and D. N. Cooper (2007). "Single base-pair substitutions in exon-intron junctions

of human genes: nature, distribution, and consequences for mRNA splicing." Hum Mutat **28**(2): 150-158.

Kuhlman, T. C., H. Cho, D. Reinberg and N. Hernandez (1999). "The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter." Mol Cell Biol **19**(3): 2130-2141.

Lacroix, M., L. Lacaze-Buzy, L. Furio, E. Tron, M. Valari, G. Van der Wier, C. Bodemer, A. Bygum, A. C. Bursztejn, G. Gaitanis, M. Paradisi, A. Stratigos, L. Weibel, C. Deraison and A. Hovnanian (2012). "Clinical expression and new SPINK5 splicing defects in Netherton syndrome: unmasking a frequent founder synonymous mutation and unconventional intronic mutations." J Invest Dermatol **132**(3 Pt 1): 575-582.

Lamond, A. I. (1993). "A glimpse into the spliceosome." Curr Biol **3**(1): 62-64.

Langford, C. J., F. J. Klinz, C. Donath and D. Gallwitz (1984). "Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast." Cell **36**(3): 645-653.

Le, T. T., L. T. Pham, M. E. Butchbach, H. L. Zhang, U. R. Monani, D. D. Coovert, T. O. Gavrilina, L. Xing, G. J. Bassell and A. H. Burghes (2005). "SMNDelta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN." Hum Mol Genet **14**(6): 845-857.

Lefebvre, S., L. Burglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau, M. Zeviani and et al. (1995). "Identification and characterization of a spinal muscular atrophy-determining gene." Cell **80**(1): 155-165.

Lew, J. M., Y. L. Fei, K. Aleck, B. J. Blencowe, R. Weksberg and P. D. Sadowski (2004). "CDKN1C mutation in Wiedemann-Beckwith syndrome patients reduces RNA splicing efficiency and identifies a splicing enhancer." Am J Med Genet A **127A**(3): 268-276.

Lewis, J. D., E. Izaurralde, A. Jarmolowski, C. McGuigan and I. W. Mattaj (1996). "A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site." Genes Dev **10**(13): 1683-1698.

Liu, H. X., M. Zhang and A. R. Krainer (1998). "Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins." Genes Dev **12**(13): 1998-2012.

Longman, D., I. L. Johnstone and J. F. Caceres (2000). "Functional characterization of SR and SR-related genes in Caenorhabditis elegans." EMBO J **19**(7): 1625-1637.

Lorson, C. L., H. Rindt and M. Shababi (2010). "Spinal muscular atrophy: mechanisms and therapeutic strategies." Human molecular genetics **19**(R1): R111-118.

Lotti, F., W. L. Imlach, L. Saieva, E. S. Beck, T. Hao le, D. K. Li, W. Jiao, G. Z. Mentis, C. E. Beattie, B. D. McCabe and L. Pellizzoni (2012). "An SMN-dependent U12 splicing event essential for motor circuit function." Cell **151**(2): 440-454.

Luco, R. F., Q. Pan, K. Tominaga, B. J. Blencowe, O. M. Pereira-Smith and T. Misteli (2010). "Regulation of alternative splicing by histone modifications." Science **327**(5968): 996-1000.

Lund, E. and J. E. Dahlberg (1984). "True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation." J Biol Chem **259**(3): 2013-2021.

Madocsai, C., S. R. Lim, T. Geib, B. J. Lam and K. J. Hertel (2005). "Correction of SMN2 Pre-mRNA splicing by antisense U7 small nuclear RNAs." Molecular therapy : the journal of the American Society of Gene Therapy **12**(6): 1013-1022.

Maniatis, T. and R. Reed (2002). "An extensive network of coupling among gene expression machines." Nature **416**(6880): 499-506.

Maniatis, T. and B. Tasic (2002). "Alternative pre-mRNA splicing and proteome expansion in metazoans." Nature **418**(6894): 236-243.

Mardon, H. J., G. Sebastio and F. E. Baralle (1987). "A role for exon sequences in alternative splicing of the human fibronectin gene." Nucleic Acids Res **15**(19): 7725-7733.

Marquis, J., K. Meyer, L. Angehrn, S. S. Kampfer, B. Rothen-Rutishauser and D. Schumperli (2007). "Spinal muscular atrophy: SMN2 pre-mRNA splicing corrected by a U7 snRNA derivative carrying a splicing enhancer sequence." Molecular therapy : the journal of the American Society of Gene Therapy **15**(8): 1479-1486.

Martinez-Contreras, R., P. Cloutier, L. Shkreta, J. F. Fisette, T. Revil and B. Chabot (2007). "hnRNP proteins and splicing control." Adv Exp Med Biol **623**: 123-147.

Martone, J., F. G. De Angelis and I. Bozzoni (2012). "U1 snRNA as an effective vector for stable expression of antisense molecules and for the inhibition of the splicing reaction." Methods in molecular biology **867**: 239-257.

Matlin, A. J., F. Clark and C. W. Smith (2005). "Understanding alternative splicing: towards a cellular code." Nat Rev Mol Cell Biol **6**(5): 386-398.

Mavilio, F., G. Pellegrini, S. Ferrari, F. Di Nunzio, E. Di Iorio, A. Recchia, G. Maruggi, G. Ferrari, E. Provasi, C. Bonini, S. Capurro, A. Conti, C. Magnoni, A. Giannetti and M. De Luca (2006). "Correction of junctional epidermolysis bullosa by transplantation of genetically modified epidermal stem cells." Nat Med **12**(12): 1397-1402.

Mayeda, A. and A. R. Krainer (1992). "Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2." Cell **68**(2): 365-375.

McCarthy, E. M. and J. A. Phillips, 3rd (1998). "Characterization of an intron splice enhancer that regulates alternative splicing of human GH pre-mRNA." Hum Mol Genet **7**(9): 1491-1496.

Medlin, J. E., P. Uguen, A. Taylor, D. L. Bentley and S. Murphy (2003). "The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA." EMBO J **22**(4): 925-934.

Merkhofer, E. C. and T. L. Johnson (2012). "U1 snRNA rewrites the "script"." Cell **150**(1): 9-11.

Meyer, K., J. Marquis, J. Trub, R. Nlend Nlend, S. Verp, M. D. Ruepp, H. Imboden, I. Barde, D. Trono and D. Schumperli (2009). "Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation." Human molecular genetics **18**(3): 546-555.

Misteli, T. and D. L. Spector (1997). "Protein phosphorylation and the nuclear organization of pre-mRNA splicing." Trends Cell Biol **7**(4): 135-138.

Moore, M. J. and P. A. Sharp (1993). "Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing." Nature **365**(6444): 364-368.

Mount, S. M., I. Pettersson, M. Hinterberger, A. Karmas and J. A. Steitz (1983). "The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro." Cell **33**(2): 509-518.

Muro, A. F., M. Caputi, R. Pariyarath, F. Pagani, E. Buratti and F. E. Baralle (1999). "Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display." Mol Cell Biol **19**(4): 2657-2671.

Muro, A. F., A. Iaconcig and F. E. Baralle (1998). "Regulation of the fibronectin EDA exon alternative splicing. Cooperative role of the exonic enhancer element and the 5' splicing site." FEBS Lett **437**(1-2): 137-141.

Nakahata, S. and S. Kawamoto (2005). "Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities." Nucleic Acids Res **33**(7): 2078-2089.

Nelissen, R. L., C. L. Will, W. J. van Venrooij and R. Luhrmann (1994). "The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins." EMBO J **13**(17): 4113-4125.

O'Keefe, R. T., C. Norman and A. J. Newman (1996). "The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast." Cell **86**(4): 679-689.

Okunola, H. L. and A. R. Krainer (2009). "Cooperative-binding and splicing-repressive properties of hnRNP A1." Molecular and cellular biology **29**(20): 5620-5631.

Pagani, F. and F. E. Baralle (2004). "Genomic variants in exons and introns: identifying the splicing spoilers." Nature reviews. Genetics **5**(5): 389-396.

Pagani, F., E. Buratti, C. Stuani and F. E. Baralle (2003). "Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9." J Biol Chem **278**(29): 26580-26588.

Pagani, F., E. Buratti, C. Stuani, R. Bendix, T. Dork and F. E. Baralle (2002). "A new type of mutation causes a splicing defect in ATM." Nature genetics **30**(4): 426-429.

Pagani, F., E. Buratti, C. Stuani, M. Romano, E. Zuccato, M. Niksic, L. Giglio, D. Faraguna and F. E. Baralle (2000). "Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element." J Biol Chem **275**(28): 21041-21047.

Pagani, F., C. Stuani, M. Tzetis, E. Kanavakis, A. Efthymiadou, S. Doudounakis, T. Casals and F. E. Baralle (2003). "New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12." Hum Mol Genet **12**(10): 1111-1120.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey and B. J. Blencowe (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nat Genet **40**(12): 1413-1415.

Paradis, C., P. Cloutier, L. Shkreta, J. Toutant, K. Klarskov and B. Chabot (2007). "hnRNP I/PTB can antagonize the splicing repressor activity of SRp30c." RNA **13**(8): 1287-1300.

Passini, M. A., J. Bu, E. M. Roskelley, A. M. Richards, S. P. Sardi, C. R. O'Riordan, K. W. Klinger, L. S. Shihabuddin and S. H. Cheng (2010). "CNS-targeted gene therapy improves survival and motor function in a mouse model of spinal muscular atrophy." The Journal of clinical investigation **120**(4): 1253-1264.

Patel, S. B. and M. Bellini (2008). "The assembly of a spliceosomal small nuclear ribonucleoprotein particle." Nucleic acids research **36**(20): 6482-6493.

Perez, B., A. Rincon, A. Jorge-Finnigan, E. Richard, B. Merinero, M. Ugarte and L. R. Desviat (2009). "Pseudoexon exclusion by antisense therapy in methylmalonic aciduria (MMAuria)." Hum Mutat **30**(12): 1676-1682.

Pinol-Roma, S. and G. Dreyfuss (1992). "Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm." Nature **355**(6362): 730-732.

Pinotti, M., L. Rizzotto, D. Balestra, M. A. Lewandowska, N. Cavallari, G. Marchetti, F. Bernardi and F. Pagani (2008). "U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency." Blood **111**(5): 2681-2684.

Pohlenz, J., A. Dumitrescu, U. Aumann, G. Koch, R. Melchior, D. Prawitt and S. Refetoff (2002). "Congenital secondary hypothyroidism caused by exon skipping due to a homozygous donor splice site mutation in the TSHbeta-subunit gene." J Clin Endocrinol Metab **87**(1): 336-339.

Pozzoli, U. and M. Sironi (2005). "Silencers regulate both constitutive and alternative splicing events in mammals." Cell Mol Life Sci **62**(14): 1579-1604.

Prokocimer, M., R. Barkan and Y. Gruenbaum (2013). "Hutchinson-Gilford progeria syndrome through the lens of transcription." Aging Cell **12**(4): 533-543.

Puttaraju, M., S. F. Jamison, S. G. Mansfield, M. A. Garcia-Blanco and L. G. Mitchell (1999). "Spliceosome-mediated RNA trans-splicing as a tool for gene therapy." Nat Biotechnol **17**(3): 246-252.

Qin, S. and H. X. Zhou (2008). "Prediction of salt and mutational effects on the association rate of U1A protein and U1 small nuclear RNA stem/loop II." The journal of physical chemistry. B **112**(19): 5955-5960.

Raker, V. A., K. Hartmuth, B. Kastner and R. Luhrmann (1999). "Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner." Mol Cell Biol **19**(10): 6554-6565.

Reed, R. (1989). "The organization of 3' splice-site sequences in mammalian introns." Genes Dev **3**(12B): 2113-2123.

Reed, R. (2000). "Mechanisms of fidelity in pre-mRNA splicing." Curr Opin Cell Biol **12**(3): 340-345.

Reed, R. and T. Maniatis (1988). "The role of the mammalian branchpoint sequence in pre-mRNA splicing." Genes Dev **2**(10): 1268-1276.

Rigo, F., Y. Hua, A. R. Krainer and C. F. Bennett (2012). "Antisense-based therapy for the treatment of spinal muscular atrophy." The Journal of cell biology **199**(1): 21-25.

Rinke, J., B. Appel, H. Blocker, R. Frank and R. Luhrmann (1984). "The 5'-terminal sequence of U1 RNA complementary to the consensus 5' splice site of hnRNA is single-stranded in intact U1 snRNP particles." Nucleic Acids Res **12**(10): 4111-4126.

Roca, X., M. Akerman, H. Gaus, A. Berdeja, C. F. Bennett and A. R. Krainer (2012). "Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides." Genes & development **26**(10): 1098-1109.

Roca, X. and A. R. Krainer (2009). "Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA." Nature structural & molecular biology **16**(2): 176-182.

Roca, X., A. R. Krainer and I. C. Eperon (2013). "Pick one, but be quick: 5' splice sites and the problems of too many choices." Genes & development **27**(2): 129-144.

Roscigno, R. F., M. Weiner and M. A. Garcia-Blanco (1993). "A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing." J Biol Chem **268**(15): 11222-11229.

Rosel, T. D., L. H. Hung, J. Medenbach, K. Donde, S. Starke, V. Benes, G. Ratsch and A. Bindereif (2011). "RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation." The EMBO journal **30**(10): 1965-1976.

Ruskin, B., A. R. Krainer, T. Maniatis and M. R. Green (1984). "Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro." Cell **38**(1): 317-331.

Sadowski, C. L., R. W. Henry, S. M. Lobo and N. Hernandez (1993). "Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE." Genes Dev **7**(8): 1535-1548.

Sanchez-Alcudia, R., B. Perez, C. Perez-Cerda, M. Ugarte and L. R. Desviat (2011). "Overexpression of adapted U1snRNA in patients' cells to correct a 5' splice site mutation in propionic acidemia." Molecular genetics and metabolism **102**(2): 134-138.

Sanford, J. R., J. Ellis and J. F. Caceres (2005). "Multiple roles of arginine/serine-rich splicing factors in RNA processing." Biochem Soc Trans **33**(Pt 3): 443-446.

Sauliere, J., A. Sureau, A. Expert-Bezancon and J. Marie (2006). "The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit." Mol Cell Biol **26**(23): 8755-8769.

Schaal, T. D. and T. Maniatis (1999). "Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA." Mol Cell Biol **19**(1): 261-273.

Schaal, T. D. and T. Maniatis (1999). "Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences." Mol Cell Biol **19**(3): 1705-1719.

Schellenberg, M. J., D. B. Ritchie and A. M. MacMillan (2008). "Pre-mRNA splicing: a complex picture in higher definition." Trends Biochem Sci **33**(6): 243-246.

Schmid, F., E. Glaus, D. Barthelmes, M. Fliegauf, H. Gaspar, G. Nurnberg, P. Nurnberg, H. Omran, W. Berger and J. Neidhardt (2011). "U1 snRNA-mediated gene therapeutic correction of splice defects caused by an exceptionally mild BBS mutation." Human mutation **32**(7): 815-824.

Schmid, F., T. Hiller, G. Korner, E. Glaus, W. Berger and J. Neidhardt (2012). "A Gene Therapeutic Approach to Correct Splice Defects with Modified U1 and U6 snRNPs." Human gene therapy.

Schmid, M. and T. H. Jensen (2008). "Quality control of mRNP in the nucleus." Chromosoma **117**(5): 419-429.

Schmucker, D., J. C. Clemens, H. Shu, C. A. Worby, J. Xiao, M. Muda, J. E. Dixon and S. L. Zipursky (2000). "Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity." Cell **101**(6): 671-684.

Senapathy, P., M. B. Shapiro and N. L. Harris (1990). "Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project." Methods Enzymol **183**: 252-278.

Shapiro, M. B. and P. Senapathy (1987). "RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression." Nucleic Acids Res **15**(17): 7155-7174.

Shen, H. and M. R. Green (2004). "A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly." Mol Cell **16**(3): 363-373.

Skordis, L. A., M. G. Dunckley, B. Yue, I. C. Eperon and F. Muntoni (2003). "Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts." Proc Natl Acad Sci U S A **100**(7): 4114-4119.

Smith, C. W. and B. Nadal-Ginard (1989). "Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing." Cell **56**(5): 749-758.

Sommer, B., M. Kohler, R. Sprengel and P. H. Seeburg (1991). "RNA editing in brain controls a determinant of ion flow in glutamate-gated channels." Cell **67**(1): 11-19.

Sperling, J., M. Azubel and R. Sperling (2008). "Structure and function of the Pre-mRNA splicing machine." Structure **16**(11): 1605-1615.

Spurdle, A. B., F. J. Couch, F. B. Hogervorst, P. Radice, O. M. Sinilnikova and I. U. G. V. W. Group (2008). "Prediction and assessment of splicing alterations: implications for clinical testing." Hum Mutat **29**(11): 1304-1313.

Stamm, S. (2008). "Regulation of alternative splicing by reversible protein phosphorylation." J Biol Chem **283**(3): 1223-1227.

Stark, H., P. Dube, R. Luhrmann and B. Kastner (2001). "Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle." Nature **409**(6819): 539-542.

Sun, H. and L. A. Chasin (2000). "Multiple splicing defects in an intronic false exon." Mol Cell Biol **20**(17): 6414-6425.

Surowy, C. S., V. L. van Santen, S. M. Scheib-Wixted and R. A. Spritz (1989). "Direct, sequence-specific binding of the human U1-70K ribonucleoprotein antigen protein to loop I of U1 small nuclear RNA." Mol Cell Biol **9**(10): 4179-4186.

Svasti, S., T. Suwanmanee, S. Fucharoen, H. M. Moulton, M. H. Nelson, N. Maeda, O. Smithies and R. Kole (2009). "RNA repair restores hemoglobin expression in IVS2-654 thalassemic mice." Proc Natl Acad Sci U S A **106**(4): 1205-1210.

Szeszel-Fedorowicz, W., I. Talukdar, B. N. Griffith, C. M. Walsh and L. M. Salati (2006). "An exonic splicing silencer is involved in the regulated splicing of glucose 6-phosphate dehydrogenase mRNA." J Biol Chem **281**(45): 34146-34158.

Tardiff, D. F. and M. Rosbash (2006). "Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly." RNA **12**(6): 968-979.

Teraoka, S. N., M. Telatar, S. Becker-Catania, T. Liang, S. Onengut, A. Tolun, L. Chessa, O. Sanal, E. Bernatowska, R. A. Gatti and P. Concannon (1999). "Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences." Am J Hum Genet **64**(6): 1617-1631.

Tian, H. and R. Kole (1995). "Selection of novel exon recognition elements from a pool of random sequences." Mol Cell Biol **15**(11): 6291-6298.

Tournier, I., M. Vezain, A. Martins, F. Charbonnier, S. Baert-Desurmont, S. Olschwang, Q. Wang, M. P. Buisine, J. Soret, J. Tazi, T. Frebourg and M. Tosi (2008). "A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects." Hum Mutat **29**(12): 1412-1424.

Trono, D. (2001). "Lentiviral vectors for the genetic modification of hematopoietic stem cells." Ernst Schering Res Found Workshop(33): 19-28.

Turner, I. A., C. M. Norman, M. J. Churcher and A. J. Newman (2004). "Roles of the U5 snRNP in spliceosome dynamics and catalysis." Biochem Soc Trans **32**(Pt 6): 928-931.

Ubby, I., E. Bussani, A. Colonna, G. Stacul, M. Locatelli, P. Scudieri, L. Galietta and F. Pagani (2013). "TMEM16A alternative splicing coordination in breast cancer." Mol Cancer **12**: 75.

Uguen, P. and S. Murphy (2003). "The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing." EMBO J **22**(17): 4544-4554.

Uguen, P. and S. Murphy (2004). "3'-box-dependent processing of human pre-U1 snRNA requires a combination of RNA and protein co-factors." Nucleic acids research **32**(10): 2987-2994.

Umen, J. G. and C. Guthrie (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**(6): 584-597.

Valadkhan, S. and J. L. Manley (2001). "Splicing-related catalysis by protein-free snRNAs." Nature **413**(6857): 701-707.

Valadkhan, S., A. Mohammadi, Y. Jaladat and S. Geisler (2009). "Protein-free small nuclear RNAs catalyze a two-step splicing reaction." Proc Natl Acad Sci U S A **106**(29): 11901-11906.

Valcarcel, J., R. K. Gaur, R. Singh and M. R. Green (1996). "Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]." Science **273**(5282): 1706-1709.

Venables, J. P. (2007). "Downstream intronic splicing enhancers." FEBS Lett **581**(22): 4127-4131.

Vidal, C., A. Cachia and A. Xuereb-Anastasi (2009). "Effects of a synonymous variant in exon 9 of the CD44 gene on pre-mRNA splicing in a family with osteoporosis." Bone **45**(4): 736-742.

Wahl, M. C., C. L. Will and R. Luhrmann (2009). "The spliceosome: design principles of a dynamic RNP machine." Cell **136**(4): 701-718.

Wang, E. T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth and C. B. Burge (2008). "Alternative isoform regulation in human tissue transcriptomes." Nature **456**(7221): 470-476.

Wang, Z., M. E. Rolish, G. Yeo, V. Tung, M. Mawson and C. B. Burge (2004). "Systematic identification and analysis of exonic splicing silencers." Cell **119**(6): 831-845.

Wessagowit, V., S. C. Kim, S. Woong Oh and J. A. McGrath (2005). "Genotype-phenotype correlation in recessive dystrophic epidermolysis bullosa: when missense doesn't make sense." J Invest Dermatol **124**(4): 863-866.

West, S., N. Gromak and N. J. Proudfoot (2004). "Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites." Nature **432**(7016): 522-525.

Will, C. L. and R. Luhrmann (2001). "Spliceosomal UsnRNP biogenesis, structure and function." Curr Opin Cell Biol **13**(3): 290-301.

Will, C. L., S. Rumpler, J. Klein Gunnewiek, W. J. van Venrooij and R. Luhrmann (1996). "In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes." Nucleic Acids Res **24**(23): 4614-4623.

Williams, J. H., R. C. Schray, C. A. Patterson, S. O. Ayitey, M. K. Tallent and G. J. Lutz (2009). "Oligonucleotide-mediated survival of motor neuron protein expression in CNS improves phenotype in a mouse model of spinal muscular atrophy." The Journal of neuroscience : the official journal of the Society for Neuroscience **29**(24): 7633-7638.

Wu, S., C. M. Romfo, T. W. Nilsen and M. R. Green (1999). "Functional recognition of the 3' splice site AG by the splicing factor U2AF35." Nature **402**(6763): 832-835.

Yuo, C. Y. and A. M. Weiner (1989). "Genetic analysis of the role of human U1 snRNA in mRNA splicing: I. Effect of mutations in the highly conserved stem-loop I of U1." Genes Dev **3**(5): 697-707.

Zahler, A. M., J. D. Tuttle and A. D. Chisholm (2004). "Genetic suppression of intronic +1G mutations by compensatory U1 snRNA changes in Caenorhabditis elegans." Genetics **167**(4): 1689-1696.

Zhang, X. H. and L. A. Chasin (2004). "Computational definition of sequence motifs governing constitutive exon splicing." Genes Dev **18**(11): 1241-1250.

Zheng, Z. M. (2004). "Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression." J Biomed Sci **11**(3): 278-294.

Zhou, J., X. Zheng and H. Shen (2012). "Targeting RNA-splicing for SMA treatment." Molecules and cells **33**(3): 223-228.

7. APPENDIX

Publications:

Fernandez-Alanis*, Pinotti*, Dal Mas*, Balestra, Cavallari, Rogalska, Bernardi, Pagani "An exon-specific U1 small nuclear RNA (U1snRNA) strategy to correct splicing defects", 2012 *Hum Mol Gen*

Pinotti, Berndardi, Dal Mas, Pagani "RNA-based therapeutic approaches for coagulation factor deficiencies", 2012 *J Thromb Haem*

Pastor*, Dal Mas*, Talotti, Bussani, Pagani "Intron cleavage affect processing of alternatively spliced transcripts", 2011 *RNA*

**Joint First Authors*

Note:

During my 4 years of PhD studies, I followed different projects in the Human Molecular Genetics Laboratory at the International Centre for Genetic Engineering and Biotechnology, under the supervision of dr. Franco Pagani.

