Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
"BIOCHIMICA, BIOLOGIA MOLECOLARE E BIOTECNOLOGIE"

CICLO XXVIII
COORDINATORE Prof. Francesco Bernardi

REGULATION OF EXON DEFINITION BY INTRINSIC ELEMENTS AND BY COMBINATION OF TAILORED U1snRNA WITH ANTISENSE OLIGONUCLEOTIDES

Settore Scientifico Disciplinare BIO/10

Dottorando
Dott. Scalet Daniela

_____

Tutore
Prof. Pinotti Mirko

_____

Co-Tutore
Dott. Balestra Dario

_____

Anni 2013/2015

"There is no diseases so rare that it does not deserve attention"

(Orphanet, the portal for rare diseases and orphan drugs)

ABSTRACT

The notion that a significant proportion of disease-causing mutations affects pre-mRNA boosted research towards the design of approaches modulating this very complex process for therapeutic purposes.

This PhD project has been focused on the development of RNA-based correction approaches for Hemophilia B (HB) or Seckel syndrome, hemorrhagic or neurologic disorders due to mutations in the *F9* or *ATR* genes, respectively.

Previous studies by our research group on a panel of HB-causing mutations led to the development of variants of the U1 small nuclear RNA (U1snRNA), key component of the spliceosomal U1 small nuclear ribonucleoprotein, which can rescue exon inclusion by targeting intronic sequences downstream of a defective exon (Exon-specific U1snRNA, ExSpeU1). Most importantly, it has been proven in cellular models that a unique ExSpeU1fx9 can rescue aberrant splicing caused by multiple exon-skipping mutations occurring at the donor (5'ss) or acceptor (3'ss) splice sites of the *F9* exon 5. Here, by expressing the human *F9* splicing-defecting expression cassettes in mouse liver, we provided the first *in vivo* proof-of-concept that the selected ExSpeU1fix9 is able to restore *F9* splicing and remarkably increase human factor IX (FIX) protein levels and coagulant activity in plasma in the presence of two model exon-skipping mutations at the 5'ss (c.519A>G) or 3'ss (c.392-8T>G)

In the model of F9 exon 2, through the expression of *F9* minigenes in mammalian cells, we characterized numerous HB-causing mutations, either missense or at the 5'ss , that promote aberrant splicing by inducing the usage of a strong exonic cryptic 5'ss. Splicing assays with natural and artificial *F9* variants indicated that the cryptic 5'ss is regulated, among a network of regulatory elements, by an exonic splicing silencer (ESS). This finding supports a compensatory mechanism aimed at minimizing unproductive splicing. To recover splicing we tested antisense oligo-ribonucleotides (AON) masking the cryptic 5'ss, which were effective on exonic changes but promoted exon 2 skipping in the presence of mutations at the authentic 5'ss. On the other hand, we observed a very poor correction effect by U1snRNA variants with increased or perfect complementarity to the defective 5'ss. Noticeably, the combination of the mutant-specific U1snRNAs with antisense oligonucleotides produced appreciable amounts of correctly spliced transcripts from several mutants of the exon 2 5'ss.

In the Seckel Syndrome (SS-1) model we characterized the synonymous c.A2101G change in exon 9 of the *ATR* gene that induces exon 9 skipping by creating an exonic splicing silencer (ESS) in the poorly defined exon. Based on this mechanism, we explored two complementary correction strategies based on AON, designed to mask the ESS and modified U1snRNA, improving exon 9 definition. In ATR minigene assays demonstrated that both the AON and the U1snRNA^{ATR} induced a robust exon 9 inclusion raising from 6% to 100% and 63%, respectively. The U1snRNA^{ATR} was then challenged in the embryonic fibroblasts from the humanized SS-1 mouse model (MEF^{SS-1}) harboring the SS-1 splicing mutation. The lentiviral-mediated delivery of the U1snRNA^{ATR} in MEF^{SS-1} resulted in partial rescue of exon 9 inclusion and in the low but appreciable increase of ATR protein expression.

Taken together these data in the hemophilia B and Seckel syndrome models demonstrate the ability of appropriately designed RNA-based approaches to counteract splicing mutations and rescue gene expression, thus encouraging their exploitation for the development of innovative therapies for genetic disorders.

Table of contents

I

RNA-based therapeutic approaches

I.1 Pre mRNA splicing

Splicing reaction

Nascent mRNAs undergo several modifications and are subjected to diverse regulatory controls before being exported from the nucleus to the cytoplasm as functional, mature messenger RNAs (mRNAs). RNA processing is carried out under the guidance of large macromolecular processing complexes which rather than acting separately are in intimate contact with each other, integrating all these functions and thus establishing genuine gene expression factories (Maniatis and Reed 2002). As soon as transcription begins, the first nucleotide of the precursor messenger RNA (pre-mRNA) is linked to a G through a 5'-5'-triphosphate bond and further methylated to create a N7-methylguanosine cap (m1G-cap) (Shatkin 1976; Babich et al. 1980; Coppola et al. 1983). At the other edge of the pre-mRNA, a string of adenine ribonucleotides s is attached upon cleavage of its 3'-end by the polyadenylation machinery components (Proudfoot 1989). These modifications stabilize the mRNA, prevent degradation by exonucleases, ensuring proper mRNA export and promote translation. However, none of the processes defines the final coding mRNA sequence. The information necessary for the synthesis of proteins is scattered across the human genome, with the coding segments (exons) being a minor proportion of the genetic information, accounting for no more than 1% of the entire genome (Lander, Linton et al. 2001). Therefore the majority of the protein-coding genes are composed mostly by non-coding, intervening sequences (introns). Within the pre-mRNA, the shorter and less abundant exons must be identified, defined, and assembled in a mRNA transcript which encodes ultimately the whole protein sequence whereas the introns are removed. This cut and paste reaction is called splicing. Recent findings indicate that almost every multiexon gene in the human genome undergo at least one alternative splicing event highlighting the central role that splicing plays in gene expression (Pan et al. 2008; Wang et al. 2008). Furthermore, splicing is a very-well conserved pre-mRNA processing mechanism, found from unicellular eukaryotic organisms as Saccharomyces cereviseae to metazoans. It displays increasing levels of regulation and complexity as the number and length of introns in multicellular eukaryotes increases (Ast 2004; Nilsen and Graveley 2010). Thus, splicing constitutes not only a crucial step for accurate transfer of the genetic information from DNA to RNA to protein, but also a step that allows for regulation of gene expression as well as increased protein diversity through alternative splicing decisions. In order to find the short exons among the sea of intronic regions some signals present at the exon/intron

boundaries are crucial. In higher eukaryotes, these elements are short consensus sequences surrounding the 3' and 5' end of the introns, which are known as 3' and 5' splice sites (ss), respectively. Their sequences exhibit a variable degree of conservation nonetheless they are fundamental for proper intron recognition and splicing catalysis (Figure I.1).

5' ss or donor splice site. The 5' ss marks the exon/intron junction at the 5' end of the intron and its sequence consensus is composed of 9 bp, located on both sides of the exon/intron boundary: 3 bases on the exonic side, and 6 on the intronic side. The 5' ss consensus sequence have been established long ago to be MAG|GURAGU (M indicates A or C; R indicates purines and the | the exon/intron boundary) (Shapiro and Senapathy 1987). The underlined GU dinucleotide is almost universally conserved as it is found in more of 98% of human donor splice sites (Sheth et al. 2006). They are critical for the splicing reaction as when one of these two nucleotides are mutated splicing is abolished or blocked at intermediate steps (Aebi et al. 1987; Lamond et al. 1987; Chanfreau and Jacquier 1993). The remaining nucleotides positions display variable conservation, with some bases at certain positions being more conserved than others, likely reflecting their different role on the splicing reaction (Carmel et al. 2004; Roca et al. 2008). Nevertheless the entire consensus donor splice site determines the 5' cleavage site, rather than the invariant GU dinucleotide (Aebi et al. 1987). Recognition of the 5' ss involves a nearly perfect base-pairing with the 5'-tail of U1 snRNA (Horowitz and Krainer 1994) and guides the early assembly of the spliceosome machinery upon the intron. However a minority of 5' ss (<1%) has a GC dinucleotide at the intron/exon boundary, defining a GC intron (Sahashi et al. 2007).