During the first year of my PhD studies I was mainly involved in a project regarding intron processing. The results were published, in co-authorship with dr. Pastor in the RNA journal (2011). This research line was not continued after the publication and I shifted on the ExSpe U1 project. Specifically, my contribution to the Human Molecular Genetic paper (2012) is related to all the studies performed on coagulation factor IX minigenes (figures 1A, 2A-B, 3A-B, 6A, 7A-B) and SMN2 minigene (figure 4) and on endogenous SMN transcripts in Hek293 cells (figure 5). In addition I contributed with figure 3 to the review about RNA therapeutic strategies for coagulation factor deficiencies published in 2012 in Journal of Thrombosis and Haemostasis

Human Molecular Genetics, 2012, Vol. 21, No. 11 **2389–2398**
doi:10.1093/hmg/dds045
Advance Access published on February 23, 2012

An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects

Eugenio Fernandez Alanis[1,†], Mirko Pinotti[2,†], Andrea Dal Mas[1,†], Dario Balestra[2], Nicola Cavallari[2], Malgorzata E. Rogalska[1], Francesco Bernardi[2] and Franco Pagani[1,*]

[1]Human Molecular Genetics, International Centre for Genetic Engineering and Biotechnology, Trieste, Italy and [2]Department of Biochemistry and Molecular Biology, University of Ferrara, Italy

Received December 28, 2011; Revised January 24, 2012; Accepted February 9, 2012

A significant proportion of disease-causing mutations affect precursor-mRNA splicing, inducing skipping of the exon from the mature transcript. Using *F9* exon 5, *CFTR* exon 12 and *SMN2* exon 7 models, we characterized natural mutations associated to exon skipping in Haemophilia B, cystic fibrosis and spinal muscular atrophy (SMA), respectively, and the therapeutic splicing rescue by using U1 small nuclear RNA (snRNA). In minigene expression systems, loading of U1 snRNA by complementarity to the normal or mutated donor splice sites (5′ss) corrected the exon skipping caused by mutations at the polypyrimidine tract of the acceptor splice site, at the consensus 5′ss or at exonic regulatory elements. To improve specificity and reduce potential off-target effects, we developed U1 snRNA variants targeting non-conserved intronic sequences downstream of the 5′ss. For each gene system, we identified an exon-specific U1 snRNA (ExSpeU1) able to rescue splicing impaired by the different types of mutations. Through splicing-competent cDNA constructs, we demonstrated that the ExSpeU1-mediated splicing correction of several *F9* mutations results in complete restoration of secreted functional factor IX levels. Furthermore, two ExSpeU1s for SMA improved SMN exon 7 splicing in the chromosomal context of normal cells. We propose ExSpeU1s as a novel therapeutic strategy to correct, in several human disorders, different types of splicing mutations associated with defective exon definition.

INTRODUCTION

Splicing errors represent a significant amount of disease-causing mutations. Considering only changes at canonical splice sites, ~15% of mutations were originally estimated to induce aberrant splicing (1) and similar frequencies are recorded in human disease databases [for example, 12% in cystic fibrosis (CF) and 10% in coagulation FIX (FIX) deficiency (Haemophilia B)]. However, the occurrence of splicing defects is significantly increased (~50%) when genomic variants are systematically evaluated for their effect on precursor-mRNA (pre-mRNA) processing (2,3). This unexpected effect is largely due to mutations located in non-canonical regulatory elements and highlights the difficulty in correctly predicting the consequence of genomic variants on pre-mRNA splicing (4–8).

To correctly identify the exonic sequences from the larger non-coding introns, the splicing machinery recognizes several *cis*-acting elements on the nascent transcripts (9,10). The core elements required for splicing are moderately conserved and consist of the classical or canonical 5′ and 3′ splice sites (named also donor and acceptor sites) and include the polypyrimidine tract as well as the branch site near the 3′ss. The consensus motif of the 5′ss is made of nine-nucleotide sequence, CAG/GURAGU where R is a purine and, with the exception of the nearly obligate GU dinucleotide, the other positions are quite degenerated. Recognition of the exon requires also splicing regulatory elements that are classified, depending on their location and effect on splicing, as exonic/intronic splicing enhancers (ESE and ISE), exonic/intronic splicing silencers (ESS and ISS)

*To whom correspondence should be addressed at: Human Molecular Genetics, International Centre for Genetic Engineering and Biotechnology, Padriciano 99, 34149 Trieste, Italy. Tel: +39 0403757342; Fax: +39 040226555; Email: pagani@icgeb.org
†The authors wish to be known that, in their opinion, the first three authors should be regarded as join First Authors.

(9,10) or composite exonic regulatory elements of splicing (CERES) (7). The exonic elements are composed by largely degenerated sequences, overlap with the coding capacity and interact with splicing factors with a positive (SR proteins) or negative (hnRNPs) effect on the exon recognition. In general, factors associated to the splice sites and to the exonic splicing regulatory elements promote a network of interactions across the exon. As a result, the final outcome of a splicing decision depends on the equilibrium between multiple positive and negative interactions over the exon in a process called exon definition (11).

Aberrant exon skipping is a common splicing defect caused by disruption of the network of interactions that define the exon. It results from mutations at the consensus donor or acceptor sites, at the polypyrimidine tract or in the exonic regulatory elements. In general, mutations at the consensus donor site reduce the complementarity with the U1 small nuclear RNA (snRNA) (see below), substitutions at the 3′ss polypyrimidine tract interfere with the complexes assembled on the 3′ss and exonic mutations affect exonic regulatory elements.

An early event in exon definition is the recognition of the donor site by the U1 snRNP (small nuclear RiboNuclearParticle). U1 snRNP is composed by a 164 bp long U1 snRNA associated to several protein factors. To initiate splicing, the 5′ end of the U1 snRNA interacts by complementarity with the moderately conserved sequence of the 5′ss. Interestingly, ∼40, 22 and 5% of normal 5′ss contains, respectively, two, three or four mismatches toward the U1 snRNA (12). U1 snRNAs with a modified 5′ tail that base pair exactly to the mutant donor sites have been used to correct 5′ss mutation (13–19). For example, we previously showed that a U1 snRNA complementary to the +5G/A mutant of the intron 7 donor splice site of coagulation *F7* rescued splicing and protein biosynthesis and function in minigene systems (14,15). However, this correction approach appears to be limited to single mutations and, through the ability of modified U1 snRNAs to target other donor sites, it can potentially interfere with splicing of other pre-mRNAs.

Here, to study the effect of mutations on splicing and identify a correction strategy on an exon basis, we considered three genes relevant in human pathology, the coagulation *F9*, the Cystic Fibrosis Transmembrane Regulator (*CFTR*) and the *SMN2* (Survival of Motoneuron 2), whose mutations are associated to coagulation factor IX (Haemophilia B), CF and spinal muscular atrophy (SMA), respectively. Expression analysis of several substitutions at the donor splice sites (in *F9* exon 5 and *CFTR* exon 12), at the polypyrimidine tract (in *F9* exon 5) and at exonic regulatory elements (in *CFTR* exon 12 and in *SMN2* exon 7), indicated their ability to cause aberrant splicing and particularly exon skipping. Screening of U1 snRNAs binding at non-conserved intronic sequences downstream the exon identified for each case an unique exon-specific U1 snRNA (ExSpeU1) able to correct different types of mutations. In *SMN*, two ExSpeU1-corrected exon 7 skipping in the normal chromosomal context. Moreover, in a *FIX* splicing-competent expression system, an ExSpeU1 restored, for several mutations, protein secretion and procoagulant function to an extent that, if achieved *in vivo*, would be therapeutic.

RESULTS

Effects of F9 exon 5 and CFTR exon 12 mutations on pre-mRNA splicing

We evaluated nine different natural mutations associated with Haemophilia B deficiency (Fig. 1A) and seven mutations associated to CF (Fig. 1C). The list of the mutations analysed and their associated clinical phenotype are shown in Supplementary Material, Table S1. Mutations are numbered according to their position relative to the donor or acceptor site. The mutations in the *F9* gene consists of two T to G transversions at the 3′ss polypyrimidine tract and seven substitutions at or near the 5′ss of exon 5. In the *CFTR* exon 12, we studied five donor-site substitutions and, as models for exonic mutations that affect splicing, two codon changes (A566T and Y577Y) located in the previously reported CERES element (7) (Supplementary Material, Table S1). To test their effect on splicing, the substitutions were inserted in appropriate minigenes followed by analysis of the splicing pattern in eukaryotic cells. For *F9* exon 5, we prepared a novel minigene, whereas the *CFTR* exon 12 minigene has been previously described and extensively validated (7,8,20). Normal *F9* ex 5 wt minigene showed that the exon was not completely included (∼85%) in the mature mRNA (Fig. 1B, lane 3). This pattern of splicing is entirely consistent with the *in vivo* splicing pattern of the *F9* gene in human liver, where ∼90% of the exon is included (Supplementary Material, Fig. S1). The appreciable level of alternative splicing indicates that the exon 5 is poorly defined, very likely because of the presence of a weak 5′ss. In fact, it differs significantly from the consensus donor site at positions 3, 5 and 6, with a C, a U and an A, respectively (Supplementary Material, Fig. S2). Expression studies of the *F9* ex 5 mutations revealed that the majority of them induced aberrant pre-mRNA processing (Fig. 1A). The −8 and −9 transversions of the polypyrimidine tract and the mutations at the consensus 5′ss at positions −2, −1, +1 and +2 induced complete exon skipping. On the other hand, all mutations in *CFTR* exon 12, either located at the donor sites (in position −1, +3 and +5) or in the exon (A566T and Y577Y), induced exon skipping (Fig. 1D). In is worth noting that the A566T and Y577Y exonic variants affect exonic splicing regulatory elements (7), which are important for correct exon definition. Thus, in two model systems, several natural mutations either at the splice sites or at exonic regulatory elements induce exon skipping.

U1 snRNAs complementary to the mutant 5′ss rescue some defective donor sites

We initially focused on mutations at donor sites, which reduce their complementarity to the 5′ tail of the U1 snRNA. To understand for each mutation the role of U1 snRNA in exon recognition, we prepared U1 snRNAs expression plasmids with compensatory changes that completely restore base-pairing with the wt or mutant 5′ss (Supplementary Material, Figs S2 and S3). These U1 snRNAs were cotransfected with the corresponding minigene variants followed by the analysis of the splicing pattern.

Cotransfection of U1FIXwt with the ex5 wt minigene improved the exon recognition inducing complete inclusion

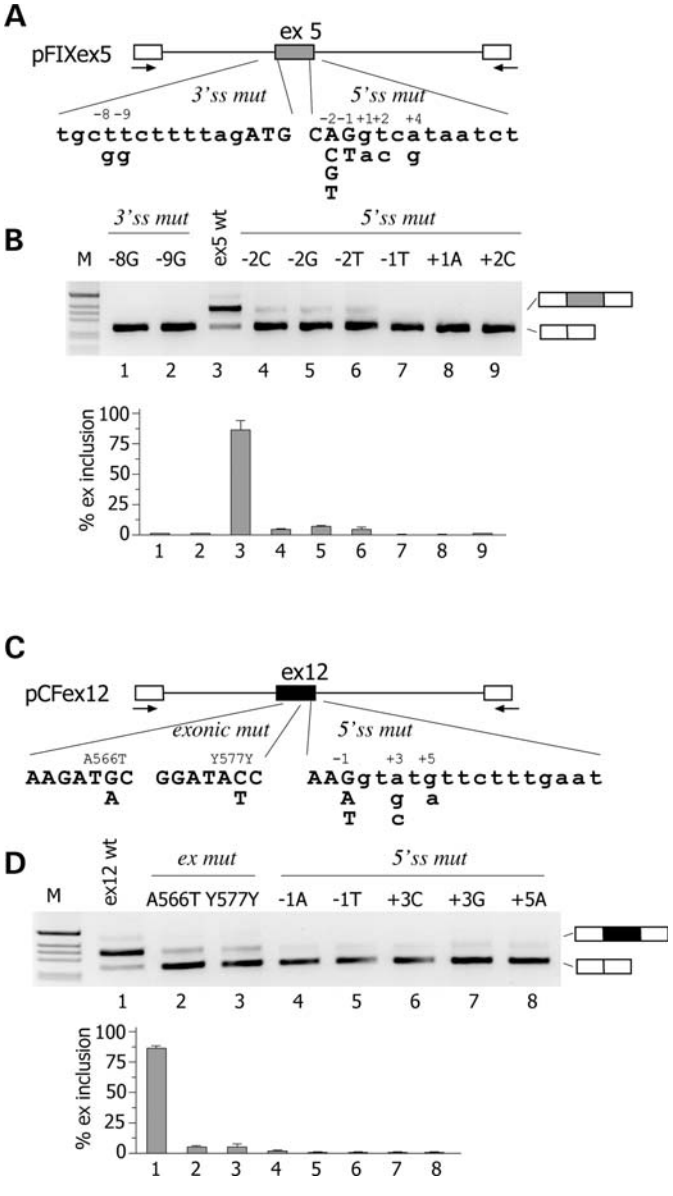


Figure 1. Effect of natural mutations on *F9* exon 5 and *CFTR* exon 12 pre-mRNA splicing. (**A** and **C**) Schematic representation of the central region of the pFIX exon 5 and pCFex12 minigenes, respectively. Exonic sequences are boxed, introns are lines and the arrows represent primers for PCR amplification. The position of genomic variants relative to the splice sites is indicated. Exonic and intronic sequences are in upper and lower case, respectively. (**B**) Analysis of pFIX exon 5-spliced transcripts. HeLa cells were transfected with pFIXex5 wt or mutant minigenes and splicing pattern evaluated by RT–PCR. Amplified products were resolved on a 2% agarose gel. The identity of the bands is indicated on the right of the panel. M is the molecular 1 Kb marker. Lower panel shows the quantification of the percentage exon 5 inclusion (mean ± SD of three independent experiments done in duplicate). (**D**) Analysis of pCF exon 12 spliced transcripts. HeLa cells were transfected with pFIX ex12 wt or mutant minigenes and splicing pattern evaluated by RT–PCR. Amplified products were resolved on a 2% agarose gel. Lower panel shows the quantification of the percentage exon 12 inclusion (mean ± SD of three independent experiments done in duplicate).

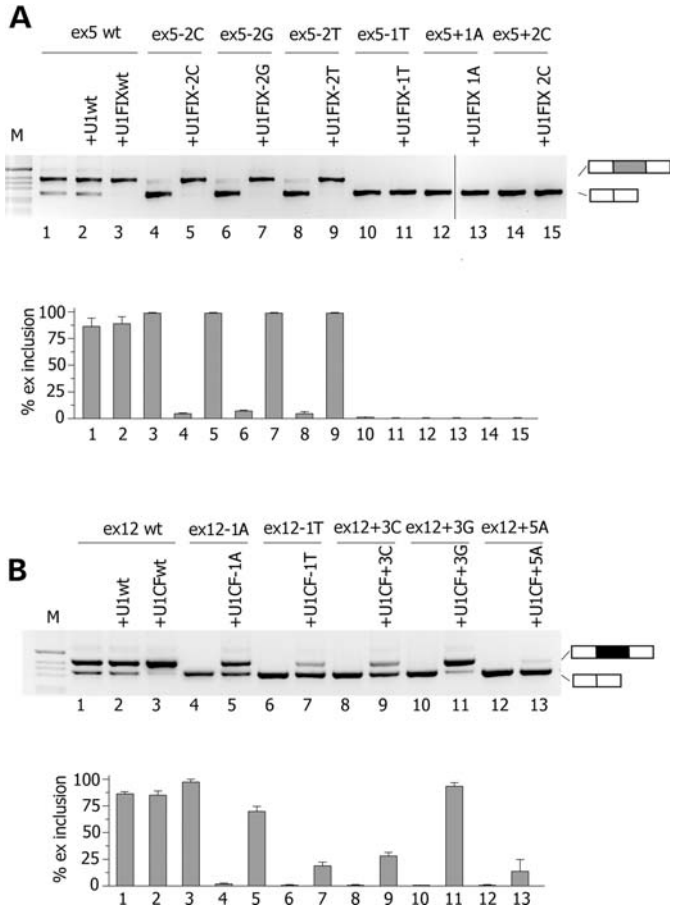


Figure 2. Effect of increased complementarity of U1 snRNAs to normal and mutated donor splice site on pre-mRNA splicing. (**A**) Effect of complementary U1 snRNA to the *F9* exon 5 5′ss. The upper panel shows the analysis of spliced transcripts. Minigenes were transfected in HeLa cells with the empty vector (lanes 1, 4, 6, 8, 10, 12, 14) or with modified U1s (lanes 3, 5, 7, 9, 11, 13, 15) and splicing pattern evaluated by RT–PCR. Amplified products were resolved on a 2% agarose gel. The identity of the U1 snRNAs is shown in Supplementary Material, Figure S2. Lower panel is the quantification of exon 5 splicing pattern after co-transfection with modified U1 snRNAs. Data are expressed as means ± SD of three independent experiments done in duplicate. (**B**) Effect of complementary U1 snRNAs to the *CFTR* exon 12 5′ss. Minigenes were transfected in HeLa cells with the empty vector (lanes 4, 6, 8) or with modified U1s (lanes 3, 5, 7, 9) and the splicing pattern evaluated by RT–PCR. Amplified products were resolved on a 2% agarose gel. The identity of the U1 snRNAs is shown in Supplementary Material, Figure S2. Lower panel is the quantification of exon 12 splicing pattern. Data are expressed as means ± SD of three independent experiments done in duplicate.

of the exon (Fig. 2A, lanes 1 and 3). Noticeably, the three synonymous changes at position −2 showed a remarkable increase in the percentage of exon inclusion after cotransfection of the corresponding complementary U1 snRNAs, with a negligible presence of aberrant transcripts (Fig. 2A, lanes 4–9). On the contrary, the splicing pattern of the other FIX mutations −1T, +1A and +2C was not affected by the corresponding complementary U1 snRNAs.

Minigene experiments with the *CFTR* exon 12 showed that all mutations were rescued by the corresponding complementarity U1 snRNAs. Whereas the −1T, +3C and +5A were slightly enhanced (Fig. 2B), the severe splicing defects caused by −1A and +3G mutations were efficiently corrected by the complementary U1 snRNAs, reaching ∼75 and ∼95% of exon inclusion. Cotransfection of U1 CFwt with the CFex12 wt also improved the splicing pattern (Fig. 2B, lane 3).

These results indicate that the three synonymous −2 variants in *F9* exon 5 and the donor 5′ss mutations in *CFTR* exon 12 do not completely disrupt the 5′ss function. They produce a defective donor site, whose recognition by complementary U1 snRNA variants recovers correct exon inclusion.

U1 snRNAs complementary to intronic sequences downstream of the 5′ss correct aberrant splicing

To promote exon definition, U1 snRNP does not necessarily have to perfectly bind at the 5′ss. Some atypical 5′ss are recognized by U1 snRNA shifted by one nucleotide (21) and U1 snRNA complementary to downstream intronic sequences were originally reported to promote exon inclusion in model gene system (22,23). Thus, we tested the effect of U1 snRNA-binding downstream mutant *CFTR* exon 12 and *F9* exon 5. For this purpose, we prepared different U1 snRNAs, whose 5′ tails were modified to base pair at variable distance downstream of the 5′ss of *F9* exon 5 and *CFTR* exon 12 (Fig. 3). This approach was also extended to the extensively investigated *SMN2* exon 7, where a weak constitutive 5′ss and a synonymous exonic substitution are associated to exon skipping (24–26). In *F9*, we systematically screened U1 snRNAs from position −7 till position 63 relative to the donor site junction and the analysis was performed on the ex5 −2C mutant minigene. As shown in Figure 3A, all U1 snRNAs tested reduced exon 5 skipping, indicating their positive effect on exon definition. In particular fix 1, fix 9 and fix 10 U1 snRNAs showed the strongest effect with a nearly complete rescue of aberrant splicing. In *CFTR* exon 12, only the U1 snRNA variant cf11 induced a significant (70%) rescue of the splicing pattern in the +3G mutant (Fig. 3D). The other U1 snRNAs had a modest effect (cf1 and cf9) or no effect (cf15 and cf33) on splicing (Fig. 3D). In the SMN2 minigene, ∼20% of exon 7 is included in the final transcript, whereas cotransfection of sm2, sm17 and sm21 resulted in >80% of exon inclusion (Fig. 4B).

These results in three different gene systems indicate that U1 snRNAs binding downstream the 5′ss corrected exon skipping caused by defective 5′ss.

SMN-specific U1 snRNAs rescue exon 7 splicing in normal cells

To study the effect of different ExSpeU1s in the chromosomal context, we focussed on SMN exon 7. Normal individuals have two SMN copies, SMN1 and SMN2, that mainly differ for a synonymous substitution in exon 7. This substitution induces SMN2 exon 7 skipping and because the majority of SMA patients lack SMN1, its splicing correction is at the basis of several therapeutic attempts (27,28). To evaluate the potential therapeutic effect of SMN-specific U1s in SMA, we transfected sm17 and sm21 in HEK293 cells and evaluated the endogenous pattern of splicing. Semiquantitative analysis of splicing isoforms showed that sm17 and sm21 induced a significant increase in the percentage of SMN exon 7 inclusion. In particular, transfection of the two ExSpeU1s increased the percentage of SMN2 exon 7 from ∼40 to ∼70% (Fig. 5).

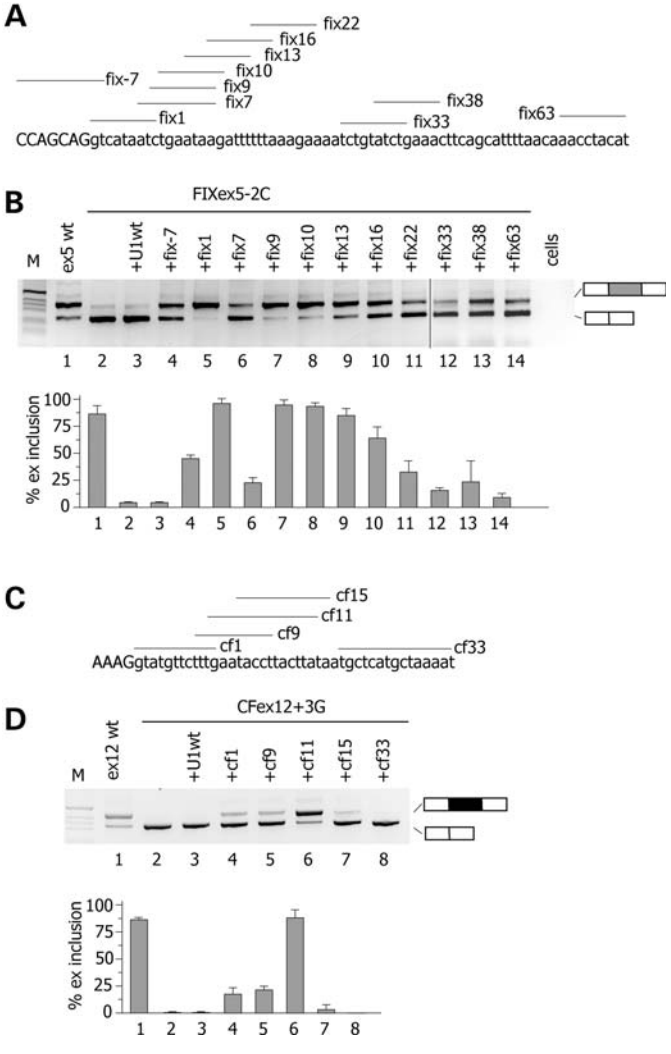


Figure 3. Identification of ExSpeU1s in *F9* exon 5 and *CFTR* exon 12. (**A** and **C**) Schematic representation of binding sites of the ExSpeU1s in *F9* exon 5 and *CFTR* exon 12, respectively. Donor site and downstream intronic region are shown and exonic and intronic sequences are in upper and lower case, respectively. Lines correspond to the target-binding sequences of ExSpeU1s U1 snRNAs on the nascent pre-mRNA. (**B**) Analysis of *F9* exon 5-spliced transcripts. The FIXex5-2C mutant minigene was transfected in HeLa cells alone (lane 1), with the empty vector (lane 2), or with plasmids encoding for the different ExSpeU1s. The pattern of splicing was evaluated by RT–PCR and amplified products were resolved on a 2% agarose gel. Lower panel is the quantification of exon 5 splicing pattern. (**D**) Analysis of *CFTR* exon 12-spliced transcripts. The mutant ex12 + 3G minigene was transfected in HeLa cells alone (lane 1), with the empty vector (lane 2), or with plasmids encoding for the ExSpeU1s. The pattern of splicing was evaluated by RT–PCR and amplified products were resolved on a 2% agarose gel. Lower panel is the quantification of exon 5 splicing pattern.

A unique exon-specific U1 snRNA can rescue splicing of defective splice site mutants and exonic variants

We subsequently tested the effect of the most active U1 snRNAs (fix9 for FIX and cf11 for CFTR), named exon-specific U1 snRNAs (ExSpeU1), on all splicing mutations. In *F9* exon 5, fix9 ExSpeU1 completely rescued the defective donor sites of the three synonymous mutations in position −2 (Fig. 6A, lanes 5–10). At variance, the −1T, +1A and +2C 5′ss mutations, not corrected by loading the U1 snRNA directly on the

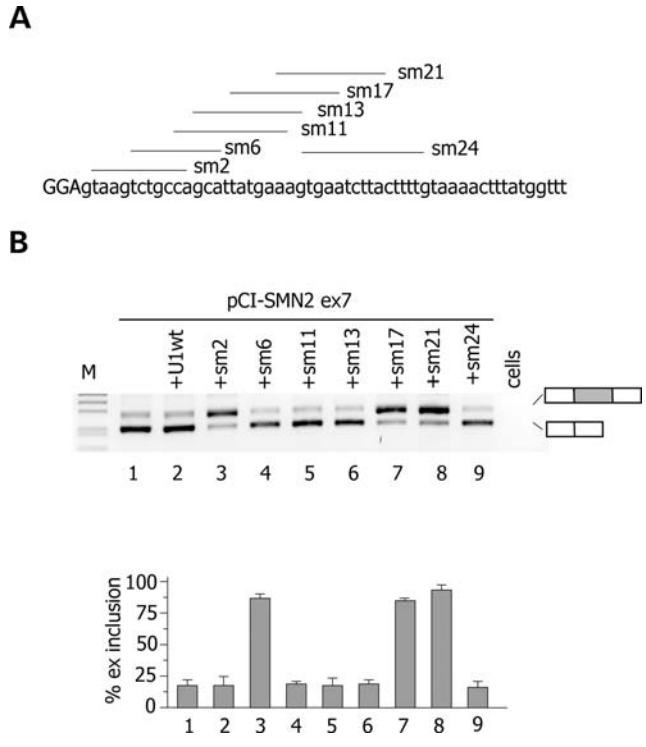


Figure 4. Identification of ExSpeU1s in *SMN2* exon 7. (**A**) Schematic representation of binding sites of the ExSpeU1s in *SMN2* exon 7. Donor site and downstream intronic region are shown and exonic and intronic sequences are in upper and lower case, respectively. Lines correspond to the target-binding sequences of ExSpeU1s on the nascent pre-mRNA. (**B**) Analysis of *SMN2* exon 7 spliced transcripts. The pCI-SMN2 minigene was transfected in HeLa cells alone (lane 1), with the U1wt (lane 2), or with plasmids encoding for ExSpeU1s. The pattern of splicing was evaluated by RT–PCR and amplified products were resolved on a 2% agarose gel. Lower panel is the quantification of *SMN2* exon 7 splicing pattern. Data are the means ± SD of three experiments done in duplicate.

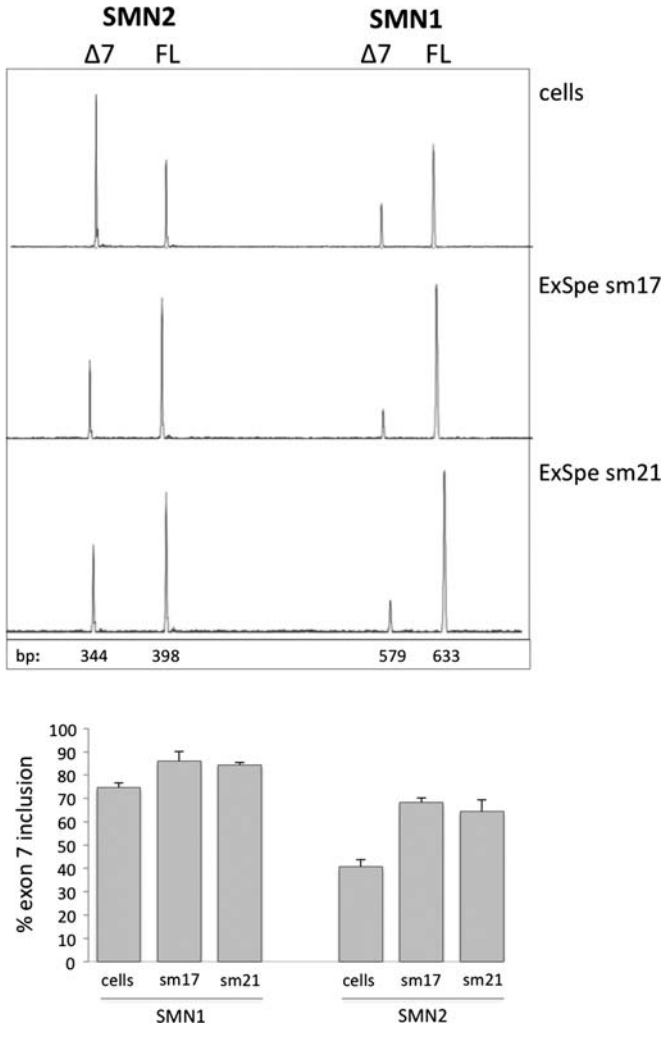


Figure 5. *SMN* exon 7 splicing correction mediated by sm17 and sm21 ExSpeU1s in normal cells. The upper panel shows a representative experiment in which the fluorescently labelled RT–PCR amplified fragments were separated on denaturing capillary electrophoresis. HEK393 cells were transfected with the indicated ExSpeU1s and amplified fragments were digested with *Dde*I to obtain SMN1 and SMN2 exon 7 inclusion (FL) and exclusion (Δ7) fragments. The size of the fragments is indicated. Lower panel shows the percentage of SMN1 and SMN2 exon 7 inclusion in ExSpeU1-treated and -untreated cells and data are expressed as means ± SD of three independent experiments done in duplicate. Transfection efficiency in the different experiments was 75–85%.

donor site (Fig. 2A), were also not influenced by ExSpeU1 fix9 (Fig. 6A, lanes 11–16). Interestingly, cotransfection of ExSpeU1 fix9 induced nearly complete exon inclusion of the two −8G and −9G intronic variants of the polypyrimidine tract (Fig. 5A, lanes 1–4). In *CFTR* exon 12, ExSpeU1 cf11 rescued all 5′ss mutations with different efficiency. After ExSpeU1 cf11 cotransfection, the +3G mutation showed ∼85% of exon inclusion (Fig. 6B); the percentage of exon inclusion in −1A, +3C and +5A increased to ∼50%, whereas in −1T it improved to ∼20% (Fig. 6B). Intriguingly, we observed complete splicing correction by ExSpeU1 cf11 of the two exonic A566T and Y577Y mutations (Fig. 6B, lanes 1–4). Sequencing of ExSpeU1-induced products showed correct usage of the normal intron–exon junction and no activation of cryptic splice sites. To clarify if the enhancing effect is specific for the U1 snRNA particle, we inserted the tail of two active ExSpeU1s, fix10 and cf11, in the U7 snRNA. U7 is not involved in splicing and a modified version has been extensively used to express antisense RNA sequences (29–35). The resulting U7 fix10 and U7 cf11 were cotransfected with −8G, −2T mutants and +3G and Y577Y, respectively. In contrast to the ExSpeU1s, the corresponding U7s had no effect (Supplementary Material, Fig. S4), suggesting that U1 snRNP-specific proteins are required for splicing rescue and that ExSpeU1s are not just targeting intronic sequences with silencer function.

Altogether, these data indicate that recruitment of ExSpeU1 by complementarity on intronic sequences downstream of the exon corrects multiple splicing mutations located either at 5′ or 3′ splice sites or in exonic regulatory elements. This provides a novel therapeutic strategy exploiting a unique ExSpeU1 to correct a panel of splicing defects.

Rescue of splicing by the F9 exon5-specific U1 snRNA results in FIX biosynthesis and coagulant activity

To establish whether the ExSpeU1-mediated correction of splicing of the different mutants result in a consistent rescue

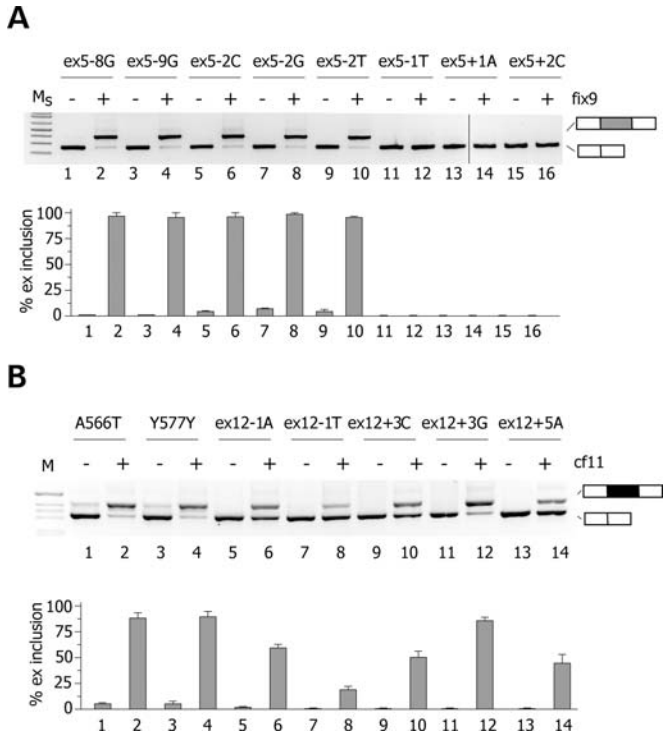


Figure 6. Splicing correction mediated by fix9 and cf11 U1 ExSpeU1s on natural mutations. (**A**) HeLa cells were transfected with the indicated FIX exon 5 mutant minigenes alone (even lanes) or along with fix9 U1 ExSpeU1 (odd lanes). Splicing pattern was evaluated by RT–PCR and amplified products were resolved on a 2% agarose gel. Ms is the molecular 1 Kb plus marker. Lower panel shows the quantification of exon 5 splicing pattern. Percentage of exon inclusion is expressed as means ± SD of three experiments done in duplicate. (**B**) HeLa cells were transfected with the indicated *CFTR* exon 12 mutant minigenes alone (even lanes) or along with cf11 ExSpeU1 (odd lanes). Splicing pattern was evaluated by RT–PCR and amplified products were resolved on a 2% agarose gel. Lower panel shows the quantification of exon 12 splicing pattern. Percentage of exon inclusion is expressed as means ± SD of three experiments done in duplicate.

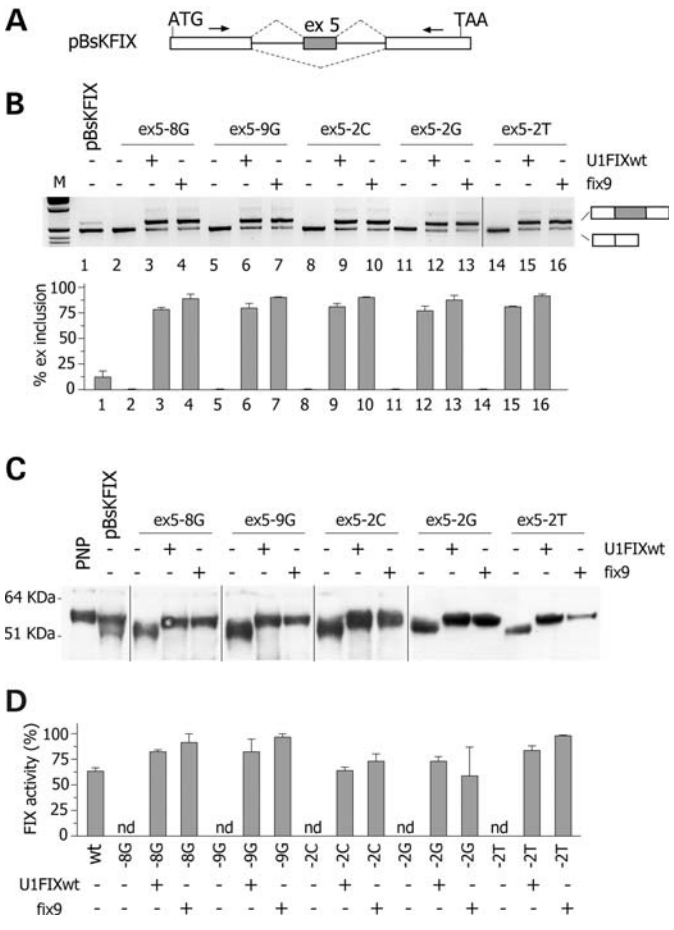


Figure 7. Rescue of splicing and FIX protein function by ExSpeU1. (**A**) Schematic representation of the pBsK-FIX minigene. Exonic sequences are boxed, introns are lines, the arrows represent primers for PCR amplification and the dotted lines indicate the two possible exon 5 alternative splicing events. The position of the ATG and TAA codons is indicated. The minigene is not drawn in scale. (**B**) Analysis of pBsK-FIX exon 5 spliced transcripts. pBsK-FIX exon 5 normal and mutant minigenes were transfected in BHK cells alone or with the indicated ExSpeU1s. The splicing pattern was evaluated by RT–PCR and amplified products were resolved on a 2% agarose gel. M is the molecular 1 Kb marker. Lower panel shows the quantification of the percentage exon 5 inclusion. Data are expressed as means ± SD of three independent experiments done in duplicate. (**C**) Western blotting of FIX in the BHK conditioned medium. PNP is pooled normal plasma. (**D**) FIX coagulant activity in the BHK conditioned medium. Data are expressed as means ± SD of three experiments done in duplicate.

of protein biosynthesis and function, we took advantage of the FIX model. FIX is a serine protease secreted from cells that can be finely monitored by protein and functional assays. The splicing mutations corrected by the ExSpeU1 fix9 (the three −2 variants at the 5′ss and the two transversions at the polypyrimidine tract) were tested in a splicing-competent cDNA minigene context (pBsKFIX) (Fig. 7A). In this minigene, exon 5 was inserted in *FIX* cDNA transcript along with part of its intronic sequences in a manner that the correction of exon skipping produces a normal transcript with secretion of a functional protein (Fig. 7A). The pBsKFIX exon 5 wt and mutant minigenes were expressed in BHK cells with or without the modified U1 snRNAs complementary to the *F9* donor site (U1FIXwt) or to the downstream intronic sequence (fix9).

Probably caused by the pBsKFIX minigene context, this construct showed exon 5 inclusion in ∼10% of the transcripts (Fig. 7, lane 1), while experiments with the pTB minigene (Fig. 1) and *in vivo* (Supplementary Material, Fig. S3) showed a ∼85% inclusion. In spite of this inefficient exon 5 recognition, the secreted FIX levels in the conditioned medium were appreciable by ELISA (13 ± 1 ng/ml), western blotting (Fig. 7C) and coagulation assays (Fig. 7D). In particular, western blotting revealed the presence of a band corresponding to the full-length FIX form (∼60 kDa), which migrated as plasma-derived FIX. In addition, a band corresponding to a smaller size FIX variant (∼51 kDa) was also clearly detectable. This arises from translation of the in frame FIX mRNA form lacking exon 5 (Fig. 7C), leading to a FIX molecule lacking 43 amino acids of the second epidermal growth factor like domain 2 (EGF2). Notably, the mRNA splicing and western blot patterns showed remarkably different relative amount of the two isoforms, with the aberrant form predominating at the mRNA (Fig. 7B) but not at the protein (Fig. 7C) level, a finding compatible with inefficient biosynthesis of the protein without the EGF2 domain.

Transfection of the five pBsKFIX mutant minigenes showed complete skipping of the exon (Fig. 7B, lanes 2, 5, 8, 11, 14), and cotransfection of either U1FIXwt or fix9 ExSpeU1 induced a complete rescue of the splicing pattern. In all five mutants, the percentage of exon inclusion increased to level above 75% (Fig. 7B). The extent of splicing rescue mediated by these U1 snRNAs was remarkable, taking into account the inefficient exon 5 recognition in this context. Results from investigations at the protein level in the conditioned medium were consistent with those from splicing assays. The expression of the five mutants resulted in secretion of deleted FIX proteins (Fig. 7C), which did not display any appreciable coagulant activity (Fig. 7D). At variance, co-expression of variants with either U1FIXwt or fix9 ExSpeU1 induced the exclusive synthesis of the full-length FIX (Fig. 7C). This was paralleled by a complete rescue of FIX activity (Fig. 7D), which for most variants was higher than that measured for the pBsKFIX wt construct.

DISCUSSION

In this study, we provide a novel strategy to correct different types of natural splicing mutations using exon specific U1 snRNAs (ExSpeU1). ExSpeU1s bind by complementarity to intronic sequences downstream of the exon and rescue different types of splicing defects associated to exon skipping. ExSpeU1s are active on several 5′ss mutations in *CFTR* exon 12 and *F9* exon 5, on two transversions at the polypyrimidine tract in *F9* exon 5, on two exonic substitutions in *CFTR* exon 12 and on defective *SMN2* exon 7 splicing (also due to an exonic variant). In this latter case, ExSpeU1 rescued *SMN2* exon 7 splicing directly in the chromosomal context of normal cells, providing a new therapeutic strategy for SMA. On the other hand, in the model of Haemophilia B, a unique ExSpeU1 induced complete splicing correction of five different *F9* mutations, thus resulting in a complete rescue of protein biosynthesis and coagulation activity.

A limited number of studies have explored the role and potential therapeutic effect of U1 snRNAs on splicing correction of donor site mutations (13–19). In all cases, the modified tails of U1 snRNA have few nucleotide changes in comparison to the WT sequences and base pair exactly to the mutant donor sites. Thus, their therapeutic effect is restricted to each mutation and the potential complementarity of these modified U1 snRNAs to other normal 5′ss might affect splicing of other exons, in particular those alternatively spliced. The ExSpeU1s we have developed here, which bind to non-conserved intronic sequences, have two major advantages, (i) they do not interact directly with normal 5′ss and (ii) correct different types of splicing defects associated to exon skipping. The mutations rescued by ExSpeU1s affect three different splicing regulatory elements important for exon recognition: the 5′ss, the polypyrimidine tract of the 3′ss and the exonic splicing regulatory sequences. To our knowledge, this is the first time that an U1 snRNA variant is shown to correct exon skipping due to polypyrimidine tract and exonic mutants. *SMN2* is a paradigmatic example of exon skipping caused by a synonymous variant in an exonic regulatory element and its correction is at the basis of several therapeutic attempts in SMA (27,28).

Compared with classical gene replacement therapies, ExSpeU1-mediated rescue has a number of additional advantages and stimulating perspectives. The direct splicing correction maintains the regulation of the gene expression in the correct cell-specific chromosomal context under the control of endogenous transcription and pre-mRNA processing regulatory elements. The short length of the ExSpeU1s cassette (~500 bp) can also be useful in gene therapy of splicing mutations in large genes such as *CFTR* and *F8*, whose full-length transcript can represent a limiting step for their insertion in viral vectors such as AAV. On the other hand, to achieve the optimal therapeutic rescue, the expression level could be modulated *in vivo* by varying the number of ExSpeU1 cassette in the viral vector. In the case of dominant-negative mutations where the replacement therapy is not feasible, the splicing correction mediated by ExSpeU1 will act directly reducing the amount of the mutated toxic protein. Moreover, binding of the ExSpeU1s to intronic sequences, which are not conserved, will significantly reduce the possibility of off-target events. Although binding of ExSpeU1s to unrelated sequences not directly involved in splicing might have an effect on other pre-mRNA processing steps (36), our result indicate that it will be possible to design ExSpeU1s with improved specificity by inserting subtle changes in sequence target and/or length. For example, ExSpeU1s sm17 and sm21 in *SMN2* (Fig. 3B) or fix9, fix10 and fix13 in *F9* exon 5 (Fig. 3B) bind at nearby or overlapping intronic sequences and are functionally active. Future studies in cellular and animal models will be required to address this point.

According to the classical exon definition model, a network of interactions across the exon put in contact the splicing complexes assembled on the splice sites (11,19,37). This includes the recognition of the 5′ss by the U1 snRNP, binding of U2AFs to the polypyrimidine tract and identification of exonic splicing regulatory elements by SR proteins. The mutations we have analysed here affect differently these regulatory elements and probably block splicing at the first step when the exon has to be recognized co-transcriptionally. At this stage, recruitment of U1 snRNA at the 5′ss or at downstream sequences by means of ExSpeU1s can rescue the different defects. This occurs not only for mutations at the consensus 5′ss that directly interfere with U1 snRNA binding but also for mutations located at distance from the donor site in the polypyrimidine tract or in exonic regulatory elements. In these cases, ExSpeU1s probably compensate the missing interactions and facilitate the formation of the correct network of splicing factors over the exon.

Interestingly, the two ExSpeU1s for *SMN2* (sm17 and sm21) are partially complementary to a previously reported intronic splicing silencer, which binds to the negative splicing factor hnRNPA1 (38). Targeting this ISS with antisense oligonucleotides resulted in inclusion of *SMN2* exon 7 in model minigenes and in cellular systems (39) and was recently shown to improve the phenotype in a SMA mouse model (27). Thus, it is possible that ExSpeU1s facilitates recruitment of the splicing machinery on the exon interfering with intronic splicing regulatory sequences located downstream the donor site. However, this mechanism seems unlikely for *F9* exon 5

and *CFTR* exon 12, as modified U7 particles binding to the target intronic sequences had no effect on splicing (Supplementary Material, Fig. S4). It is possible that ExSpeU1 acts at different levels depending on the architecture and relative strength of splicing regulatory elements in each exon. More in general, ExSpeU1s-mediated induction of exon inclusion could be used to regulate the synthesis of specific alternatively spliced isoforms for therapeutic purposes. For example, therapeutic induction of specific alternative splicing isoforms affect tumour progression (40), angiogenesis (41,42) or aging (43), and modified U1 snRNAs have been recently used to promote splicing to inhibit HIV replication (44).

Even if the disease-causing splicing mutations here investigated have a comparable effect on mature transcript, they behave differently regarding their sensitivity to ExSpeU1-mediated correction. The −1T +1T and +2C mutations in *F9* exon 5 did not respond to the ExSpeU1 fix9, whereas the 5′ss mutations in *CFTR* exon 12 showed a variable response to cf11 ExSpeU1: in this case, the rescue efficiency was optimal for the +3G, intermediate for the −1A, +3C and +5A and low for the −1T (Fig. 5). As we have previously suggested (19), the different rescue efficiency to ExSpeU1 indicates that donor site mutations are mechanistically different. Interestingly, mutations that respond to ExSpeU1 (Fig. 2) were also sensitive to U1 snRNA complementary to the mutated 5′ss sequence (Fig. 5), suggesting a common effect of U1 snRNAs regardless of its loading position. Non-responsive mutations could affect the splicing progression after the exon definition step and their rescue would require the complementation with additional splicing factors. Even if mutations in positions +1 and +2 are not expected to respond to ExSpeU1 due to the obligate presence of the GT dinucleotide at the 5′ss, the mechanism that regulates the rescue efficiency of the other mutations is less obvious. Some positions are known to interact with other splicing factors in later splicing steps (like the −1 and +5 positions to U5 snRNA and U6 snRNA, respectively) (45,46). More studies are necessary to fully address the mechanism and identify the factors involved in determining the ExSpeU1 responsiveness of this type of splicing defects at the donor site.

In conclusion, a single ExSpeU1 appropriately loaded by complementarity to intronic sequences downstream of the exon rescues correct splicing impaired by different mutations and recover the corresponding protein biosynthesis and function in model gene systems. The functional levels obtained in the model of Haemophilia B and in *SMN2* would produce *in vivo* a therapeutic correction of the bleeding defect and of the SMA phenotype, respectively. ExSpeU1s constitute a promising novel therapeutic strategy to correct splicing defects associated to defective exon definition in several human disorders.

MATERIALS AND METHODS

Hybrid minigene constructs

pCF exon 12 minigene has been previously described (7) and exonic and donor splice site mutations were introduced substituting the *Acc*I-*Bam*HI cassette by polymerase chain reaction (PCR)-mediated site-directed mutagenesis. To obtain the pFIX exon 5 minigene, the FIX exon 5 fragment consisting of the last 314 bp of intron 4, exon 5 (129 bp) and the first 278 bp of intron 5 was amplified from normal genomic DNA using FIXex5dir and FIXex5rev and cloned into the pTB *Nde*I-minigene (7). Polypyrimidine tract and donor splice site mutations were introduced in pFIXex5 between the unique *Pst*I and *Xba*I sites of pFIX exon 5 by PCR-mediated site-directed mutagenesis. pCI-SMN2 exon 7 was obtained from Dr Adrian Krainer (CSHL, NY, USA). Modified and exon-specific U1 snRNAs were created by replacing the sequence between the sites *Bcl*I and *Bgl*II with oligonucleotides as previously reported (47). Sequences of oligonucleotides are provided in Supplementary Material, Table S2. Hybrid minigenes were verified by the sequence analysis. Modified U7 snRNAs (U7 fix10 and U7 cf11) were created by PCR amplification of U7SmOPT vector using cfsh11U7 and FIXsh10U7 and Sp6 primers. PCR products were digested with *Hin*dIII and *Stu*I and ligated into *Hin*dIII/*Stu*I sites of the U7SmOPT vector and the resulting clones verified by the sequence analysis.