3' ss, a composite signal. The intronic element that identifies the 3' ss usually appears thousands bases downstream of the 5' ss. It is composed by three different moderately conserved elements: the branch point (BP), the polypyrimidine tract (PPT) and the terminal conserved AG dinucleotide (Reed 1989). The BP is characterized by the presence of a conserved A surrounded by a highly degenerated motif YNYURAY (Y=pyrimidine and R=purine)(Reed and Maniatis 1985). It is commonly found about 18-40 nucleotides upstream of the AG dinucleotide (Ruskinet al. 1984; Reed and Maniatis 1985) although some exceptions can be found hundreds of nucleotides away (Reed 1989). The recognition of the branch site involves a base-pairing with the U2 snRNP in order to form the spliceosome A complex (Berglund et al. 1997).The PPT is a run of pyrimidines (eight

bases in the average intron) located between the branch site and the terminal AG at the intron/exon junction (Reed 1989). It can display variable pyrimidine content, length and distance to the branch-point and the AG. The PPT is essential for efficient branch-point utilization and correct AG recognition as it has been shown that progressive deletion of the polypyrimidine tract impairs splicing while elongating its length can improve its efficiency (Roscigno et al. 1993). The terminal AG dinucleotide or acceptor site defines the 3' border of the intron. This site is characterized by the short YAG/G sequence (Y pyrimidines; the slash is the intron-exon boundary and the underlined nucleotides are conserved). Even if it is essential for the second step of splicing reaction (see below) no base-pairing interactions with snRNAs are involved in recognizing this sequence (Wu et al. 1999). The sequences between the branch-point and the acceptor site are commonly devoid of AG dinucleotides (Gooding et al. 2006).

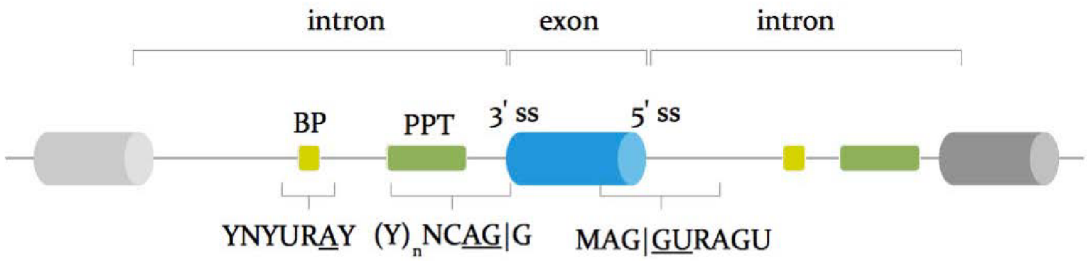


Figure I.1 Schematic representation of exon-intron boundaries
The cylinders represent the exons and the grey line depict the introns. The branch--point (BP) and the polypyrimidine tract (PPT) are illustrated by the rectangles (yellow and green, respectively). Below the consensus sequences for each canonical signal are shown. The nearly invariant GU, AG at 5' and 3' ss, respectively and the conserved A at the branch-point (BP) are underlined. The polypyrimidine tract (PPT) is also represented. Y, polypyrimidines (U or C); R, purines (A or G); M, A or C.; N, any nucleotide.

I.2 The spliceosome: Assembly of Spliceosomal snRNPs

Structure of Spliceosomal snRNPs:

The removal of introns requires the five major spliceosomal small nuclear ribonucleoprotein particles (snRNPs; U1, U2, U4, U5 and U6). Each snRNP consists of a uridylic acid-rich small nuclear RNA (U1, U2, U4, U5 and U6 snRNAs) that is post-transcriptionally modified and a cortege of associated proteins (Fabrizio et al. 1994; Nagai et al. 2001; Jurica and Moore 2003; Stanek and Neugebauer 2006). The 2,2,7-trimethyl guanosine (m_3G) capped U1, U2, U4 and U5 snRNAs (Sm snRNAs) contain an Sm site ($RAU_{3-6}GR$, where R is a purine) flanked by stem-loops, which collectively constitute domain A (Branlant et al. 1982). Sm proteins (B/B′, D1, D2, D3, E, F and G) assemble into a heteroheptameric ring around the Sm site to form the core of the snRNP particle. Similarly, the γ-methyl triphosphate (γ-m-P_3) capped U6 snRNA acquires a heteroheptameric ring of LSm proteins (Like Sm). The LSm proteins (LSm2-8) assemble around the U6 snRNA 3′-terminus, which consists of a uridine tract ending in a 2′, 3′-cyclic phosphate (U_4-2′, 3′cP_i) (Singh and Reddy 1989; Achsel et al. 1999; Mayes et al. 1999; Vidal et al. 1999). Proteins of the L/Sm lineage share an ancient signature motif, the Sm fold. In addition to the core proteins, each snRNP is decorated with an ensemble of proteins unique to a given snRNP, the snRNP-specific proteins (Stanek and Neugebauer 2006).