The splicing-competent FIX cDNA expression cassette was synthesized by GeneScript Inc. (Piscataway, NJ, USA) and its sequence is available upon request. The cassette, inserted in the pCDNA 3.1+ backbone to make the pBsK-FIX, consists of a simian virus 40 (SV40) promoter followed by (i) the human FIX cDNA sequence from exons 1 to 4 and the 5′ portion (544 bp) of intron 4, (ii) a unique *Nde*I restriction site, (iii) the last 314 bp of intron 4, exon 5 and the 5′ portion (278 bp) of intron 5, (iv) a *Nde*I restriction site, (v) the 3′ end (908 bp) of intron 5 followed by the cDNA sequence from exons 6 to 8, and (vi) the SV40 polyadenylation site. The FIX cDNA fragments were amplified with high-fidelity DNA polymerase from the pCMV5-FIX vector. The *Nde*I–*Nde*I cassettes containing the last 314 bp of intron 4, the exon 5 and the 5′ portion (278 bp) of intron 5 were subcloned into pBsK-FIX to make the mutant variants for investigation of rescue at the protein level.

Analysis of hybrid minigene expression and SMN splicing

HeLa (human cervical carcinoma), HEK393 and BHK cell lines were grown in Dulbecco's modified Eagle's medium with Glutamax I (Gibco) (DMEM with glutamine, sodium pyruvate, pyridoxine and 4.5 g/l glucose) supplemented with 10% fetal calf serum (Euro Clone) and Antibiotic Antimycotic (Sigma) according to the manufacturer's instructions. HeLa cells grown on six well plates were transfected with Effectene reagents (Qiagen) according to the manufacturer's protocol. 0.5 μg of hybrid minigenes were transfected either alone or with 0.5 μg of wt/mutant U1 snRNA-encoding plasmids. Total RNA extraction was performed after 24 h of incubation using TRIreagent (Invitrogen) and reverse transcription reaction was carried out as described (47). alpha2,3 and Bra2 oligonucleotides were used for amplification of pCF exon 12 and pFIX exon 5; T7-F2 and E8-75 + 5′R oligonucleotides for pCI-SMN2. The conditions used for the PCRs were 94°C for 5 min for the initial denaturation, 94°C for 45 s, 56°C for 45 s, 72°C for 45 s for 35 cycles and 72°C for 10 min for the final extension. PCR products were resolved on 2%

agarose gel electrophoresis. Quantification of exon inclusion was performed using the ImageJ software.

For *SMN2* exon 7 analysis, HEK393 cells grown in six wells were transfected with 2 μg of each ExSpeU1s expression plasmid with the calcium-phosphate method and RT–PCR performed with E8-467-R and fluorescently labelled FAM-E6-F primers. Reactions were incubated at 95°C for 3 min followed by 35 cycles at 95°C for 30 s, 55°C for 30 s and 72°C for 36 s. Amplified fragments were digested with *Dde*I and separated on denaturing capillary electrophoresis (ABI-3100). Quantification of intensity of SMN1 and SMN2 exon 7 inclusion and exclusion bands was performed with Peakscanner™ software.

Transfection of BHK splicing-competent pBsK-FIX expression vector and reverse transcription was done as described (14). PCR reaction was carried out using pBsK-FIXdir and pBsK-FIXrev oligonucleotides. PCR products were resolved by 2% agarose gel electrophoresis. Liver RNA (First Choice Human Total RNA Survey Panel, Ambion, Inc.) was retrotranscribed in standard conditions and amplified with FIX140 and FIX279 primers.

FIX activity and protein assays

FIX coagulant activity was assessed by the aPTT coagulation assay (48). FIX antigen levels in the conditioned medium were evaluated by ELISA (Factor IX antigen, FIX; Affinity Biologicals, Ancaster, Canada). For western blotting analysis, 26 μl of the conditioned medium were incubated 5 min at 95°C and run on 4–12% SDS–PAGE (NuPAGE Bis–Tris gel, Invitrogen®, Carlsbad, CA, USA). Proteins were transferred onto a 0.2 μm nitrocellulose membrane (Whatman®, Dassel, Germany), which was blocked over night with PBS buffer supplemented with 0.1% Tween-20 (PBS-T) and 5% low fat dry milk (Bio-Rad, Hercules, CA, USA). Membranes were then incubated for 3 h at room temperature with an anti-Human F.IX peroxidase conjugated (GAFIX-APHRP; Affinity Biologicals). The Supersignal® West Femto reagent (Thermo Scientific, Rockford, IL, USA) was exploited for detection. Plasma derived FIX or rFIX-wt were used to optimize the assay.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank Andres Muro and Marco Baralle for critical reading of the manuscript, Adrian Krainer for the pCI-SMN2 exon 7 minigene and D. Schümperli for the U7SmOPT plasmid.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by grants from the Telethon foundation (GGP09183), the Italian Cystic Fibrosis Foundation no (FFC#9/2009) adopted by Dedikato, CIV and Delegazione FFC del Lago di Garda, the Ministero dell'Università e della Ricerca (MIUR)-Progetti di Ricerca di Interesse Nazionale (PRIN), the University of Ferrara and Fondazione CARIFE. Funding to pay the Open Access publication charges for this article was provided by Comitato Telethon Fondazione Onlus.

REFERENCES

1. Krawczak, M., Reiss, J. and Cooper, D.N. (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.*, **90**, 41–54.
2. Teraoka, S.N., Telatar, M., Becker-Catania, S., Liang, T., Onengut, S., Tolun, A., Chessa, L., Sanal, O., Bernatowska, E., Gatti, R.A. *et al.* (1999) Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *Am. J. Hum. Genet.*, **64**, 1617–1631.
3. Ars, E., Serra, E., Garcia, J., Kruyer, H., Gaona, A., Lazaro, C. and Estivill, X. (2000) Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum. Mol. Genet.*, **9**, 237–247.
4. Pagani, F. and Baralle, F.E. (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat. Rev. Genet.*, **5**, 389–396.
5. Cartegni, L., Chew, S.L. and Krainer, A.R. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.*, **3**, 285–298.
6. 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.
7. Pagani, F., Stuani, C., Tzetis, M., Kanavakis, E., Efthymiadou, A., Doudounakis, S., Casals, T. and Baralle, F.E. (2003) New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum. Mol. Genet.*, **12**, 1111–1120.
8. Pagani, F., Raponi, M. and Baralle, F.E. (2005) Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. *Proc. Natl Acad. Sci. USA*, **102**, 6368–6372.
9. Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.*, **10**, 741–754.
10. Nilsen, T.W. and Graveley, B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
11. Berget, S.M. (1995) Exon recognition in vertebrate splicing. *J. Biol. Chem.*, **270**, 2411–2414.
12. Carmel, I., Tal, S., Vig, I. and Ast, G. (2004) Comparative analysis detects dependencies among the 5′ splice-site positions. *RNA*, **10**, 828–840.
13. Baralle, M., Baralle, D., De Conti, L., Mattocks, C., Whittaker, J., Knezevich, A., Ffrench-Constant, C. and Baralle, F.E. (2003) Identification of a mutation that perturbs NF1 agene splicing using genomic DNA samples and a minigene assay. *J. Med. Genet.*, **40**, 220–222.
14. Pinotti, M., Balestra, D., Rizzotto, L., Maestri, I., Pagani, F. and Bernardi, F. (2009) Rescue of coagulation factor VII function by the U1+5A snRNA. *Blood*, **113**, 6461–6464.
15. Pinotti, M., Rizzotto, L., Balestra, D., Lewandowska, M.A., Cavallari, N., Marchetti, G., Bernardi, F. and Pagani, F. (2008) U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency. *Blood*, **111**, 2681–2684.
16. Schmid, F., Glaus, E., Barthelmes, D., Fliegauf, M., Gaspar, H., Nurnberg, G., Nurnberg, P., Omran, H., Berger, W. and Neidhardt, J. (2011) U1 snRNA-mediated gene therapeutic correction of splice defects caused by an exceptionally mild BBS mutation. *Hum. Mutat.*, **7**, 815–824.
17. Tanner, G., Glaus, E., Barthelmes, D., Ader, M., Fleischhauer, J., Pagani, F., Berger, W. and Neidhardt, J. (2009) Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. *Hum. Mutat.*, **30**, 255–263.
18. Glaus, E., Schmid, F., Da Costa, R., Berger, W. and Neidhardt, J. (2011) Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells. *Mol. Ther.*, **19**, 936–941.
19. Susani, L., Pangrazio, A., Sobacchi, C., Taranta, A., Mortier, G., Savarirayan, R., Villa, A., Orchard, P., Vezzoni, P., Albertini, A. *et al.* (2004) TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and *in vitro* rescue by U1 snRNA. *Hum. Mutat.*, **24**, 225–235.

20. Raponi, M., Baralle, F.E. and Pagani, F. (2007) Reduced splicing efficiency induced by synonymous substitutions may generate a substrate for natural selection of new splicing isoforms: the case of CFTR exon 12. *Nucleic Acids Res.*, **35**, 606–613.
21. 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.
22. Cohen, J.B., Snow, J.E., Spencer, S.D. and Levinson, A.D. (1994) Suppression of mammalian 5′ splice-site defects by U1 small nuclear RNAs from a distance. *Proc. Natl Acad. Sci. USA*, **91**, 10470–10474.
23. Hwang, D.Y. and Cohen, J.B. (1997) U1 small nuclear RNA-promoted exon selection requires a minimal distance between the position of U1 binding and the 3′ splice site across the exon. *Mol. Cell Biol.*, **17**, 7099–7107.
24. 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. USA*, **96**, 6307–6311.
25. Cartegni, L. and Krainer, A.R. (2002) Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat. Genet.*, **30**, 377–384.
26. Lorson, C.L. and Androphy, E.J. (2000) An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. *Hum. Mol. Genet.*, **9**, 259–265.
27. 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.
28. Lorson, C.L., Rindt, H. and Shababi, M. (2010) Spinal muscular atrophy: mechanisms and therapeutic strategies. *Hum. Mol. Genet.*, **19**, R111–R118.
29. Meyer, K., Marquis, J., Trub, J., Nlend Nlend, R., Verp, S., Ruepp, M.D., Imboden, H., Barde, I., Trono, D. and Schumperli, D. (2009) Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation. *Hum. Mol. Genet.*, **18**, 546–555.
30. Gorman, L., Suter, D., Emerick, V., Schumperli, D. and Kole, R. (1998) Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. *Proc. Natl Acad. Sci. USA*, **95**, 4929–4934.
31. Madocsai, C., Lim, S.R., Geib, T., Lam, B.J. and Hertel, K.J. (2005) Correction of SMN2 Pre-mRNA splicing by antisense U7 small nuclear RNAs. *Mol. Ther.*, **12**, 1013–1022.
32. Liu, S., Asparuhova, M., Brondani, V., Ziekau, I., Klimkait, T. and Schumperli, D. (2004) Inhibition of HIV-1 multiplication by antisense U7 snRNAs and siRNAs targeting cyclophilin A. *Nucleic Acids Res.*, **32**, 3752–3759.
33. Asparuhova, M.B., Marti, G., Liu, S., Serhan, F., Trono, D. and Schumperli, D. (2007) Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing Tat and Rev exon skipping. *J. Gene Med.*, **9**, 323–334.
34. Francois, V., Klein, A.F., Beley, C., Jollet, A., Lemercier, C., Garcia, L. and Furling, D. (2011) Selective silencing of mutated mRNAs in DM1 by using modified hU7-snRNAs. *Nat. Struct. Mol. Biol.*, **18**, 85–87.
35. Uchikawa, H., Fujii, K., Kohno, Y., Katsumata, N., Nagao, K., Yamada, M. and Miyashita, T. (2007) U7 snRNA-mediated correction of aberrant splicing caused by activation of cryptic splice sites. *J. Hum. Genet.*, **52**, 891–897.
36. Beckley, S.A., Liu, P., Stover, M.L., Gunderson, S.I., Lichtler, A.C. and Rowe, D.W. (2001) Reduction of target gene expression by a modified U1 snRNA. *Mol. Cell Biol.*, **21**, 2815–2825.
37. Faustino, N.A. and Cooper, T.A. (2003) Pre-mRNA splicing and human disease. *Genes Dev.*, **17**, 419–437.
38. 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.
39. Hua, Y., Vickers, T.A., Baker, B.F., Bennett, C.F. and Krainer, A.R. (2007) Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon. *PLoS Biol.*, **5**, e73.
40. Ghigna, C., De Toledo, M., Bonomi, S., Valacca, C., Gallo, S., Apicella, M., Eperon, I., Tazi, J. and Biamonti, G. (2010) Pro-metastatic splicing of Ron proto-oncogene mRNA can be reversed: therapeutic potential of bifunctional oligonucleotides and indole derivatives. *RNA Biol.*, **7**, 495–503.
41. Nowak, D.G., Amin, E.M., Rennel, E.S., Hoareau-Aveilla, C., Gammons, M., Damodoran, G., Hagiwara, M., Harper, S.J., Woolard, J., Ladomery, M.R. *et al.* (2009) Regulation of vascular endothelial growth factor (VEGF) splicing from pro-angiogenic to anti-angiogenic isoforms: a novel therapeutic strategy for angiogenesis. *J. Biol. Chem.*, **285**, 5532–5540.
42. Merdzhanova, G., Gout, S., Keramidas, M., Edmond, V., Coll, J.L., Brambilla, C., Brambilla, E., Gazzeri, S. and Eymin, B. (2010) The transcription factor E2F1 and the SR protein SC35 control the ratio of pro-angiogenic versus antiangiogenic isoforms of vascular endothelial growth factor-A to inhibit neovascularization *in vivo*. *Oncogene*, **29**, 5392–5403.
43. Fong, L.G., Vickers, T.A., Farber, E.A., Choi, C., Yun, U.J., Hu, Y., Yang, S.H., Coffinier, C., Lee, R., Yin, L. *et al.* (2009) Activating the synthesis of progerin, the mutant prelamin A in Hutchinson-Gilford progeria syndrome, with antisense oligonucleotides. *Hum. Mol. Genet.*, **18**, 2462–2471.
44. Mandal, D., Feng, Z. and Stoltzfus, C.M. (2010) Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *J. Virol.*, **84**, 12790–12800.
45. Sawa, H. and Shimura, Y. (1992) Association of U6 snRNA with the 5′-splice site region of pre-mRNA in the spliceosome. *Genes Dev.*, **6**, 244–254.
46. Sontheimer, E.J. and Steitz, J.A. (1993) The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science*, **262**, 1989–1996.
47. Pagani, F., Buratti, E., Stuani, C., Bendix, R., Dork, T. and Baralle, F.E. (2002) A new type of mutation causes a splicing defect in ATM. *Nat. Genet.*, **30**, 426–429.
48. Bernardi, F., Dolce, A., Pinotti, M., Shapiro, A.D., Santagostino, E., Peyvandi, F., Batorova, A., Lapecorella, M., Schved, J.F., Ingerslev, J. *et al.* (2009) Major differences in bleeding symptoms between factor VII deficiency and hemophilia B. *J. Thromb. Haemost.*, **7**, 774–779.

Journal of Thrombosis and Haemostasis, **9**: 2143–2152 DOI: 10.1111/j.1538-7836.2011.04481.x

REVIEW ARTICLE

RNA-based therapeutic approaches for coagulation factor deficiencies

M. PINOTTI,* F. BERNARDI,* A. DAL MAS† and F. PAGANI†

**Department of Biochemistry and Molecular Biology, University of Ferrara, Ferrara; and †International Centre for Genetic Engineering and Biotechnology, Trieste, Italy*

To cite this article: Pinotti M, Bernardi F, Dal Mas A, Pagani F. RNA-based therapeutic approaches for coagulation factor deficiencies. *J Thromb Haemost* 2011; **9**: 2143–52.

Summary. Substitutive therapy has significantly ameliorated the quality of life of patients with coagulation factor deficiencies. However, there are some limitations that support research towards alternative therapeutic approaches. Here we focus on the rescue of coagulation factor biosynthesis by targeting the RNA processing and translation, which would permit restoration of the altered gene expression while maintaining the gene regulation in the physiological tissues. The essential prerequisite of the three reported RNA-based correction approaches (i–iii), which rely on mutation types and are applicable even to large size mRNAs, is the presence in cells of the precursor (pre-mRNA) or mature mRNA forms. (i) In the *F7* gene, modification of the small nuclear RNA U1 (U1 snRNA), the key component of the spliceosomal U1 ribonucleoprotein, re-directs correct usage of a mutated exon–intron junction, triggering synthesis of correct mRNA and secretion of functional factor (F)VII. (ii) Spliceosome-mediated RNA trans-splicing (SMaRT) between mutated and engineered pre-mRNAs produces normal FVIII mRNA and secretion of functional protein. (iii) Aminoglycoside drugs induce ribosome readthrough and suppress premature translation termination caused by nonsense mutations in FVII, VIII and IX. The rescued expression levels ranged from very low (aminoglycosides) to moderate (U1 snRNA and SMaRT), which could result in amelioration of the disease phenotypes. These findings prompt further studies aimed at demonstrating the clinical translatability of RNA-based strategies, which might open new avenues in the treatment of coagulation factor deficiencies.

Keywords: coagulation factor deficiencies, modified U1 snRNA, ribosome readthrough, RNA-based correction approaches.

Correspondence: Francesco Bernardi, Department of Biochemistry and Molecular Biology, University of Ferrara, Via Fossato di Mortara n 74, 44121 Ferrara, Italy.
Tel.: +39 532 974425; fax: +39 532 974484.
E-mail: ber@unife.it

Introduction

The inherited deficiency of procoagulant factors is associated with bleeding diathesis in patients, whose clinical manifestations are related to the clotting factor involved and the reduction extent of its plasma levels.

Deficiencies of factors (F)VIII (hemophilia A) or IX (Hemophilia B), which are caused by mutations in the X-linked *F8* and *F9* genes, respectively, represents 95%–97% of all inherited hemorrhagic coagulation factor disorders [1]. The other rare deficiencies are transmitted as autosomal recessive traits with a prevalence ranging from approximately 1 in 2 million for prothrombin and FXIII deficiency to 1 in 500 000 for FVII deficiency [2].

Current treatment for hemophiliacs is based on the intravenous administration of the missing proteins (replacement therapy), either plasma derived or produced by recombinant DNA technology [3], in response to bleeding episodes or prophylactically [4].

Although protein replacement has significantly increased the quality of life and prolonged the life expectancy of patients suffering from coagulation factor disorders, the cost and short half-life of these proteins impose limitations on this therapy that motivated research towards alternative therapeutic approaches.

As even a tiny increase in coagulation factor levels would result in a significant amelioration of the clinical phenotype, coagulation factor deficiencies represent preferred models to investigate innovative therapeutic approaches in a quantitative manner, by virtue of functional and protein assays in plasma.

Enormous efforts have made with regards to substitutive gene therapy that consists of viral or non-viral mediated delivery of a copy of the defective gene (or better of the coding DNA sequence) into the patient's cells, thus triggering stable endogenous expression of the missing protein [5,6].

Another area of research has been focused on the correction of the gene expression of the mutated clotting factor gene by modulation of the messenger RNA (mRNA) processing and translation, which has been successfully explored for the treatment of other human genetic disorders [7–9]. Notably, RNA targeting would permit restoration of gene expression

while maintaining the gene promoter regulation in the cells belonging to the physiological tissue. Moreover, it has the potential to circumvent some limitation owing to the large size of certain human disease genes, and could be also effective in addressing dominant-negative disease forms.

An essential prerequisite for an RNA-based therapy is the presence of the transcribed RNA either in its precursor (pre-mRNA) or mature form (mRNA), the targets of the therapeutic compounds. Depending on the mutation type, three approaches can be foreseen (Table 1). Whereas *trans*-splicing can theoretically act on several types of mutations, the strategy aimed at inducing ribosome readthrough over stop codon or at restoring splicing represent interventions for disease forms caused by nonsense substitutions or mutations affecting pre-mRNA processing, respectively.

The different strategies for splicing correction include the usage of small nuclear RNAs and antisense oligonucleotides. In general, antisense oligonucleotides bind to target sequences on pre-mRNA and can either directly mask splicing regulatory elements or recruit additional binding sites for positive splicing factors. The usage of antisense oligonucleotides for splicing correction has been illustrated in previous excellent reviews [10–13] and will not be described here.

Two forms of spliceosomal small nuclear RNAs (U1 and U7 snRNA) have been used so far to modulate splicing. These molecules were exploited in human disease models other than coagulation factor deficiencies as antisense molecules to inhibit splicing, for example by masking the disease-causing mutation and inducing exon skipping and/or by interfering with the spliceosome formation [11,13].

On the other hand, U1 snRNA is becoming an attractive molecule for its specific capacity to correct splicing mutations at the donor splice site [14–21].

In the present review, we will focus on the principle of the different approaches, their application and future perspective in the treatment of inherited coagulation factor disorders. Particular attention will be given to novel approaches with modified U1 snRNA to restore normal splicing in donor splice site defects, and with drug-mediated ribosome readthrough over nonsense mutations.

Table 1 Summary of correction strategies targeting pre-messenger RNA(mRNA) processing and mRNA translation

RNA-based therapeutic strategies in coagulation factor deficiencies		
Strategy	Target molecule	Mutation type
Splicing correction (U1 snRNAs. oligos, drugs)	pre-mRNA	Splicing defects
Trans-splicing	pre-mRNA	Mutations not impairing pre-mRNA levels
Ribosome readthrough (drugs)	mRNA	Nonsense mutations*

*Nonsense mutations associated with major nonsense-mediated mRNA degradation may not respond to treatment.

Rescue of coagulation factor expression by splicing modulation

The splicing process and the splicing pathology

The splicing process Exonic sequences located on primary transcripts (pre-mRNAs) must be precisely identified and joined together to form the mature mRNA along with the removal of the usually larger intronic sequences. In the past, this process has revealed a large complexity, rendering it particularly susceptible to derangements caused by human mutations [22–24]. Pre-mRNA splicing requires the correct identification of several *cis*-acting elements in an ordered fashion [25,26]. The core elements are moderately conserved and consist of the classic or canonical splice sites (the 5′ and 3′ splice sites named also donor and acceptor sites), the polypyrimidine tract and the branch site (Fig. 1). However, the nascent pre-mRNA transcript contains several potential cryptic donor (5′ss) and acceptor (3′ss) splice sites, which are never used. To correctly identify the correct splice site, the splicing machinery requires the assistance of additional splicing regulatory elements. According to their position and functional effect, these elements are classified as exonic or intronic, and enhancer or silencer, depending on their effect on splicing [25,26].

Several previous studies have demonstrated that these regulatory elements play important roles in modulating RNA splicing in normal and pathological conditions. A classic example is represented by purine-rich Exonic Splicing Enhancers (ESE). A mutation in this element may induce, independently of its effect on the amino acid code, a pathological skipping of the exon from the mRNA. In general, the splice sites and the additional splicing regulatory elements are recognized by several *trans*-acting factors that form a network of interactions on the exon and contribute to its definition. As a result, the final outcome of a splicing decision

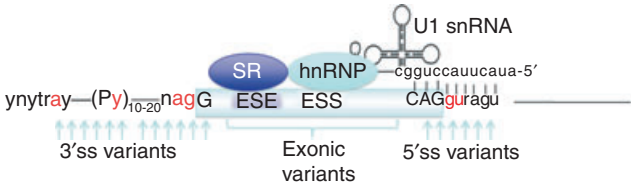


Fig. 1. Regulatory elements in pre-messenger RNA (mRNA) splicing. Schematic representation of the major splicing regulatory elements on pre-mRNA along with corresponding *trans*-acting factors and the genomic variants that affect splicing. The canonical splicing signals (in red) are composed of the branch site, the polypyrimidine tract (Py), the acceptor (3′ss) and the donor (5′ss) splice sites. In the first step of splicing, the 5′ tail of U1 snRNA recognizes the consensus 5′ss mainly by base pair complementarity. The boxed exon contains regulatory sequences, Exonic Splicing Enhancer (ESE) and Silencer (ESS), which contribute positively or negatively to exon recognition through interaction with SR proteins and hnRNPs (heterogeneous RiboNuclearParticles), respectively. Genomic variants (arrows) at canonical splicing signals, in the exonic or intronic (not shown) splicing regulatory sequences can affect splicing.

depends on the equilibrium between multiple positive and negative interactions.

The splicing pathology As a result of the complexity of the process, splicing defects can originate from mutations at the splice sites or at the splicing regulatory elements located either in the exon or in the intron (Fig. 1). A thorough investigation of the composition of splicing regulatory elements is not only important to improve our knowledge of the basic splicing mechanisms, but it provides potential targets for splicing correction. A paradigmatic example is offered by Spinal Muscular Athrophy (SMA), in which an intronic splicing silencer (ISS) in the *SMN2* gene can be specifically targeted with an oligonucleotide [11,27]. This results in improved inclusion of SMN2 exon 7, which compensates for the lack of the paralogue *SMN1* gene, and ameliorated phenotype in a mouse model of SMA [28].

Role of U1 snRNP in splicing regulation A key molecule involved in the initial discrimination between exons and introns is the U1 small nuclear RiboNuclearParticle (U1 snRNP) [29]. In common with other snRNPs, it contains a small RNA (165bp for U1 snRNA) complexed with several proteins. The 5′ end of the U1 snRNA interacts by complementarity with the donor splice site, whose sequence is moderately conserved. In fact, its consensus motif consists of a nine-nucleotide sequence, CAG/GURAGU where R is a purine. However, with the exclusion of the nearly obligate GU dinucleotide, the other positions can tolerate some substitutions. About 40%, 22% and 5% of normal donor splice sites contain, respectively, two, three or four mismatches towards the U1 snRNA [30]. As a consequence, in some cases the effect of mutations flanking the canonical GU site may be difficult to predict without an analysis of the mature mRNA produced. Furthermore, additional splicing regulatory sequences are involved in the recognition, and can compensate for the defective complementarity between the U1 snRNA and the donor splice site.

Splicing correction using U1 snRNA

Splicing defects caused by mutations at the donor splice site consensus can be corrected in target cells with a molecular approach that re-directs the endogenous spliceosome machinery to the proper splicing junction of the mutant mRNA, thus rescuing exon definition and gene expression.

Donor splice site substitutions represent approximately 8% of all different mutations found to be associated with human inherited diseases [23,31–33]. In coagulation factor genes, mutations affecting the donor splice sequences (GU dinucleotide or the nearby positions) occur at a similar rate (7% in *F7*; 8% in *F9*; 3% in *F8*) (http://www.hgmd.cf.ac.uk/ac/index.php). These mutations, reducing the complementarity of the U1 snRNA to the donor splice site, can result in exon skipping, intron retention or activation of cryptic splice sites. *In silico* analysis through a number of computer-assisted tools available on the web helps the prediction of their disease-causing effect. However, for many mutations, in particular those nearby the invariant GU dinucleotide, it can be difficult to clearly predict their effect on splicing based only on the DNA sequence.

For example, the intron 7 donor splice site of *F7* gene, where a cluster of mutations has been mapped [34], displays a very low score, and significantly deviates from the consensus sequence (Fig. 2).

As donor splice site mutations disrupt the complementarity of the donor site with the endogenous U1 snRNA, restoring the complementarity through engineered modification of the U1 snRNA represents a valuable approach. For this purpose, the normal 5′ end sequence of the U1 snRNA is substituted with a

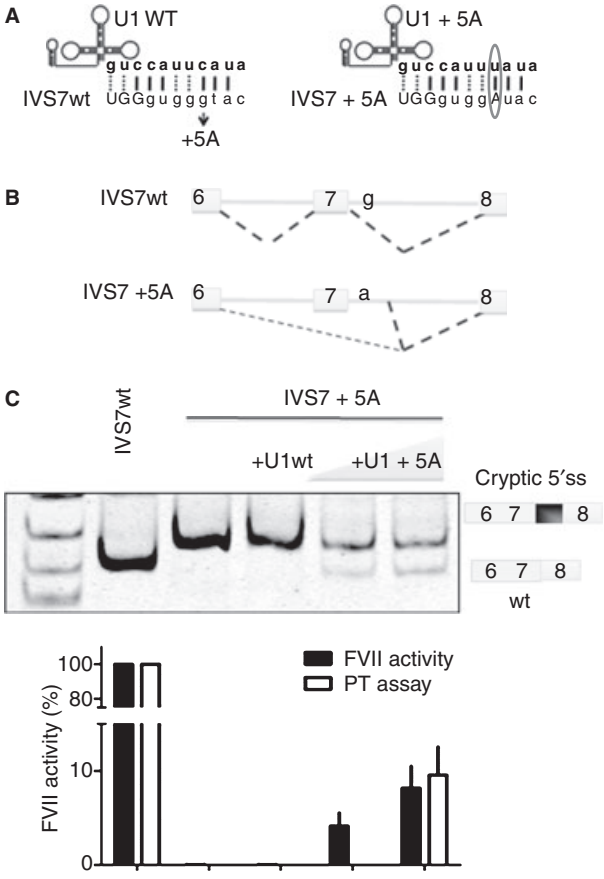


Fig. 2. U1 snRNA-mediated rescue of coagulation factor (F)VII expression impaired by the IVS7 + 5G/A mutation. (A) Schematic representation of the U1 snRNA with its 5′ tail paired to the 5′ donor splice site of intron 7 (IVS7) of the *F7* gene (exonic and intronic sequences are shown in upper and lower case letters, respectively). The normal (left) and mutated (right) IVS7 sequences targeted by the wild-type and modified U1 snRNA (U1 + 5A) are reported. The circle indicates the rescued complementarity at the mutated nucleotide position. (B) Schematic representation of the splicing events (dotted lines) occurring in normal (upper panel) and mutated (lower panel) conditions. (C) Rescue of FVII mRNA processing (upper panel) and of FVII function (lower panel) by the modified U1 + 5A in co-transfection experiments. The figure has been adapted from data previously reported by Pinotti *et al.* [14,15].

sequence complementary to the target mutation (Fig. 2A), and the modified U1 snRNA gene is tested in appropriate splicing assays.

This approach has been originally employed to demonstrate the role of U1 snRNA in donor splice site recognition in model systems, and recently studied for the rescue of splicing defects in coagulation disorders [14,15].

The 9726 + 5G/A substitution in the *F7* intron 7 (IVS7 + 5G/A) reduces the complementarity with the normal U1 snRNA (Fig. 2A), activates an intronic downstream cryptic donor splice site and induces exon skipping (Fig. 2B). The resulting transcripts do not encode for functional proteins.

Co-expression of a modified U1 snRNA with increased complementarity to the defective donor splice site, named U1 + 5A (Fig. 2A), induced a significant rescue of the splicing defect with synthesis of the normal transcript (Fig. 2C). Interestingly, the splicing correction resulted in the synthesis of functional FVII molecules, whose coagulant activity levels reached 9% of those of the normal construct (Fig. 2C), an extent that would be theoretically sufficient to correct the coagulation defect *in vivo*.

The modified U1 snRNA approach has also been used for some other diseases caused by donor splice site mutations either to correct splicing defects or to detail mechanisms of aberrant splicing [16–21]. It is worth noting that mutations at the invariant GU site are generally not rescued, consistent with the critical importance of this dinucleotide in splicing. Differently, splicing rescue is possible for several substitutions nearby the GU, which account for approximately 40% of all donor splice site mutations. Among them, there are position and sequence preferences influenced by the gene context that could render some variants more susceptible to the rescue effect of U1 snRNAs.

The use of modified U1 snRNA represents a complementary strategy to classic gene therapy with which it shares the delivery issue as well as the potential capability of guaranteeing long-term correction of the genetic defect.

For those genes, like *F8*, in which the large size of the coding region limits their package into some viral vectors such as those derived from the Adeno-associated virus (AAV) [35], modified U1 snRNA might represent a valid option.

The U1 snRNA gene used for splicing rescue includes promoter and regulatory sequences in < 1 Kb and can be easily inserted, even in multiple copies, in AAV vectors, which are now actively studied for gene therapy of coagulation factor disorders [35]. In addition, the U1 snRNA approach, acting on the pre-mRNA, has the advantage of maintaining the expression regulation of the targeted gene in the normal chromosomal context. Future studies in appropriate animal models will be necessary to prove the efficacy *in vivo*.

In common with other rescue strategies based on targeting RNA by complementarity (i.e. oligonucleotides), modified U1 snRNAs have to deal with potential off-target effects that might affect splicing of other genes. This could be dangerous for modified U1 snRNAs that have only one base change from the natural U1 snRNA, and thus might activate normally silent cryptic donor splice sites and induce aberrant splicing in other genes. However, the design of novel Exon-Specific U1 snRNAs (ExSpeU1) complementary to non-conserved sequences downstream of mutant donor splice sites are expected to reduce this type of interaction (F. Pagani and M. Pinotti, unpublished data). ExSpeU1, reminiscent of originally described shifted U1 snRNAs [36], would reduce off-targets and, by binding downstream of the donor splice site, correct different splicing mutations. In this case, a single ExSpeU1 rescues multiple splicing defects that affect a single exon.

The potential therapeutic effect of modified U1 snRNAs on pre-mRNA relies also on the gene context and on the targeting sequence. For example, in Duchenne Muscular Dystrophy (DMD) modified U1 snRNAs that bind extensively to exonic sequences and splice sites induce skipping of an exon that contains a disease-causing stop codon [37]. The resulting mature mRNA devoid of the defective exon is in frame and codes for a functionally active protein. On the other hand, binding of modified U1 snRNA to 3′UTR can be used to silence genes by interfering with polyadenylation and stability of the mRNA [38].

In factor (F)IX (FIX), the importance of the context where the U1 snRNAs binds is highlighted by the analysis of a A to T mutation occurring at position + 13 downstream exon 5 of *F9* gene. This mutation is associated with mild hemophilia B and activates a cryptic donor splice site (Fig. 3). We tested whether an ExSpeU1, designed to bind to the not conserved intronic sequence of the cryptic donor splice site (Fig. 3A), improves selection of the correct one. Expression of this ExSpeU1 + 13 partially rescued the splicing pattern and led to an increase in levels of correctly spliced mRNA from approximately 8% to 40% (Fig. 3B–C). Interestingly, this ExSpeU1 did not stimulate usage of the cryptic donor splice site where it binds, suggesting that it acts on intronic splicing regulatory elements important for donor splice site definition. Thus, fine-tuning of U1 snRNA binding sites might reveal interesting targets for splicing correction.

It should be noticed that an important requirement for developing novel strategies for splicing correction relies also on our knowledge of basic mechanisms of normal and pathological splicing, and very few data are available regarding the architecture of defective splicing in coagulation factor genes. In these genes, the identification of splicing regulatory elements with enhancer and silencers function, and their role in pathological processing of pre mRNA, could reveal novel pathways for therapeutic intervention.

Rescue by spliceosome-mediated RNA trans-splicing (SMaRT)

Normally, splicing occurs in *cis* within the same pre-mRNA molecule. Although to a very low frequency in mammalian cells, the spliceosome machinery can also catalyze splicing between two separate pre-mRNA molecules, leading to a process called RNA trans-splicing [39].

The induced spliceosome-mediated RNA trans-splicing (SMaRT) (extensively reviewed by Yang *et al.* [40]) is a recent

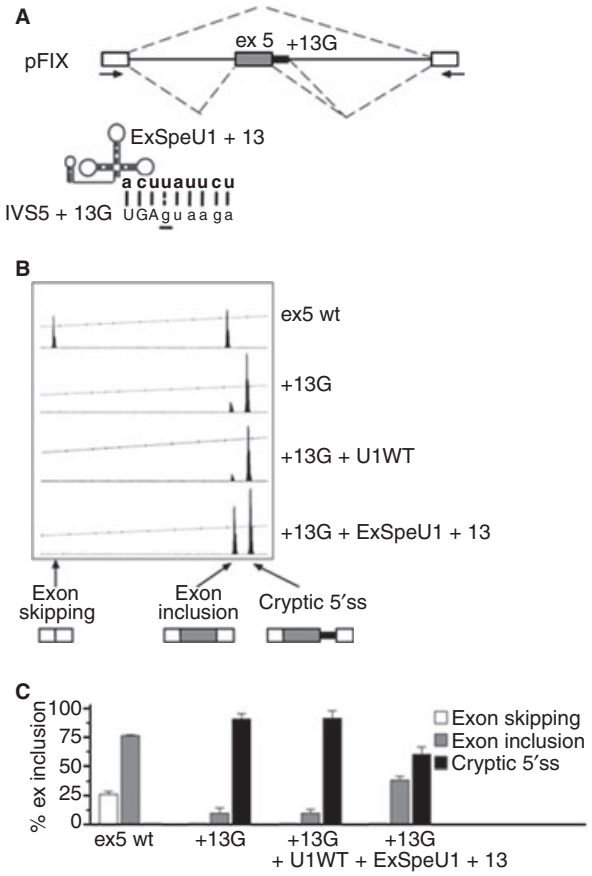


Fig. 3. Partial rescue of factor (F)IX) splicing by the ExSpeU1 + 13. (A) Schematic representation of the minigene construct (pFIX) (upper panel). Exonic and intronic sequences are boxes and lines, respectively. The black box represents the sequences inserted by the +13G mutation in the final mRNA and the dotted lines above and below the construct indicate the splicing events in normal and mutated conditions, respectively. The arrows are the primers used in RT-PCR amplification. The lower panel shows the RNA complementarity between the 5′ tails of the ExSpeU1 + 13 and the intronic FIX sequence affected by the +13a/g mutation (underlined). (B) Separation on a denaturing capillary system of fluorescently labeled RT-PCR products obtained from total RNA of HeLa cells expressing the normal (ex5 wt) or the mutated (+13G) minigenes without or upon overexpression of the normal (U1wt) or the modified (ExSpeU1 + 13G) U1 snRNAs. The identity of transcripts deriving from exon 5 skipping, usage of the correct (exon 5 inclusion) or the cryptic (cryptic 5′ss) donor splice sites is depicted at the bottom. (C) Relative percentage of transcripts in the different experimental conditions.

RNA reprogramming strategy in which the final mRNA results from splicing of two independently transcribed RNA molecules. During splicing of the endogenously expressed gene, a new pre-mRNA sequence, driven by an exogenous vector, is trans-spliced into the target to generate a new gene product. Through this approach, depending on the design of the pre-trans-splicing RNA molecules (PTMs), the target mRNA can be modified to include the novel sequences at its 3′ or 5′ region, or within its sequence (Fig. 4).

PTMs consist of three main elements: a binding domain (BD) complementary to a chosen target pre-mRNA, an unpaired splice site and a coding domain containing sequences to be trans-spliced into the target. To prevent direct PTM expression, 3′ exon replacement PTMs should be constructed without translation start codons, 5′ PTMs without polyadenylation signals and double trans-splicing PTMs should lack both these sequence elements.

Hemophilia A was the first human inherited disease in which the therapeutic potential of SMaRT has been successfully demonstrated *in vivo* [41]. Hemophilia A mice created by insertion of the neomycin resistance gene into *F8* exon 16 did not display detectable levels of functional FVIII. It should be noted that replacement gene therapy approaches for this bleeding disorder are made difficult by the large size of the *F8* gene, thus making alternative approaches of great interest.

To induce synthesis of the correct FVIII mRNA, a PTM was designed to trans-splice to exon 15 of the endogenous FVIII pre-mRNA target and to replace the downstream coding sequence. By exploiting adenoviral vector driving the expression of the PTM, the authors demonstrated the presence of the repaired FVIII mRNA in the mouse liver, a significant increase (25% of normal) in circulating functional FVIII and an amelioration of the bleeding phenotype.

Based on these results and on the principle of the SMaRT, virtually all mutations in coagulation factor deficiencies could be approached, with the exception of those significantly affecting gene transcription (i.e. sequence variations in the promoter), which prevent pre-mRNA from being available to trans-splicing.

Although extremely interesting, a number of open issues have to be addressed before translating SMaRT into the clinic. Among them, improvement in SMaRT efficiency for therapeutic purposes will probably require the development of more effective molecules and/or the combination with other approaches that, by specifically targeting the spliceosome, will facilitate the process. For example, an antisense oligonucleotide has been used to selectively block a downstream splice site and facilitate *trans*-splicing [42] in a mouse model of SMA. This approach could be exploited in inherited coagulation disorders, to improve its efficacy in a transcript specific manner.

The potential and limitations of the SMaRT approach have been extensively reviewed elsewhere [40,43] and will not be discussed further in the present review.

As for any gene therapy approach, the induction of trans-splicing or splicing correction by modified U1 snRNA *in vivo* requires a safe and efficient delivery of the expression cassette that should guarantee a long-term correction of the genetic defect. As a result of their limited size, the expression cassettes could fit into adeno-associated virus (AAV) vectors that are extensively used for gene replacement therapy purposes [35]. These studies, and particularly those on gene therapy of hemophilia B in the liver, will hopefully provide the ideal vector and protocol to translate the RNA-based mediated approaches into the clinic.

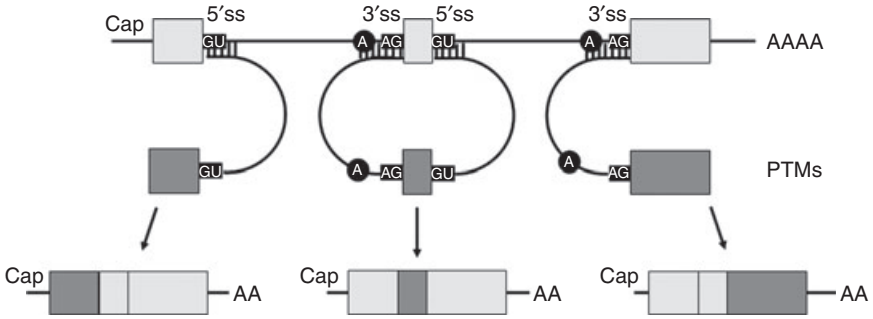


Fig. 4. Spliceosome-mediated RNA *trans*-splicing (SmaRT). The therapeutic trans-splicing is based on the substitution of the 5′ (left), internal (middle) or 3′ (right) exons (light grey rectangles). The three different trans-splicing events are triggered by pre-trans-splicing molecules (PTMs), appropriately designed to reprogram 3′, internal or 5′ exons (dark grey rectangles). Products of SmaRT, in which the new exons are inserted, are shown in the lower panels.

Rescue of coagulation factor expression by translation modulation

Nonsense mutations, translation termination efficiency and ribosome readthrough

Mutations introducing premature stop codons (nonsense mutations) are relatively frequent in coagulation factor deficiencies, particularly in the most severe forms. Nonsense mutations account for approximately 10% of all molecular defects leading to hemophilia A (http://europium.csc.mrc.ac.uk) and B (http://www.kcl.ac.uk/ip/ petergreen/haemBdatabase.html) [44,45].

Depending on its localization in the gene sequence, premature stop codons can trigger the rapid degradation of the mRNA through a process termed nonsense-mediated mRNA decay (NMD) [46]. Moreover, the translation of the mutated mRNA may lead to synthesis of truncated polypeptides that usually are not properly folded and hence degraded. Based on these detrimental mechanisms, nonsense mutations virtually produce null phenotypes.

A multitude of mechanisms and players participate in the exquisitely regulated process of translation termination at nonsense codons (UAA, UAG and UGA) that is essential for the correct expression of proteins. When a stop codon is presented in the A site of the ribosome, it is recognized and bound by release factors that trigger release of the polypeptide chain [47]. However, albeit with a very low rate in normal conditions (10^{-4}), an aminoacyl-tRNA can also enter the ribosomal A site at the stop codon position, thus leading to amino acid misincorporation and keep the the protein synthesis going. This process is termed ribosome readthrough [48]. As the sequence context of the natural stop codon has a role in its functional definition, the premature nonsense triplets could lack the proper definition, which would favour ribosome readthrough [49].

Translation fidelity has a crucial role in the gene expression control and is the target of antibiotic drugs such as aminoglycosides. These drugs are able to specifically bind to the highly conserved 16S ribosomal RNA (rRNA) at the decoding center of the small bacterial ribosomal subunit [50]. The aminoglycosides induce codon misreading and misincorporation of amino acids at a nonsense codon, thus leading to translational readthrough rather than chain termination.

Although with a much lower affinity, aminoglycosides can also bind to the eukaryotic ribosome, thus explaining their toxicity in humans upon long-term administration or at high doses [51].

In the past decades, many studies have focused on turning this side effect into a therapeutic opportunity for human genetic diseases caused by nonsense mutations [52]. The aminoglycoside-mediated restoration of protein biosynthesis in the presence of a nonsense mutation may, even to a lower extent, be functionally significant (Fig. 5).

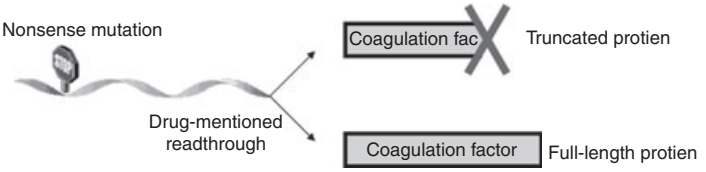
To date, the restoration of protein and functional levels by aminoglycosides has been demonstrated *in vitro* and *in vivo* for a number of human diseases such as cystic fibrosis (CF) and DMD (reviewed in Zingman *et al.* [53]). Here we will focus on studies conducted in hemophilia and FVII deficiency.

Aminoglycoside-mediated rescue

James *et al.* [54] tested the effect of the aminoglycoside gentamicin (7 mg kg^{-1} once a day) in severe hemophilia A and B patients with nonsense mutations (Fig. 5). In one patient with hemophilia A (S1395X mutation) and one with hemophilia B (R333X mutation) the treatment produced a transient shortening in the activated partial thromboplastin time (approximately 20 s) and an increase in FVIII/FIX activity (approximately 2%). However, such an effect on functional parameters could not be detected in the other three patients.

Notably, in the hemophilia A patient displaying a 1.6% increase in FVIII activity the plasma FVIII antigen levels were remarkably higher (7%), thus suggesting that ribosome readthrough induced the synthesis of dysfunctional molecules. This hypothesis was corroborated by the observation of a persistent increase in FIX antigen levels (approximately 2%) in one hemophilia B patient (R252X mutation) in whom FIX activity levels did not appreciably change upon treatment.

A



B

	Mutation	Cellular model		Mouse model		Patients	
		Drug	Response	Drug	Response	Drug	Response
FVII deficiency	K316X	G418	Yes (>3%)[57]				
	K316X	Gentamicin	Yes (>2%)[57]			Gentamicin	Sub-therapeutic[58]
	VV364X	G418	Yes (>3%)[57]				
	VV364X	Gentamicin	Yes (>2%)[57]			Gentamicin	Sub-therapeutic[58]
FIX deficiency	R29X			G418	No[56]		
	R29X			Gentamicin	No[56]		
	R252X					Gentamicin	No[54]
	R333X					Gentamicin	Yes (2%)[54]/NO[55]
	R338X			G418	Yes, (>8%)[56]		
	R338X			Gentamicin	Yes, (>3%)[56]		
FVIII deficiency	R427X					Gentamicin	No[54]
	S1395X					Gentamicin	Yes (1.6%)[54]
	R2116X					Gentamicin	No[54]

Fig. 5. Aminoglycoside-mediated ribosome readthrough in coagulation factor deficiencies. (A) Schematic representation of the effect of premature nonsense codons and of ribosome readthrough on protein products. (B) Summary of results obtained with the aminoglycosides neomycin (G418) and gentamicin in factor (F)VII, FIX or FVIII deficiencies caused by nonsense mutations. The percent within parentheses indicate the activity levels measured upon treatment in the experimental models or patients.

A negligible effect on FIX activity levels was also previously reported by Srivastava *et al.* [55] in four hemophilic B patients treated with gentamicin.

A remarkable variation of FIX levels in response to aminoglycosides geneticin and gentamicin has been documented in hemophilic B mice expressing the human FIX mutations R338X and R29X [56]. Geneticin treatment resulted in a significant increase (approximately 5%) in antigen and activity in R338X mice but not in R29X mice. Noticeably, residual antigen could be detected at 3 weeks, and activity levels could be detected at 6 days.

The aminoglycosides-mediated rescue of FVII expression impaired by the K316X and W364X nonsense mutations was explored by us both in cellular models [57] and in patients [58].

In a fluorescent model with the FVII-green fluorescent protein (GFP) chimera, geneticin treatment of cells expressing the nonsense variants resulted in an appreciable fluorescence expression, to indicate partial restoration of full-length protein synthesis. Consistently, expression studies with the native FVII revealed that geneticin (G418) and gentamicin induced a dose-dependent increase of secreted FVII molecules with activity levels reaching 3% to 4% of wild-type FVII (Fig. 5). However, the increase in secreted protein levels was more pronounced, thus indicating the synthesis of dysfunctional proteins [57].

The pilot clinical study with gentamicin (3 mg kg^{-1} once a day) was conducted in two patients bearing the K316X and W364X mutations [58]. At some time-points prothrombin time (PT) values were slightly shortened (approximately 10 s) and, using activated FX generation assays, the FVII activity (0.1%–0.25% of pooled normal plasma [PNP]) was significantly above the baseline (Fig. 5). However, the FVII antigen and coagulant activity levels were undetectable.

Altogether *in vitro* and *in vivo* data, both in humans and animal models, revealed a low and extremely variable response to aminoglycosides, thus making it difficult to define their therapeutic potential. Several elements might explain the variable response:

1 the levels of the mutated mRNA, differentially affected by nonsense mediated decay;
2 the presence of nonsense mutations that may undergo readthrough with differential efficiency, depending on the nonsense triplet and the sequence context [59–62];
3 the introduction of amino acids other than the natural one at the nonsense positions, which could impair protein biosynthesis, stability and/or function. This would explain discrepancies between antigen and activity levels detected upon treatment in some hemophilic patients [54] and in cellular models of FVII deficiency [57]. Therefore, the beneficial effects could be low or negligible for nonsense mutations occurring at crucial functional positions of the protein; and
4 the different duration and method of application of the aminoglycosides between studies, and variability in drug metabolism among individuals.

The toxicity issue should be also considered. In fact, the clinical benefit of gentamicin, and of aminoglycosides in general,

is limited because high concentrations and/or long-term treatments can cause severe side effects such as kidney damage and hearing loss [63]. Novel molecules, able to discriminate between premature and normal termination codons, are needed to improve selectivity and efficacy, and thus safety.

A high-throughput screening revealed a new small molecule, PTC124 (Ataluren, PTC therapeutics, South Plainfield, NJ, USA), which can readthrough premature termination codons [64,65] and thus potentially treat a variety of genetic diseases. The molecule demonstrated its correction efficacy in a number of human disease models [66] and, in phase I safety studies, was found to be well tolerated [65].

Phase II clinical trials of PTC124 have been initiated in patients with different human genetic diseases caused by nonsense mutations, including CF, DMD or Becker Muscular Dystrophy (BMD) and hemophilia.

In a study involving 23 CF individuals, more than half of the patients showed convincing changes in the nasal chloride-channel defect in the first cycle of treatment, but this improvement was observed in only one-third of them in the second cycle [67]. In a second trial, PTC124 treatment significantly increased the proportion of nasal epithelial cells expressing apical full-length cystic fibrosis transmembrane conductance receptor (CFTR) and induced a nasal chloride transport response or hyperpolarization in 50% and 47% of patients [68]. However, in a phase II clinical trial in DMD/BMD patients, a 48-week treatment of PTC124 did not result in a significant change in 6-min walking distance [69]. The ongoing clinical trial in hemophilia A and B patients will provide insights into this promising 'mutation specific' therapeutic approach for coagulation factor deficiencies.

General conclusion

The intervention at the post-transcriptional (modulation of cis-splicing or induction of trans-splicing) and translational (induction of ribosome readthrough) levels represents an innovative therapeutic approach for coagulation factor diseases. Although the modulation of cis-splicing and of translation termination is mutation specific, it must be underlined that splicing and nonsense mutations are relatively frequent in the severe coagulation factor deficiencies.

One might argue that the extent of rescue of expression levels so far obtained by strategies at the RNA level is moderate and, for *in vivo* application of modified U1 snRNA and SMaRT, depends on a safe and efficient delivery of expression cassettes. However, even a very low increase in plasma levels of coagulation factors would result in a significant amelioration of a clinical phenotype of patients, thus encouraging further studies aimed at demonstrating their clinical translatability.

Acknowledgements

The financial support of Telethon – Italy (GGP09183) (M.P., F.P), Ministero dell'Università e della Ricerca (MIUR)-Progetti di Ricerca di Interesse Nazionale (PRIN)(M.P.), the Italian Cystic Fibrosis Foundation n. FFC#9/2009 adopted by Dedikato, CIV and Delegazione FFC del Lago di Garda (F.P., A.D.M.), University of Ferrara (M.P., F.B.) and Fondazione CARIFE (M.P.) is gratefully acknowledged.

Disclosure of Conflict of Interests

The authors state that they have no conflict of interest.

References

1 Bolton-Maggs PH, Pasi KJ. Haemophilias A and B. *Lancet* 2003; **361**: 1801–9.
2 Mannucci PM, Duga S, Peyvandi F. Recessively inherited coagulation disorders. *Blood* 2004; **104**: 1243–52.
3 Pipe SW, High KA, Ohashi K, Ural AU, Lillicrap D. Progress in the molecular biology of inherited bleeding disorders. *Haemophilia* 2008; **14**: 130–7.
4 Carcao MD, Aledort L. Prophylactic factor replacement in hemophilia. *Blood Rev* 2004; **18**: 101–13.
5 Murphy SL, High KA. Gene therapy for haemophilia. *Br J Haematol* 2008; **140**: 479–87.
6 Petrus I, Chuah M, Vandendriessche T. Gene therapy strategies for hemophilia: benefits versus risks. *J Gene Med* 2010; **12**: 797–809.
7 Wilton SD, Fletcher S. RNA splicing manipulation: strategies to modify gene expression for a variety of therapeutic outcomes. *Curr Gene Ther* 2005; **5**: 467–83.
8 Wood M, Yin H, McClorey G. Modulating the expression of disease genes with RNA-based therapy. *PLoS Genet* 2007; **3**: e109.
9 Bonetta L. RNA-based therapeutics: ready for delivery? *Cell* 2009; **136**: 581–4.
10 Du L, Gatti RA. Progress toward therapy with antisense-mediated splicing modulation. *Curr Opin Mol Ther* 2009; **11**: 116–23.
11 Nlend Nlend R, Meyer K, Schumperli D. Repair of pre-mRNA splicing: prospects for a therapy for spinal muscular atrophy. *RNA Biol* 2010; **7**: 430–40.
12 Ward AJ, Cooper TA. The pathobiology of splicing. *J Pathol* 2010; **220**: 152–63.
13 Hammond SM, Wood MJ. Genetic therapies for RNA mis-splicing diseases. *Trends Genet* 2011; **27**: 196–205.
14 Pinotti M, Rizzotto L, Balestra D, Lewandowska MA, Cavallari N, Marchetti G, Bernardi F, Pagani F. U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency. *Blood* 2008; **111**: 2681–4.
15 Pinotti M, Balestra D, Rizzotto L, Maestri I, Pagani F, Bernardi F. Rescue of coagulation factor VII function by the U1 +5A snRNA. *Blood* 2009; **113**: 6461–4.
16 Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, French-Constant C, Baralle FE. Identification of a mutation that perturbs NF1 agene splicing using genomic DNA samples and a minigene assay. *J Med Genet* 2003; **40**: 220–2.
17 Susani L, Pangrazio A, Sobacchi C, Taranta A, Mortier G, Savarirayan R, Villa A, Orchard P, Vezzoni P, Albertini A, Frattini A, Pagani F. TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA. *Hum Mutat* 2004; **24**: 225–35.
18 Tanner G, Glaus E, Barthelmes D, Ader M, Fleischhauer J, Pagani F, Berger W, Neidhardt J. Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. *Hum Mutat* 2009; **30**: 255–63.
19 Hartmann L, Neveling K, Borkens S, Schneider H, Freund M, Grassman E, Theiss S, Wawer A, Burdach S, Auerbach AD, Schindler D, Hanenberg H, Schaal H. Correct mRNA processing at a mutant

TT splice donor in FANCC ameliorates the clinical phenotype in patients and is enhanced by delivery of suppressor U1 snRNAs. *Am J Hum Genet* 2010; **87**: 480–93.
20 Schmid F, Glaus E, Barthelmes D, Fliegauf M, Gaspar H, Nürnberg G, Nürnberg P, Omran H, Berger W, Neidhardt J. U1 snRNA-mediated gene therapeutic correction of splice defects caused by an exceptionally mild BBS mutation. *Hum Mutat* 2011; **32**: 815–24.
21 Glaus E, Schmid F, Da Costa R, Berger W, Neidhardt J. Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells. *Mol Ther* 2011; **19**: 936–41.
22 Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 2002; **3**: 285–98.
23 Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. *Genes Dev* 2003; **17**: 419–37.
24 Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet* 2004; **5**: 389–96.
25 Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* 2009; **10**: 741–54.
26 Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 2010; **463**: 457–63.
27 Sendtner M. Therapy development in spinal muscular atrophy. *Nat Neurosci* 2010; **13**: 795–9.
28 Passini MA, Bu J, Richards AM, Kinnecom C, Sardi SP, Stanek LM, Hua Y, Rigo F, Matson J, Hung G, Kaye EM, Shihabuddin LS, Krainer AR, Bennett CF, Cheng SH. Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Sci Transl Med* 2011; **3**: 72ra18.
29 Horowitz DS, Krainer AR. Mechanisms for selecting 5′ splice sites in mammalian pre-mRNA splicing. *Trends Genet* 1994; **10**: 100–6.
30 Carmel I, Tal S, Vig I, Ast G. Comparative analysis detects dependencies among the 5′ splice-site positions. *RNA* 2004; **10**: 828–40.
31 Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base–pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum Genet* 1992; **90**: 41–54.
32 Krawczak M, Thomas NS, Hundrieser B, Mort M, Wittig M, Hampe J, Cooper DN. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Hum Mutat* 2007; **28**: 150–8.
33 Buratti E, Chivers M, Královicová J, Romano M, Baralle M, Krainer AR, Vorechovsky I. Aberrant 5′ splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 2007; **35**: 4250–63.
34 Pinotti M, Toso R, Redaelli R, Berrettini M, Marchetti G, Bernardi F. Molecular mechanisms of FVII deficiency: expression of mutations clustered in the IVS7 donor splice site of factor VII gene. *Blood* 1998; **92**: 1646–51.
35 Mingozzi F, High KA. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat Rev Genet* 2011; **12**: 341–55.
36 Cohen JB, Snow JE, Spencer SD, Levinson AD. Suppression of mammalian 5′ splice-site defects by U1 small nuclear RNAs from a distance. *Proc Natl Acad Sci USA* 1994; **91**: 10470–4.
37 Incitti T, De Angelis FG, Cazzella V, Sthandier O, Pinnarò C, Legnini I, Bozzoni I. Exon skipping and duchenne muscular dystrophy therapy: selection of the most active U1 snRNA antisense able to induce dystrophin exon 51 skipping. *Mol Ther* 2010; **18**: 1675–82.
38 Beckley SA, Liu P, Stover ML, Gunderson SI, Lichtler AC, Rowe DW. Reduction of target gene expression by a modified U1 snRNA. *Mol Cell Biol* 2001; **21**: 2815–25.
39 Solnick D. Trans-splicing of mRNA precursors. *Cell* 1985; **42**: 157–64.
40 Yang Y, Mitchell Mansfield SG, Chao H, Walsh CE. RNA repair using spliceosome-mediated RNA trans-splicing. *Trends Mol Med* 2004; **10**: 263–8.
41 Chao H, Mansfield SG, Bartel RC, Hiriyanna S, Mitchell LG, Garcia-Blanco MA, Walsh CE. Phenotype correction of hemophilia A mice by spliceosome-mediated RNA trans-splicing. *Nat Med* 2003; **9**: 1015–9.
42 Coady TH, Lorson CL. Trans-splicing-mediated improvement in a severe mouse model of spinal muscular atrophy. *J Neurosci* 2010; **30**: 126–30.
43 Chao H, Walsh CE. RNA repair for haemophilia A. *Expert Rev Mol Med* 2006; **8**: 1–8.
44 Giannelli F, Green PM, Sommer SS, Poon M, Ludwig M, Schwaab R, Reitsma PH, Goossens M, Yoshioka A, Figueiredo MS, Brownlee GG. Haemophilia B: database of point mutations and short additions and deletions – eighth edition. *Nucleic Acids Res* 1998; **26**: 265–8.
45 Kemball-Cook G, Tuddenham EG, Wacey AI. The factor VIII structure and mutation resource site: HAMSTeRS version 4. *Nucleic Acids Res* 1998; **26**: 216–9.
46 Khajavi M, Inoue K, Lupski JR. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *Eur J Hum Genet* 2006; **14**: 1074–81.
47 Kisselev LL, Buckingham RH. Translational termination comes of age. *Trends Biochem Sci* 2000; **25**: 561–6.
48 Rospert S, Rakwalska M, Dubaquié Y. Polypeptide chain termination and stop codon readthrough on eukaryotic ribosomes. *Rev Physiol Biochem Pharmacol* 2005; **155**: 1–30.
49 Janzen DM, Geballe AP. Modulation of translation termination mechanisms by cis- and trans-acting factors. *Cold Spring Harb Symp Quant Biol* 2001; **66**: 459–67.
50 Recht MI, Fourmy D, Blanchard SC, Dahlquist KD, Puglisi JD. RNA sequence determinants for aminoglycoside binding to an A-site rRNA model oligonucleotide. *J Mol Biol* 1996; **262**: 421–36.
51 Böttger EC, Springer B, Prammananan T, Kidan Y, Sander P. Structural basis for selectivity and toxicity of ribosomal antibiotics. *EMBO Rep* 2001; **2**: 318–23.
52 Burke JF, Mogg AE. Suppression of a nonsense mutation in mammalian cells in vivo by the aminoglycoside antibiotics G-418 and paromomycin. *Nucleic Acids Res* 1985; **13**: 6265–72.
53 Zingman LV, Park S, Olson TM, Alekseev AE, Terzic A. Aminoglycoside-induced translational read-through in disease: overcoming nonsense mutations by pharmacogenetic therapy. *Clin Pharmacol Ther* 2007; **81**: 99–103.
54 James PD, Raut S, Rivard GE, Poon MC, Warner M, McKenna S, Leggo J, Lillicrap D. Aminoglycoside suppression of nonsense mutations in severe hemophilia. *Blood* 2005; **106**: 3043–8.
55 Srivastava A, Viswabandya A, Baidya S, Jayandharan G, George B, Shaji RV, Chandy M. Administration of gentamicin does not increase factor levels in severe hemophilia-B due to premature termination codons [abstract]. *J Thromb Haemost* 2003; **1** (Suppl. 1): CD043.
56 Yang C, Feng J, Song W, Wang J, Tsai B, Zhang Y, Scaringe WA, Hill KA, Margaritis P, High KA, Sommer SS. A mouse model for nonsense mutation bypass therapy shows a dramatic multiday response to geneticin. *Proc Natl Acad Sci USA* 2007; **104**: 15394–9.
57 Pinotti M, Rizzotto L, Pinton P, Ferraresi P, Chuansumrit A, Charoenkwan P, Marchetti G, Rizzuto R, Mariani G, Bernardi F; International Factor VII Deficiency Study Group. Intracellular readthrough of nonsense mutations by aminoglycosides in coagulation factor VII. *J Thromb Haemost* 2006; **4**: 1308–14.
58 Pinotti M, Rizzotto L, Chuansumrit A, Mariani G, Bernardi F; International Factor VII Deficiency Study Group. Gentamicin induces sub-therapeutic levels of coagulation factor VII in patients with nonsense mutations. *J Thromb Haemost* 2006; **4**: 1828–30.
59 Kopelowitz J, Hampe C, Goldman R, Reches M, Engelberg-Kulka H. Influence of codon context on UGA suppression and readthrough. *J Mol Biol* 1992; **225**: 261–9.
60 Poole ES, Brown CM, Tate WP. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in Escherichia coli. *EMBO J* 1995; **14**: 151–8.

61 Bonetti B, Fu L, Moon J, Bedwell DM. The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in Saccharomyces cerevisiae. *J Mol Biol* 1995; **251**: 334–45.
62 Manuvakhova M, Keeling K, Bedwell DM. Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system. *RNA* 2000; **6**: 1044–55.
63 Mingeot-Leclercq MP, Tulkens PM. Aminoglycosides: nephrotoxicity. *Antimicrob Agents Chemother* 1999; **43**: 1003–12.
64 Welch EM, Barton ER, Zhuo J, Tomizawa Y, Friesen WJ, Trifillis P, Paushkin S, Patel M, Trotta CR, Hwang S, Wilde RG, Karp G, Takasugi J, Chen G, Jones S, Ren H, Moon YC, Corson D, Turpoff AA, Campbell JA, *et al*. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* 2007; **447**: 87–91.
65 Hirawat S, Welch EM, Elfring GL, Northcutt VJ, Paushkin S, Hwang S, Leonard EM, Almstead NG, Ju W, Peltz SW, Miller LL. Safety, tolerability, and pharmacokinetics of PTC124, a nonaminoglycoside nonsense mutation suppressor, following single- and multiple-dose administration to healthy male and female adult volunteers. *J Clin Pharmacol* 2007; **47**: 430–44.
66 Rowe SM, Clancy JP. Pharmaceuticals targeting nonsense mutations in genetic diseases: progress in development. *BioDrugs* 2009; **23**: 165–74.
67 Kerem E, Hirawat S, Armoni S, Yaakov Y, Shoseyov D, Cohen M, Nissim-Rafinia M, Blau H, Rivlin J, Aviram M, Elfring GL, Northcutt VJ, Miller LL, Kerem B, Wilschanski M. Effectiveness of PTC124 treatment of cystic fibrosis caused by nonsense mutations: a prospective phase II trial. *Lancet* 2008; **372**: 719–27.
68 Sermet-Gaudelus I, Boeck KD, Casimir GJ, Vermeulen F, Leal T, Mogenet A, Roussel D, Fritsch J, Hanssens L, Hirawat S, Miller NL, Constantine S, Reha A, Ajayi T, Elfring GL, Miller LL. Ataluren (PTC124) induces cystic fibrosis transmembrane conductance regulator protein expression and activity in children with nonsense mutation cystic fibrosis. *Am J Respir Crit Care Med* 2010; **182**: 1262–72.
69 Guglieri M, Bushby K. Molecular treatments in Duchenne muscular dystrophy. *Curr Opin Pharmacol* 2010; **10**: 331–7.

Intron cleavage affects processing of alternatively spliced transcripts

Tibor Pastor, Andrea Dal Mas, Gabriele Talotti, et al.

RNA 2011 17: 1604-1613 originally published online June 14, 2011
Access the most recent version at doi:10.1261/rna.2514811

Intron cleavage affects processing of alternatively spliced transcripts

TIBOR PASTOR,[1] ANDREA DAL MAS,[1] GABRIELE TALOTTI, ERICA BUSSANI, and FRANCO PAGANI[2]
International Centre for Genetic Engineering and Biotechnology, Padriciano 99, 34149 Trieste, Italy

ABSTRACT

We previously showed that the insertion of a hammerhead ribozyme (Rz) in a critical intronic position between the EDA exon and a downstream regulatory element affects alternative splicing. Here we evaluate the effect of other intronic cotranscriptional cleavage events on alternative pre-mRNA processing using different ribozymes (Rz) and Microprocessor target sequences (MTSs). In the context of the fibronectin EDA minigene, intronic MTSs were cleaved very inefficiently and did not affect alternative splicing or the level of mature transcripts. On the contrary, all hammerhead Rz derivatives and hepatitis δ Rz were completely cleaved before a splicing decision and able to affect alternative splicing. Despite the very efficient Rz-mediated cleavage, the levels of mature mRNA were only reduced to ~40%. We show that this effect on mature transcripts occurs regardless of the type and intronic position of Rzs, or changes in alternative splicing and exon definition. Thus, we suggest that intron integrity is not strictly required for splicing but is necessary for efficient pre-mRNA biosynthesis.

Keywords: ribozymes; alternative splicing; cotranscriptional processing; Microprocessor target sequence

INTRODUCTION

A typical mammalian protein-coding gene transcribed by RNA polymerase II (Pol II) contains relatively short exonic sequences separated by long introns, which average 3 kb but are frequently longer than 10 kb. Introns are removed from pre-mRNAs by the spliceosome and, according to EST data, >90% of human genes undergo alternative splicing (Pan et al. 2008; Wang et al. 2008). The efficiency and accuracy of pre-mRNA splicing depend on complex interactions on the nascent pre-mRNAs. Small nuclear ribonuclear proteins assemble on splice sites along with splicing factors such as SR proteins and hnRNPs that interact with exonic and/or intronic regulatory sequences (Maniatis and Reed 2002). Several lines of evidence indicate that pre-mRNA splicing and transcription are coupled events. Nuclear imaging studies in *Drosophila* and polytene chromosomes indicated that intron looping can occur while the transcript is tethered to the chromatin template (Beyer et al. 1981; Beyer and Osheim 1988). During transcription, splicing factors are recruited to actively transcribed genes (Gornemann et al. 2005; Lacadie et al. 2006; Listerman et al. 2006), and the accumulation of spliceosomal components requires Pol II with an intact C-terminal domain (CTD) (Misteli and Spector 1999). Pol II transcripts are more efficiently processed than T7 transcripts both in in vivo and in vitro assays in which transcription and splicing are linked (Das et al. 2006; Hicks et al. 2006; Natalizio et al. 2009). Published data support a key role for Pol II and specifically for its CTD in the coupling of transcription and pre-mRNA processing (McCracken et al. 1997), including a direct contact with splicing factors (Morris and Greenleaf 2000; Rosonina et al. 2005). Recently, SR proteins and components of U1 snRNP have been suggested to be involved in coupling transcription to pre-mRNA processing via the CTD (Das et al. 2007). Coupling of pre-RNA processing and transcription also has an influence on alternative splicing. Promoter type, rate of elongation, transcriptional activators, and chromatin-remodeling factors have been shown to affect alternative splicing (Kornblihtt 2005, 2007; Luco et al. 2010).

Coupling of pre-RNA processing and transcription is thought to facilitate processing of intronic sequences through exon tethering by Pol II (Dye et al. 2006). The existence of a molecular tether between the nascent pre-mRNA and RNA Pol II transcripts comes from the original observation that RNAP II transcripts derived from β-globin are efficiently spliced in vivo when the intron is cotranscriptionally cleaved

[1]These authors contributed equally to this work.
[2]**Corresponding author.**
E-mail pagani@icgeb.org.
Article published online ahead of print. Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.2514811.

by a hammerhead ribozyme (Dye et al. 2006). In addition, evidence that intronic cotranscriptional cleavage does not affect splicing has been reported in yeast (Lacadie et al. 2006). Exon tethering is also consistent with the fact that several introns contain pri-miRNAs hairpins, which have been reported to be cotranscriptionally cleaved by the Microprocessor without affecting the splicing of adjacent constitutive exons (Kim and Kim 2007; Morlando et al. 2008). The excision of pri-miRNAs from introns involves a specific complex called the Microprocessor, containing Drosha, an RNase III–like enzyme, and its cofactor DGCR8 (Denli et al. 2004). However, it is not known if the Microprocessor complex crops pri-miRNAs cotranscriptionally with a different efficiency and if this affects the regulation of alternative splicing.

In contrast to the less efficient hammerhead Rz, δ hepatitis Rz was recently reported to inhibit splicing in the β-globin system (Fong et al. 2009), suggesting that exon tethering of Pol II to rescue splicing of a severed transcript depends on the relative activity of the Rz in comparison to the rate of splicing (Fong et al. 2009). A rapid cotranscriptional cleavage of the β-globin intron, as provided by the δ hepatitis Rz, but not by the hammerhead Rz, aborts pre-mRNA processing with a nearly complete inhibition of gene expression.

We previously showed that, in contrast to the constitutive β-globin, two alternatively spliced exons (the fibronectin EDA and α-TM) are more sensitive to cotranscriptional cleavage mediated by an Rz and more prompt to changes in their pattern of splicing. In particular, placing the N117 derivative of a *Schistosoma mansoni* (*Sm*) hammerhead Rz between the EDA exon and its negative intronic downstream regulatory element (DRE) affects EDA alternative splicing (Gromak et al. 2008). The N117 Rz was selected for its extremely high cleavage activity in vivo (Yen et al. 2004), and, when placed within an intron of an expression vector, it did not apparently inhibit the expression of a reporter gene (Yen et al. 2004). When the N117 Rz was positioned in the middle of the relatively long introns flanking the EDA exon and outside the DRE, it did not affect the regulation of alternative splicing (Gromak et al. 2008). Thus, in the case of alternative splicing, the continuity of the nascent transcript between the EDA exon and DRE is essential for the alternative splicing of the fibronectin EDA minigene. In addition, the DRE has been recently shown to affect the relative order of intron removal (de la Mata et al. 2010). As cotranscriptional cleavage between the EDA and its regulatory element affects alternative splicing, this system provides a useful model to evaluate exon tethering and investigate the effect of different Rzs and MTSs on pre-mRNA processing.

In this study, using the EDA alternative splicing minigene system, we evaluated intronic cleavage events mediated by different Rzs and Microprocessor target sequences and their effect on alternative splicing regulation. We show that all Rzs, in contrast to MTSs and consistent with their rapid cotranscriptional cleavage, affect EDA alternative splicing. MTS are processed less efficiently but, interestingly, showed a variable residual cleavage activity. On the other hand, we observed that all Rzs reduce the mRNA levels to ∼40% regardless of the distance from splice sites, alternative splicing pattern and exon definition, and relative cleavage efficiency. We suggest that in our system the disruption of the intron integrity is not completely rescued by Pol II–mediated exon tethering and that defective transcripts are to some extent subjected to abortive processing.

RESULTS

The hammerhead ribozyme interrupts the continuity of the nascent transcript and facilitates splicing of downstream introns in a DRE-dependent manner

We previously showed that the insertion of the N117 hammerhead Rz between the alternatively spliced EDA exon and its downstream regulatory element (DRE) induces exon inclusion (Gromak et al. 2008). To study in more detail the cleavage efficiency of the hammerhead Rz, we have analyzed the nascent transcripts by means of RT-PCR. To amplify nascent pre-mRNAs across the Rz cleavage site and distinguish them from endogenous fibronectin transcripts, we inserted a short 25-nt-long spacer within the intron (Fig. 1A, pEDAL and pEDAs minigenes derivatives). As shown in Figure 1B, the spacer insertion did not modify the previously reported splicing pattern, whereas the wild-type Rz pEDAL2N117 induced exon inclusion, and the inactive Rz had no effect on spliced mRNA (Fig. 1B, lanes 1–3). Analysis of the pre-mRNAs showed that the active Rz had no significant effect on the abundance of upstream and downstream nascent transcripts (Fig. 1C, lanes 1–3, *a* and *c* products, respectively), while the nascent unspliced pre-mRNA across the Rz was completely absent (Fig. 1C, lanes 1–3, *b* product). To confirm the absence of nascent pre-mRNA across the Rz, we transfected pEDAL alone or together with pEDAL2N117 or pEDAL2N117m constructs, followed by amplification of the *b* pre-mRNA. Cotransfection of pEDAL with the inactive mutant Rz minigene pEDAL117m, followed by amplification of the *b* fragment, showed the two nascent pre-mRNAs originating from the two minigenes, as expected. The upper band incorporates the mutant Rz, whereas the lower band derives from the empty pEDAL minigene (Fig. 1D, lane 3). On the contrary, cotransfection of pEDAL along with pEDALN117 did not show any upper band derived from the active Rz that is consistent with its high cleavage efficiency (Fig. 1D, lane 2). We also evaluated the abundance of the nascent pre-mRNA in the pEDs minigenes (Fig. 1A) in which the 5′ss of the exon was strengthened, rendering the exon constitutively included (Fig. 1B, lanes 4–6). Again, amplification across the active Rz did not generate any product (Fig. 1C,D, lane 5), whereas the amplification of the upstream and downstream sequences (*a* and *c* products) was not affected (Fig. 1C,

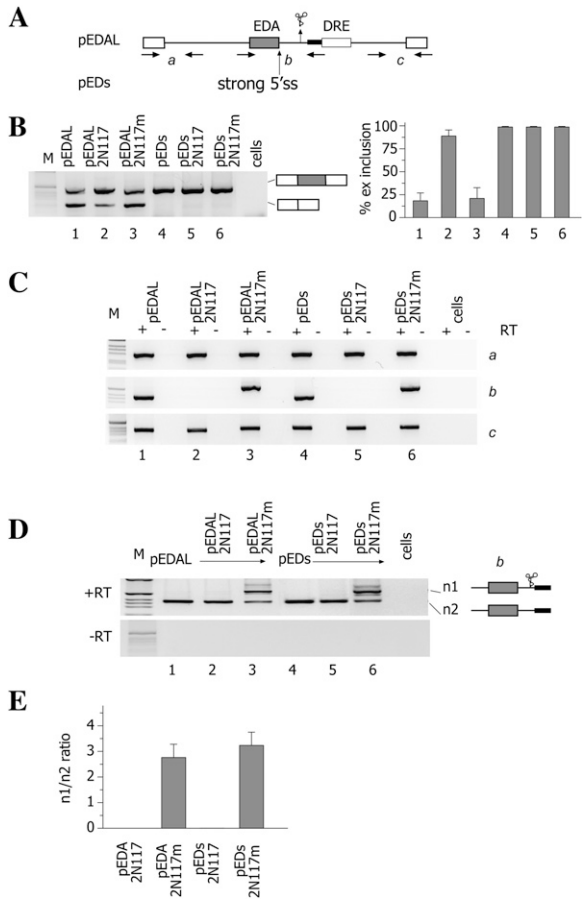


FIGURE 1. Cleavage efficiency mediated by the N117 hammerhead ribozyme affects EDA alternative splicing. (*A*) Diagram of the pEDAL and pEDs minigenes. The position of the N117 ribozyme is indicated. To distinguish amplification products in cotransfection experiments, two active or mutant N117 hammerhead ribozymes were inserted in tandem between the EDA exon and the DRE. The DRE is indicated, and the thick line represents the 25-nt-long spacer (see text). The primers used in RT-PCR analysis to amplify nascent transcripts are indicated. (*B*) Analysis of mature spliced transcripts. RT-PCR analysis of total RNA from Hep3B cells expressing different minigenes was performed, and the resulting amplification products were resolved on 2% agarose gel. The two splicing products are schematically indicated. (M) The 1-kb molecular weight marker. (Histogram on the *right*) The percentage of EDA exon inclusion (±SD) based on at least three independent duplicate experiments. (*C*) Analysis of nascent transcripts. Amplification of nascent mRNA with (+) and without (−) RT resolved on 2% agarose gel. The identity of the amplified fragments (*a*, *b*, and *c*) is reported on the *right* of each panel. (*D*) Analysis of nascent *b* transcript across the Rzs. The indicated minigenes (0.45 μg) were cotransfected with pEDAL (0.15 μg) (lanes *1–3*) or with pEDs (0.15 μg) (lanes *4–6*) in Hep3B cells and the *b* fragment amplified from total RNA. The identity of *b* bands originating from pEDAL or pEDS (n1) and other minigenes (n2) is reported. (*Lower* panel) The PCR control, without reverse transcriptase (−RT). (*E*) Ratio of the intensity of the n1/n2 bands in *D* expressed as the mean ±SD of three independent experiments done in duplicate. Quantification of pre-mRNA was done using a fluorescent primer.

lanes 4–6). We also tested in vitro the activity of N117Rz in the context of the pEDA intron, and we it found it to be extremely effective, cleaving the RNA with an almost 100% efficiency (Supplemental Fig. 1). The absence of any nascent transcript deriving from the minigenes containing the wild-type hammerhead N117 Rz indicates that this element is very efficiently cleaved before splicing and that it interrupts the continuity of the nascent transcript.

To understand why the Rz-mediated cleavage of the nascent transcript induces exon inclusion, we initially considered the possibility that 5′–3′ exonucleases can cotranscriptionally degrade the DRE. However, silencing of the Xrn2 5′–3′ exonuclease did not affect the splicing pattern in minigenes with either wild-type or mutant Rzs, indicating that a cotranscriptional degradation of the DRE is not implicated in the changes of EDA alternative splicing (Supplemental Fig. 2).

An alternative explanation for the effect of the Rz on mature transcript could be a change in the splicing efficiency of the intron that contains the Rz. Previous studies both in the natural chromosomal context and in minigene systems have shown that there is a preferential removal of the intron downstream from the EDA exon before the upstream intron has been removed (Pandya-Jones and Black 2009; de la Mata et al. 2010). Interestingly, deletion of the DRE was shown to affect the relative rate of the downstream intron excision, decreasing the amount of the 5′-splicing intermediate (de la Mata et al. 2010). To evaluate the possible effect of the Rz-mediated interruption of the nascent transcript on the intron splicing efficiency, we analyzed the 5′-splicing intermediate by semiquantitative RT-PCR amplification. Consistent with the previous observation, deletion or inversion of the DRE in our minigene system reduces the amount of 5′-splicing intermediate (Fig. 2C, lanes 2,3). The diminished level of the 5′-splicing intermediate reflects an increase in the splicing efficiency of the downstream intron. A similar effect was observed when the active Rz, but not the inactive one, was inserted between the exon and the DRE (Fig. 2C, lanes 4 and 5, respectively). On the contrary, insertion of the Rz after the DRE did not affect the amount of 5′ intermediate (Fig. 2C, lane 7), indicating that the Rz reduces the 5′-splicing intermediate only when it cleaves the nascent transcript between the EDA exon and the DRE. Thus, the interruption of the continuity of the nascent transcript mediated by the Rz facilitates splicing of the downstream intron in a DRE-dependent manner.

The intronic hammerhead N117 ribozyme modulates alternative splicing in different minigene contexts

To further explore the effect of the Rz on alternative splicing, we tested both the N117 Rz and the DRE regulatory element in the context of the CFTR exon 12 (Fig. 3A). Transient transfection of the normal CFTR minigene produced a splicing pattern in which ~85% of the exon is included in the final mRNA (Fig. 3B, lane 1; Pagani et al. 2003). Cloning of the DRE downstream from

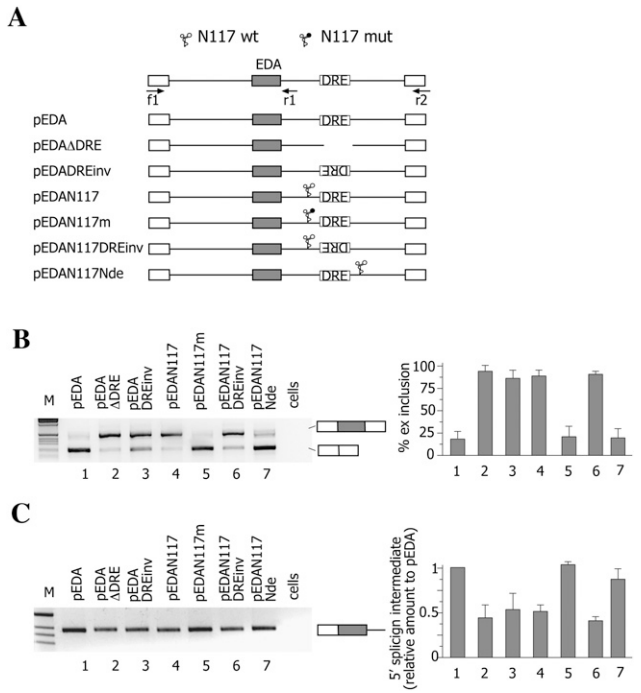


FIGURE 2. The N117 ribozyme facilitates splicing of the downstream intron in a DRE-dependent manner. (*A*) Diagram of the pEDA minigenes. The position of the ribozyme (wild-type [WT] or mutant) and of the DRE sequence is indicated. (Arrows) The locations of primers used in RT-PCR analysis. Exonic sequences are boxed. (*B*) Analysis of mature spliced transcripts. Minigenes (0.5 μg) were transfected into Hep3B cells and the pattern of splicing evaluated using f1 and r2 primers. (M) The 1-kb molecular weight marker. The identity of the splicing products is schematically reported on the *right* of the panel. The histogram shows the percentage of EDA exon inclusion. Data represent the mean ± SD of three independent experiments performed in duplicates. (*C*) Analysis of 5′ splicing intermediates. Minigenes (0.5 μg) were transfected into Hep3B cells along with 0.1 μg of pCF1 plasmid, which was used as control for normalization of transfection and reverse transcriptase efficiencies. RT-PCR was carried out with a1 and r1 primers and amplification products resolved on 2% agarose gel. The identity of the band is indicated on the *right* of the panel. The histogram shows the relative amount of EDA 5′-splicing intermediates expressed as normalized ratio to pEDA. Calculations show mean ± SD at two independent experiments done in duplicate. Real-time PCR was carried out on pCF1 transcript to assess transfection efficiency (see Materials and Methods).

the CFTR exon 12 reduced the percentage of exon inclusion to ~40% (Fig. 3B, lane 4), whereas the DRE in inverted orientation did not change the splicing pattern (Fig. 3B, lane 5). The insertion of an active Rz, but not of an inactive one, between the CFTR exon 12 and the DRE restored the original splicing pattern, suggesting that the continuity of the nascent pre-mRNA molecule is required for the DRE-dependent splicing inhibition also in the case of CFTR (Fig. 3B, lanes 6 and 7, respectively). On the other hand, in the absence of an intronic DRE, wild-type and mutant Rzs alone did not affect the splicing pattern (Fig. 3B, lanes 2 and 3, respectively), consistent with the absence of a natural downstream intronic splicing regulatory element in CFTR exon 12.

To study the cleavage efficiency of the hammerhead Rz in the CFTRex12 context, we have analyzed the pre-mRNAs: The nascent transcripts showed that the active Rz had no significant effect on pre-mRNA abundance of upstream and downstream nascent transcripts (Fig. 3C,D, lanes 1–3, *a1–a2* and *c1–c2* products), whereas the nascent pre-mRNA across the Rz was completely absent (Fig. 3C, lanes 2, *b1–b2*; Fig. 3D, *b1–b3*). We confirmed the absence

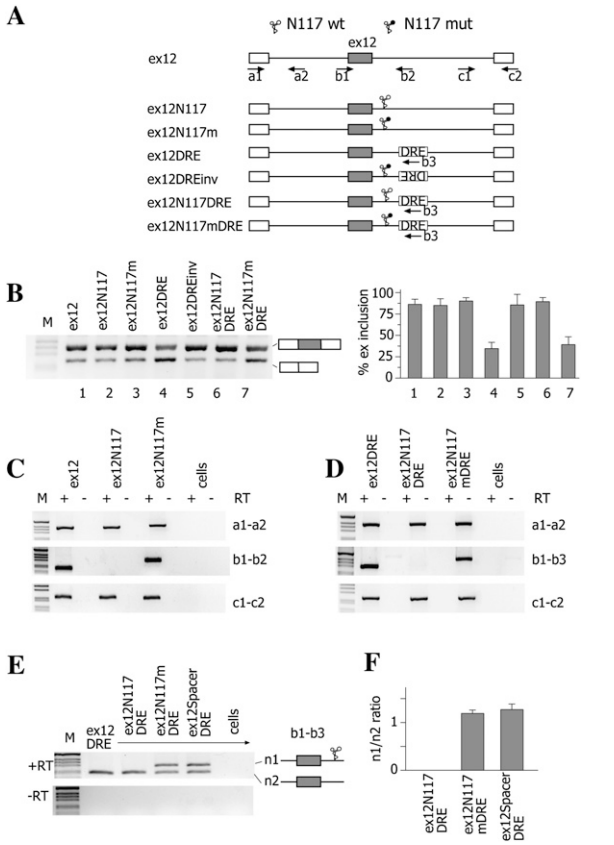


FIGURE 3. The N117 hammerhead ribozyme affects CFTR exon 12 alternative splicing in a DRE-dependent manner. (*A*) Schematic representation of the CFTR exon 12 minigenes. The positions of wild-type (WT) and mutant N117 ribozymes and of DRE are indicated. (Arrows) The locations of primers used in RT-PCR analysis. Exonic sequences are boxed. (*B*) Analysis of mature spliced transcripts. Minigenes were transfected into Hep3B cells, and RT-PCR products amplified with a1 and c2 primers were resolved on a 2% agarose gel. Exon 12 inclusion or exclusion forms are indicated. (M) The 1-kb molecular weight marker. The histogram shows the percentage of CFTR exon 12 inclusion (±SD) based on at least three independent duplicate experiments. (*C,D*) Analysis of nascent transcripts. Ex12 minigenes were transfected in Hep3B cells and analyzed with the indicated pairs of primers reported on the *right* of each panel. RT(−) is the control in absence of reverse transcriptase. (M) The 1-kb molecular weight marker. (*E*) Analysis of the *b1–b3* pre-mRNA transcript. The indicated minigenes (0.45 μg) were cotransfected with pTB ex 12 DRE (0.15 μg). Total RNA was amplified with b1 and b3 primers, and the identity of the two resulting products (n1 and n2) is indicated. (*Lower* panel) The amplification without reverse transcriptase (−RT). (*F*) Quantification of pre-mRNA n1 and n2 ratio expressed as mean ± SD of two independent experiments done in duplicate.

of nascent pre-mRNA across the active Rz through cotransfection experiments in which no amplification of the *b1–b3* product was observed (Fig. 3E,F). Furthermore, the mutant Rz and a control minigene (ex12 spacer DRE) showed a band of similar intensity, indicating that no residual cleavage activity is present in the mutant Rz, as also evaluated by quantitative analysis (Fig. 3F). On the other hand, the small difference in the isoform ratio between the ex12DRE and ex122N117mDRE is not due to the residual activity of the mutant Rz but possibly as a consequence of a minor effect of the transcribed mutant Rz sequences on the DRE.

We also evaluated the effect of the N117 Rz in the context of the EDA combined with a strong downstream intronic silencer consisting of a GAA repeat expansion (Baralle et al. 2008). The wild-type (WT) Rz but not the mutant was able to rescue EDA exon skipping, and analysis of nascent intermediates did not show any PCR product derived from amplification of the WT Rz (Supplemental Fig. 3; Supplemental Data). These results, obtained with different minigene systems, are entirely consistent with the hypothesis that the Rz-mediated interruption of the continuity of nascent RNA between the exon and the intronic splicing regulatory element affects alternative splicing regulation.

Both hammerhead and hepatitis δ ribozymes efficiently cleave nascent mRNA and affect EDA alternative splicing regulation

To clarify the role of the catalytic activity of different Rzs on alternative splicing regulation, we evaluated the hepatitis δ Rz, which has been suggested to be more active in vivo when compared to the hammerhead Rz (Fong et al. 2009), as well as derivatives of the hammerhead itself (Yen et al. 2004). The hammerhead derivatives were specifically developed in vivo and showed different cleavage efficiencies in comparison to the original hammerhead, with 1401 ± 228–, 225 ± 70–, and 53 ± 13–fold increase for N117, N79, and N93, respectively (Yen et al. 2004). In addition, transcripts containing the N79 ribozyme were recently shown to be completely cleaved in vivo in transfected cells (Chen et al. 2009). The Rzs or their corresponding inactive variants were inserted between the EDA exon and the DRE, and the resulting minigenes were analyzed for the EDA splicing pattern and nascent pre-mRNA. In all cases, only active Rzs induced EDA exon inclusion (Fig. 4A) and showed no nascent pre-mRNA (Fig. 4B, *b* product). Quantitative analysis in cotransfection experiments confirmed the absence of nascent products (Fig. 4C,D). Furthermore, the pre-mRNA abundance across the upstream or downstream exon junctions was not significantly affected (Fig. 4B, lanes 1–3, *a* and *c* products). Thus, we did not detect in our system significant differences in the cleavage activity between hammerhead and hepatitis δ Rzs. More precisely, both types of Rzs possess a sufficient cleavage activity in vivo that leads to equally efficient interruption of the nascent transcript and comparable changes in the EDA alternative splicing pattern.

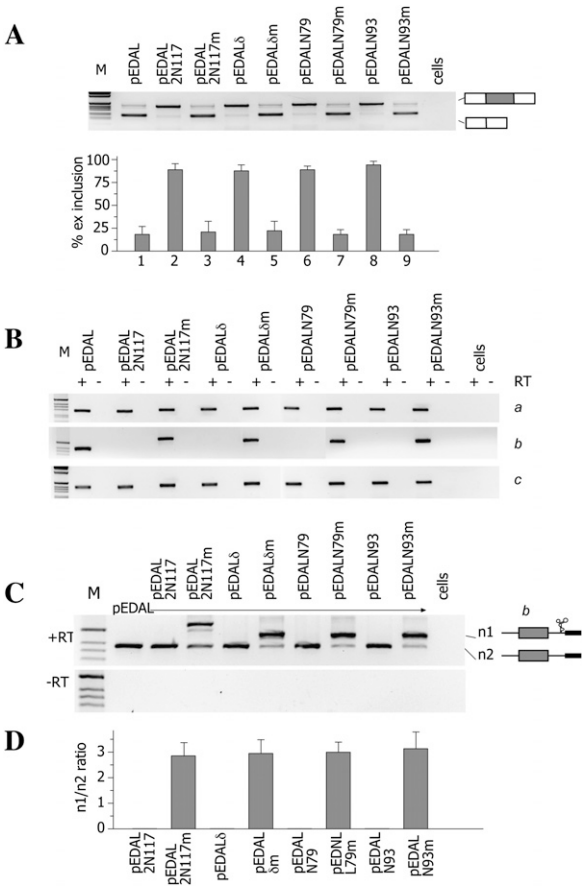


FIGURE 4. Effects of different ribozymes on EDA alternative splicing and nascent transcripts. (*A*) Analysis of mature spliced transcripts. The indicated minigenes derived from pEDAL (Fig. 1A) contain different Rzs or their corresponding inactive variant inserted between the EDA and the DRE. Minigenes were transfected into Hep3B cells, total RNA was extracted, and its splicing pattern was analyzed by RT-PCR. (M) The 1-kb molecular weight marker. The identity of the corresponding amplified bands is shown. The histogram shows the percentage of EDA exon inclusion, and data are the mean ± SD of three independent experiments performed in duplicate. (*B*) Analysis of nascent transcripts. Nascent transcripts derived from transfections of the indicated minigenes in cells were analyzed by RT-PCR using the specific pairs of primers indicated in Figure 1A. The identity of the bands (*a*, *b*, *c*) is shown. (*C*) Analysis of nascent *b* pre-mRNA transcript across the ribozyme. Minigenes (0.45 μg) were cotransfected with pEDAL (0.15 μg). Total RNA was amplified with the specific pair of primers indicated in Figure 1A. (M) The 1-kb molecular weight marker. The identity of the bands (n1 and n2) is shown. (*Lower* panel) A control PCR, without reverse transcriptase (−RT). (*D*) Quantification of n1 and n2 ratio of *C* expressed as the mean ± SD of two independent experiments done in duplicate.

Microprocesssor target sequences are cleaved less efficiently and do not affect EDA alternative splicing

As primary miRNA transcripts are processed cotranscriptionally (Kim and Kim 2007; Morlando et al. 2008; Pawlicki

and Steitz 2008), we evaluated the effect of different Microprocessor target sequences (MTSs) on the regulation of EDA alternative splicing. We hypothesized that if the cropping by the Microprocessor is as fast as by Rzs, it will affect EDA alternative splicing. We tested miR-26b and miR-330 pri-miRNA hairpins, a cropping-deficient miR-26b mutant (miR-26bcrop) (Kim and Kim 2007), and two MTS hairpins of DGCR8 mRNA. The last two sequences are located in the 5′ UTR region and in exon 2 of DGCR8, respectively, and are the target for the Microprocessor complex and act in a DGCR8 autoregulatory feedback to degrade their own mRNA (Han et al. 2009; Shenoy and Blelloch 2009; Triboulet et al. 2009). Different MTS hairpin sequences, along with ~60 bp of corresponding natural flanking regions, were inserted between the EDA exon and its intronic regulatory element. The resulting minigenes were transfected in Hep3b cells for the analysis of the splicing pattern (Fig. 5A) and evaluation of miRNA production. Evaluation of miRNA production revealed that transfection of the miR-26b and miR-330 minigenes, but not of the defective miR-26bcrop, produced the corresponding mature miRNAs (Supplemental Fig. 4): This indicated that in pEDA minigene context, the pri-miRNA hairpins are correctly processed by the Microprocessor complex. Analysis of the splicing pattern showed that, in contrast to N117 Rz, which increases the percentage of exon inclusion, none of the MTSs affected the EDA splicing pattern significantly, suggesting an inefficient cotranscriptional cleavage of nascent pre-mRNA (Fig. 5A). As MTSs did not affect EDA alternative splicing when inserted between the exon and the intronic regulatory element, it is possible that their relative cotranscriptional cleavage activity is reduced in comparison with the activity of Rzs. To study the efficiency of cleavage mediated by different MTSs, we evaluated the abundance of nascent pre-mRNA across the MTSs in cotransfection experiments. Each minigene, containing MTSs or Rzs, was cotransfected with the control pEDAL construct, and the resulting RNAs were amplified to obtain the *b* product. Two major bands were identified, the lower one resulting from amplification of pEDA-derived transcripts and the higher one from amplification of MTS- or Rz-derived splicing intermediates. While no upper band was evident in the case of pEDAL2N117 due to its very efficient Rz cleavage, in all the other cases 5′-splicing intermediate and nascent pre-mRNA products were detected (Fig. 5B). Interestingly, analysis of nascent pre-mRNA revealed different amounts of transcripts derived from MTSs (Fig. 5B), suggesting a different relative cleavage efficiency of the MTSs. To evaluate the ratio between the intensity of the two bands, we performed RT-PCR using a fluorescent oligonucleotide. The highest ratio between the two bands was observed for miR-330 and DGCR8 minigenes, whereas pEDA26b showed the lowest ratio (Fig. 5C). In comparison to miR-26b, miR-26bcrop showed an increase in the ratio, compatible with a reduced

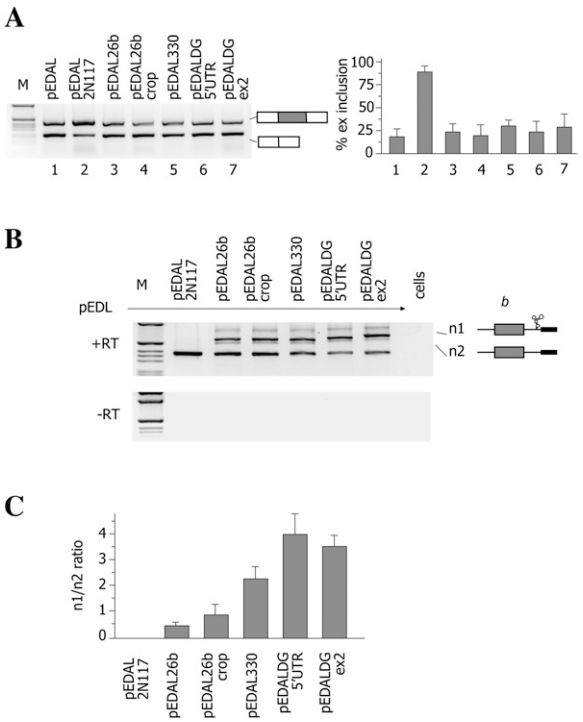


FIGURE 5. Effects of Microprocessor target sequences (MTSs) on EDA alternative splicing and pre-mRNA processing. (*A*) Analysis of mature spliced transcripts. The indicated minigenes derived from pEDAL (Fig. 1A) contain different Microprocessor target sequences inserted between the EDA and the DRE. Minigenes were transfected in Hep3B cells and total RNA analyzed by RT-PCR to detect mature transcripts. (M) The 1-kb molecular weight marker. The identity of the corresponding amplified bands is indicated. The histogram shows the percentage of EDA exon inclusion. Data represent the mean ± SD of three independent experiments performed in duplicate. (*B*) Analysis of pre-mRNA transcripts. The indicated minigenes (0.45 μg) were cotransfected with pEDAL (0.15 μg) in Hep3B cells, and the nascent *b* fragment was amplified with the specific pair of primers indicated in Figure 1A. The identity of the *b* bands originating from pEDAL (n1) or other minigenes (n2) is indicated. (*Lower* panel) A control PCR, without reverse transcriptase (−RT). (*C*) Quantification of n1 and n2 ratio expressed as the mean ± SD of two independent experiments done in duplicate.

but not completely abolished activity in processing of the nascent mRNA by the Microprocessor complex. No transcripts were observed from minigenes with wild-type N117 Rz. Altogether these results indicate that the Rzs are rapidly and more efficiently cleaved than the MTSs; this could explain the lack of effect of MTSs on the splicing pattern of the alternative EDA exon (Fig. 5A). However, among the MTSs, miR-26b is cotranscriptionally processed faster than the DGCR8 and miR-330 hairpins.

Different Rzs but not MTSs reduce mRNA abundance regardless of exon definition, alternative splicing pattern, and intronic position

As the relative cotranscriptional cleavage efficiency has been recently shown to severely affect processing of α-globin

transcripts (Fong et al. 2009), we have evaluated how insertion of the different Rzs or the MTSs in the fibronectin intron can affect the mRNA levels in vivo. Considering the competition between the Rz cleavage and splicing efficiency, as previously proposed (Fong et al. 2009), one can speculate that the differences in the amount of mature transcripts produced can depend on the cleavage efficiency, on the intronic position of the Rz, or on alternative splicing pattern. To quantify the total amount of spliced mRNA derived from the minigenes, we used an RT-PCR TaqMan assay, which amplifies the spliced α-globin portion common to all the minigenes. After normalization for transfection and RT efficiency, we compared the total amount of spliced mRNA produced for each wild type and its corresponding inactive ribozyme. Even when the N117 hammerhead ribozyme in the Nco position was completely cleaved (Fig. 2, pEDALN117 construct), it did not entirely reduce the amount of mature mRNA (Fig. 6B). In fact, relative to the inactive Rz, ~40% of expression is retained. Interestingly, the hammerhead N79 and N93 derivatives and the hepatitis δ Rz showed a similar level of expression when inserted at the same position (Fig. 6). We also evaluated the role of exon definition, studying the N117 Rz in the context of the pEDs minigene in which, due to a strong 5′ss, the EDA exon is completely included (Fig. 2, pEDs construct). Also in this case, we observed ~40% of mRNA levels. Lastly, we studied the role of the intronic position since the N117 Rz at the Nco site is only 30 bp downstream from the EDA exon. We evaluated the Rz-inserted 433 bp and 900 bp downstream at the Nde, Bcl positions as well as at the Kpn site, which is 622 bp upstream of the alternatively spliced EDA exon (Fig. 6A). The analysis of relative mRNA levels again gave a constant reduction to 40% levels. As the active Rz induced the exon inclusion only when inserted into the NcoI site (Fig. 1), with no effect on splicing pattern when introduced into other intronic positions (Fig. 2B, pEDAN117Nde construct; Gromak et al. 2008), this means that the reduced level of mRNA is not due to the changes associated with alternative splicing mediated by the Rzs. On the other hand, miR-26b and miR-330 pri-miRNAs are very inefficiently cleaved and do not significantly affect the amount of mRNAs (Fig. 6C). Altogether these data indicate that, in the FN minigenes, a constant amount of transcript is processed regardless of exon definition, alternative splicing pattern, or distance between the Rz cleavage site and the flanking exons.

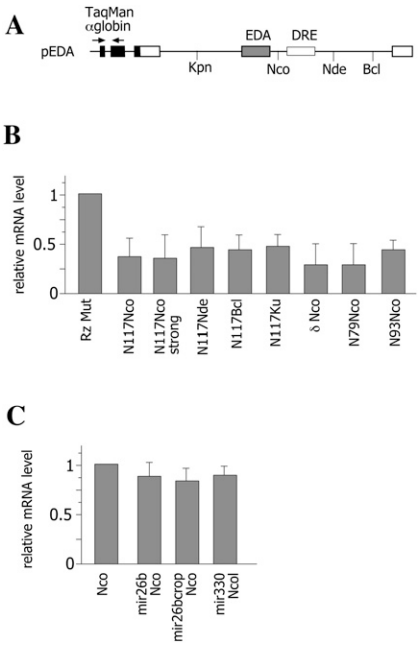


FIGURE 6. Analysis of gene expression in ribozyme- and MTS-derived minigenes. (*A*) Diagram of the fibronectin EDA minigene. Fibronectin exons (white boxes); α-globin exons (black boxes); introns (lines). The unique restriction sites in the introns used for wild-type (WT) or mutant ribozyme insertion are indicated. (Arrows) The positions of the TaqMan α-globin amplification primers. Minigenes were transfected in HeLa cells, and their expression levels were monitored by real-time PCR using a specific α-globin TaqMan assay. Each minigene (0.5 μg) was cotransfected with the control pCF1 (0.1 μg), which was used for normalization of transfection and reverse transcriptase efficiencies. (*B*) Relative expression levels of mRNAs produced from minigenes carrying WT or mutant ribozymes. The relative WT Rzs levels were normalized to the corresponding mutant Rz levels, and data are shown as the mean of at least two transfection experiments done in duplicate. Error bars represent the SEM. (*C*) Relative expression levels of mRNAs produced from miRNA-derived minigenes. The relative miRNA levels were normalized to the pEDA Nco empty vector levels, and data are shown as the mean of at least two transfection experiments done in duplicate. Error bars represent the SEM.

DISCUSSION

Intronic sequences in Pol II–derived pre-mRNA transcripts are cotranscriptionally processed by the spliceosome and by the Microprocessor to generate mature mRNAs and miRNAs, respectively. We have used different ribozymes and Microprocessor target sequences to study the effect of intronic cotranscriptional cleavage events on pre-mRNA processing and regulation of alternative splicing. Consistent with the exon tethering model (Dye et al. 2006), an Rz-mediated interruption of the continuity of the nascent transcript affects alternative splicing when this occurs between the exon and its downstream splicing regulatory element, in the fibronectin EDA minigene (Fig. 1) and in similar minigene systems (Fig. 3; Supplemental Fig. 3). On the other hand, MTSs are cleaved very inefficiently and do not affect either alternative splicing (Fig. 5) or the level of mature transcripts (Fig. 6). The DRE deletion itself affects the order of intron removal as previously reported (de la Mata et al. 2010), and we show here that the Rz-mediated interruption of the continuity of the nascent transcript facilitates splicing of the downstream intron in a DRE-dependent manner increasing exon inclusion (Fig. 2). When the Rz was inserted in other intronic

positions, neither alternative splicing of mature mRNA nor splicing of the downstream intron was affected (Fig. 2).

Insertion of an artificial hammerhead Rz in the short intron of the constitutively included α-globin minigene was originally reported to have no effect on the abundance of mature mRNA or on splicing, suggesting that exons are efficiently kept in place by tethering to the Pol II (Dye et al. 2006). However, this exon-tethering model has been recently challenged by the observation that a different type of ribozyme, the hepatitis δ, aborts cotranscriptional pre-mRNA processing in a similar system, resulting in severe damage of the nascent transcript (Fong et al. 2009). This effect was related to a faster catalytic activity of the hepatitis δ relative to the hammerhead, suggesting that there may be a competition between cleavage by the Rz and splicing efficiency (Fong et al. 2009). We tested different types of Rzs (hammerhead and hepatitis δ), and all of them, when inserted between the EDA and the DRE, induced exon inclusion (Fig. 4A). Through the amplification of nascent transcripts across the Rz insertions, we did not detect significant differences in their catalytic activity (Fig. 4C,D), suggesting that all Rzs are efficiently cleaved before the splicing decision. In comparison to the α-globin construct, the EDA minigene contains relatively longer introns, and the central exon, being alternatively spliced, is less defined. These aspects might explain the comparable cleavage efficiency of the Rzs in the fibronectin minigene and their similar effect on the pattern of EDA alternative splicing.

The Rz-mediated cleavage of intronic sequences is associated with a constant reduction in the level of mature mRNAs with ~40% of mRNA produced (Fig. 6B). However, we did not observe a relationship between different amounts of mature transcripts and type of Rz, intronic position, or strength of the 5′-splice site (Fig. 6B). If intronic cleavage occurs before the engagement of Pol II to the successive exon, this will result in the release of transcripts from RNA Pol II and will therefore abort cotranscriptional pre-mRNA processing. In our alternatively spliced system, with relatively long introns, the constant effect of the Rzs on the level of mature transcripts indicates that cleavage of the nascent transcript can affect pre-mRNA processing to a certain extent. As this effect was independent of the alternative splicing pattern, EDA exon definition, Rzs type, and intronic position, a competition between the Rz cleavage and splicing efficiency is not involved. Thus, the Rz-dependent disruption of intron integrity is not completely rescued by Pol II–mediated exon tethering. A certain amount of severed transcript is probably not tethered by the CTD of Pol II and is consequently subjected to abortive pre-mRNA processing, or the spliced transcript is not properly exported from the nucleus. The efficiency of rescue of interrupted intronic transcripts might be gene-specific and dependent on Pol II processivity. Pol II elongation rates and/or its CTD phosphorylation status might regulate exon tethering in a transcript-specific manner, thus facilitating processing of long intronic sequences. Intriguingly, as splicing of long introns in *Drosophila* is facilitated by components of the pre-exon junction complex (EJC) (Ashton-Beaucage et al. 2010; Roignant and Treisman 2010), this complex might be involved in modulating exon tethering by Pol II in a transcript-specific manner.

Using the alternatively spliced EDA system, we have also evaluated to what extent different MTSs are cropped before splicing and what effect they have on alternative splicing regulation. Several lines of evidence indicate that pri-miRNAs are cropped cotranscriptionally (Kim and Kim 2007; Morlando et al. 2008; Pawlicki and Steitz 2008). In our minigene system, MTSs showed a significantly lower cleavage activity in comparison to the Rzs (Fig. 5) and do not affect the splicing pattern when inserted between the EDA exon and the DRE, suggesting that Microprocessor cropping occurs after transcription of the downstream regulatory element. This supports the idea that miRNA processing by the Drosha complex is a relatively slow process, as previously suggested (Kim and Kim 2007; Morlando et al. 2008). On the other hand, we cannot exclude that some components of the Microprocessor complex might remain after cropping on the pre-mRNA, maintaining the exon and the DRE in close proximity. In this last case, MTS-mediated cleavage would not disrupt the interaction between the EDA exon and the DRE. Through the evaluation of the levels of pre-mRNA transcripts in cotransfection experiments, we observed that MTSs are processed with different efficiencies (Fig. 5). These differences could be important for pri-miRNA maturation and might be due to the relative affinity of pri-miRNAs for the Microprocessor complex, as reported to occur in vitro (Han et al. 2006) and/or to some specific regulatory mechanisms. Biosynthesis of a number of miRNAs is regulated at the level of pri-miRNA processing, with the involvement of RNA binding proteins. Several proteins interact with Drosha and/or with the pri-miRNA sequence and can either inhibit or promote the processing efficiency of individual miRNAs at the level of pri-mRNA cropping (Guil and Caceres 2007; Piskounova et al. 2008; Viswanathan et al. 2008; Trabucchi et al. 2009). As Drosha cleavage occurs cotranscriptionally (Kim and Kim 2007; Morlando et al. 2008; Pawlicki and Steitz 2008), it is conceivable that the regulation of biogenesis of some miRNAs at the level of Microprocessor processing occurs early on the nascent transcripts. The inefficient processing of two hairpins in the 5′ UTR and coding regions of the DGCR8 mRNAs might be related to the unique role of the Microprocessor in direct destabilization of a coding mRNA (Han et al. 2009; Shenoy and Blelloch 2009; Triboulet et al. 2009). In this case, it might be possible that the Microprocessor complex degrades the coding mRNA after intron splicing.

In conclusion, we show that efficient cleavage of nascent introns mediated by different Rzs partially reduces the amount of mature mRNA and, when inserted at critical positions between the exon and its regulatory element, affects alternative splicing. Intronic Microprocessor target sequences are cleaved inefficiently relative to Rzs and do not change the pattern of splicing but retain some individual residual activity that can be important for a regulated cotranscriptional cropping. The comparable effect of Rzs on mRNA levels, which is independent from the Rz type, the intronic distance from the neighboring exons, and exon definition, suggests that severed nascent transcripts are inefficiently processed to the mature mRNA. The different efficiencies in processing of severed nascent transcripts can be due to the contribution of some splicing factors like SR proteins. These proteins, involved in coupling RNA Pol II to pre-mRNA splicing (Das et al. 2007), might modulate Pol II interaction with exonic splicing enhancers and regulate exon tethering in a transcript-specific manner.

MATERIALS AND METHODS

Cell culture, transfection, and reverse transcription-PCR analysis

Hep3B and HeLa cells (5×10^6) were grown under standard conditions and were transiently transfected with 0.5 μg of plasmid DNA with the Effectene Transfection Reagent (QIAGEN) kit according to the manufacturer's instructions. Cells were harvested after 24 h, and RNA was isolated using TRIReagent (Ambion). RT-PCR was performed as previously described (Pagani et al. 2003; Baralle et al. 2008). For 5′-splicing intermediate and nascent pre-mRNAs analysis, total RNA was treated with Dnase–RNase-free and loaded on RNA cleanup columns (QIAGEN). Cotransfection experiments were performed with 0.15 μg of pEDA or pEDs, along with 0.45 μg of the indicated minigenes. Quantification of band intensities was performed using a fluorescent primer. Amplifications were performed in the linear range; the resulting products were run on a capillary electrophoresis with a fluorescent internal MW standard and analyzed with Peak Scanner software (Applied Biosystems). The intensity of bands in agarose gels has been determined with ImageJ software as previously reported (Goina et al. 2008).

siRNA transfections

siRNA transfections were performed in HeLa cells using Oligofectamine reagents (Invitrogen). The sense strand for Xrn2 RNA interference oligonucleotides (Dharmacon) used for silencing was 5′-aagaguacagaugaucauguu-3′. Silencing was performed as previously described with two rounds of siRNA transfections (Goina et al. 2008). Cells were harvested and divided in two parts for RNA and protein extractions. Protein extract was analyzed in Western Blotting for Xrn2 (a generous gift from N. Gromak) and tubulin as a control.

Quantitative RT-PCR

For expression level, each minigene (500 ng) was cotransfected with the control pCF1 (100 ng), which was used for normalization of transfection and reverse transcriptase efficiencies. Real-time PCR amplification of pCF1 was performed with Syber green supermix (Bio-Rad) with CFex1_248dir and CFex3_292rev that gave a spliced product of 233 bp. For quantitative RT-PCR on total mRNA, the minigene expression was detected with an HBA1 TaqMan gene expression assay (Applied Biosystems). For semiquantitative RT-PCR analysis of the 5′-splicing intermediates, minigenes were amplified with f1 and r1 primers loaded on agarose gel, and the band intensity was evaluated with the ImageJ software.

Quantification of hsa-mir-26b and hsa-mir-330 expression was performed with TaqMan MicroRNA Assays (Applied Biosystems).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This work was supported by the AIRC grant IG10387. We thank Natalia Gromak (Sir William Dunn School of Pathology, University of Oxford, Oxford, UK) for the Xrn2 antibody and Marcos Morgan (ICGEB, Trieste, Italy) for the pCMV26b plasmid. We thank also Rodolfo Garcia and F.E. Baralle for helpful discussions.

Received October 25, 2010; accepted May 2, 2011.

REFERENCES

Ashton-Beaucage D, Udell CM, Lavoie H, Baril C, Lefrancois M, Chagnon P, Gendron P, Caron-Lizotte O, Bonneil E, Thibault P, et al. 2010. The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in *Drosophila*. *Cell* **143:** 251–262.

Baralle M, Pastor T, Bussani E, Pagani F. 2008. Influence of Friedreich ataxia GAA noncoding repeat expansions on pre-mRNA processing. *Am J Hum Genet* **83:** 77–88.

Beyer AL, Osheim YN. 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2:** 754–765.

Beyer AL, Bouton AH, Miller OL Jr. 1981. Correlation of hnRNP structure and nascent transcript cleavage. *Cell* **26:** 155–165.

Chen X, Denison L, Levy M, Ellington AD. 2009. Direct selection for ribozyme cleavage activity in cells. *RNA* **15:** 2035–2045.

Das R, Dufu K, Romney B, Feldt M, Elenko M, Reed R. 2006. Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev* **20:** 1100–1109.

Das R, Yu J, Zhang Z, Gygi MP, Krainer AR, Gygi SP, Reed R. 2007. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell* **26:** 867–881.

de la Mata M, Lafaille C, Kornblihtt AR. 2010. First come, first served revisited: Factors affecting the same alternative splicing event have different effects on the relative rates of intron removal. *RNA* **16:** 904–912.

Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432:** 231–235.

Dye MJ, Gromak N, Proudfoot NJ. 2006. Exon tethering in transcription by RNA polymerase II. *Mol Cell* **21:** 849–859.

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

Goina E, Skoko N, Pagani F. 2008. Binding of DAZAP1 and hnRNPA1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant. *Mol Cell Biol* **28:** 3850–3860.

Gornemann J, Kotovic KM, Hujer K, Neugebauer KM. 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol Cell* **19:** 53–63.

Gromak N, Talotti G, Proudfoot NJ, Pagani F. 2008. Modulating alternative splicing by cotranscriptional cleavage of nascent intronic RNA. *RNA* **14:** 359–366.

Guil S, Caceres JF. 2007. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* **14:** 591–596.

Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT, Kim VN. 2006. Molecular basis for the recognition of primary microRNAs by the Drosha–DGCR8 complex. *Cell* **125:** 887–901.

Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, Yeom KH, Yang WY, Haussler D, Blelloch R, Kim VN. 2009. Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* **136:** 75–84.

Hicks MJ, Yang CR, Kotlajich MV, Hertel KJ. 2006. Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. *PLoS Biol* **4:** e147. doi: 10.1371/journal.pbio.0040147.

Kim YK, Kim VN. 2007. Processing of intronic microRNAs. *EMBO J* **26:** 775–783.

Kornblihtt AR. 2005. Promoter usage and alternative splicing. *Curr Opin Cell Biol* **17:** 262–268.

Kornblihtt AR. 2007. Coupling transcription and alternative splicing. *Adv Exp Med Biol* **623:** 175–189.

Lacadie SA, Tardiff DF, Kadener S, Rosbash M. 2006. In vivo commitment to yeast cotranscriptional splicing is sensitive to transcription elongation mutants. *Genes Dev* **20:** 2055–2066.

Listerman I, Sapra AK, Neugebauer KM. 2006. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol* **13:** 815–822.

Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. 2010. Regulation of alternative splicing by histone modifications. *Science* **327:** 996–1000.

Maniatis T, Reed R. 2002. An extensive network of coupling among gene expression machines. *Nature* **416:** 499–506.

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385:** 357–361.

Misteli T, Spector DL. 1999. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol Cell* **3:** 697–705.

Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* **15:** 902–909.

Morris DP, Greenleaf AL. 2000. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **275:** 39935–39943.

Natalizio BJ, Robson-Dixon ND, Garcia-Blanco MA. 2009. The carboxyl-terminal domain of RNA polymerase II is not sufficient to enhance the efficiency of pre-mRNA capping or splicing in the context of a different polymerase. *J Biol Chem* **284:** 8692–8702.

Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, Baralle FE. 2003. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Hum Mol Genet* **12:** 1111–1120.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40:** 1413–1415.

Pandya-Jones A, Black DL. 2009. Co-transcriptional splicing of constitutive and alternative exons. *RNA* **15:** 1896–1908.

Pawlicki JM, Steitz JA. 2008. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* **182:** 61–76.

Piskounova E, Viswanathan SR, Janas M, LaPierre RJ, Daley GQ, Sliz P, Gregory RI. 2008. Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. *J Biol Chem* **283:** 21310–21314.

Roignant JY, Treisman JE. 2010. Exon junction complex subunits are required to splice *Drosophila* MAP kinase, a large heterochromatic gene. *Cell* **143:** 238–250.

Rosonina E, Ip JY, Calarco JA, Bakowski MA, Emili A, McCracken S, Tucker P, Ingles CJ, Blencowe BJ. 2005. Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Mol Cell Biol* **25:** 6734–6746.

Shenoy A, Blelloch R. 2009. Genomic analysis suggests that mRNA destabilization by the Microprocessor is specialized for the auto-regulation of Dgcr8. *PLoS ONE* **4:** e6971. doi: 10.1371/journal.pone.0006971.

Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A, Gherzi R, Rosenfeld MG. 2009. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* **459:** 1010–1014.

Triboulet R, Chang HM, Lapierre RJ, Gregory RI. 2009. Post-transcriptional control of DGCR8 expression by the Microprocessor. *RNA* **15:** 1005–1011.

Viswanathan SR, Daley GQ, Gregory RI. 2008. Selective blockade of microRNA processing by Lin28. *Science* **320:** 97–100.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456:** 470–476.

Yen L, Svendsen J, Lee JS, Gray JT, Magnier M, Baba T, D'Amato RJ, Mulligan RC. 2004. Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* **431:** 471–476.

RINGRAZIAMENTI

Solo ultimamente mi sono reso realmente conto di quanto siano importanti, da un punto di vista sia scientifico, sia umano, gli anni di dottorato. Ti insegnano che non sei mai abbastanza aggiornato, abbastanza al passo, ti spingono a informarti, a studiare, a documentarti. Ti insegnano ad essere una persona critica e – auspicabilmente- un buon futuro ricercatore. Ti insegnano ad entusiasmarti per i piccoli traguardi e a non abbatterti per qualche sconfitta.

Il mio primo ringraziamento va quindi a ***Franco*** *(*aka *the boss o FP)* che in momenti e in modi diversi mi ha spinto a essere curioso, a provare nuove strade, a seguire le mie idee lasciandomi anche a volte sbagliare. Non è stato sempre facile, ma siamo sopravvissuti entrambi!
Grazie a tutti i componenti passati e presenti dello *Human Molecular Genetics Group* per l'aiuto e il supporto, tecnico e morale, per il clima familiare e di collaborazione che c'è sempre stato in lab. Grazie a ***Gosia*** (combattiva collega polacca e soprattutto amica)*,* ***Erica, Elisa, Tibor, Gabriele, Chiara, Giulia, Irving, Mojca, Luca e Francesca****.*
Grazie alle ***Giulie*** e alla new entry *Nicoletta* per le pause pranzo goliardiche, per la leggerezza che spesso ha alleviato delle giornate no.

Ci sono persone che nella vita incontri, per caso, e sai che non perderai mai. Persone con le quali non ti vergogni a condividere nulla, persone con cui puoi stare bene anche in silenzio e alle quali basta uno sguardo per capire come ti senti, sapendo di non venire mai giudicato. ***Linda***, questo grazie è per te. Ed è solo l'ennesimo di una serie di grazie che dovrò farti anche in futuro.

Roy, imprescindibile presenza, amico attento e sempre disponibile. Grazie, per tutto e per i momenti in cui ci sei stato senza nemmeno bisogno che te lo chiedessi.

Grazie ***Vale***. Avrei così tanti "grazie" da dirti e per così tante cose che non basterebbe una pagina intera. Fortuna vuole che noi due ci capiamo e ci capiremo sempre al volo, senza bisogno d'altro che di uno sguardo. O di un sms con una scimmietta.

Grazie a ***voi***. Mamma e papà. Perché ci siete sempre stati, e so che sempre ci sarete, a prescindere dalla lontananza. Grazie per avermi aiutato, grazie per non avermi dato sempre ragione, grazie per avermi fatto diventare la persona che sono contento di essere. Spero di avervi resi, almeno un po', fieri di me. E spero di continuare a farlo in futuro. Il mio abbraccio più grande e il mio grazie più sentito sono per *voi.*

Infine, chiudo con un augurio a me stesso, ricalcando il titolo di una celebre commedia italiana: ***"Io… speriamo che me la cavo!"*.**