The mono-snRNPs just described do not represent their *in vivo* functional forms; rather, they are organized into higher order particles. The U4, U5 and U6 snRNPs exist largely in their functional form as a U4/U6. U5 tri-snRNP (Behrens and Luhrmann 1991; Liu et al. 2006). The same holds true of the U4atac, U5 and U6atac snRNPs, which form the minor spliceosomal U4atac/U6atac.U5 tri-snRNP (Schneider et al. 2002). Interestingly, minor spliceosomal counterparts of the U1 and U2 snRNPs, the U11 and U12 snRNPs, respectively, are known to assemble into the minor spliceosomal U11/U12 di-snRNP. Furthermore, penta-snRNP complexes, which consist of all five major splicing snRNPs and may represent a `splicing holoenzyme', have been shown to exist in both yeast and humans (Nilsen 2002; Stevens et al. 2002; Malca et al. 2003) (Figure I.2). The snRNPs, along with over 300 other splicing factors, assemble onto pre-mRNA to form the spliceosome, and it is this dynamic macromolecular machine that orchestrates the excision of introns and the ligation of exons through two successive transesterification reactions (Patel and Steitz 2003). Prior to participating in splicing, however, snRNPs must be

assembled through a series of intricate steps that, in all organisms, begins in the nuclear compartment. In animals, protista and plantae, a brief transit to the cytoplasm is essential for the assembly of Sm snRNPs, but the assembly of the U6 snRNP is uninterrupted by a cytoplasmic phase.

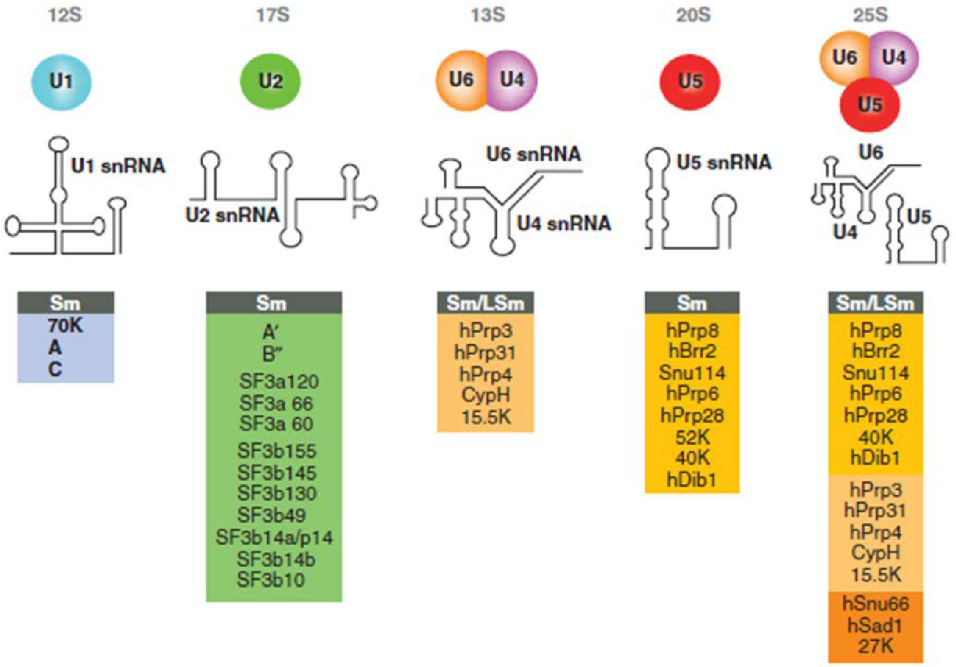


Figure I.2 Protein composition and snRNA secondary structures of the major human spliceosomal snRNPs. All seven Sm proteins (B/B', D3, D2, D1, E, F, and G) or LSm proteins (Lsm2-8) are indicated by "Sm" or "LSm" at the top of the boxes showing the proteins associated with each snRNP. The U4/U6.U5 tri-snRNP contains two sets of Sm proteins and one set of LSm proteins.

The Biogenesis of Spliceosomal snRNPs

The assembly of all spliceosomal snRNPs begins with the transcription of a U snRNA. The genes for the U snRNAs reside in the nuclear genome and are transcribed by either RNA polymerase (RNAP) II or III. During evolution, multiple copies (20-100) of U1, U2, U4 and U5 genes have arisen by gene duplication; however, the U6 gene is present only in approximately five functional copies in the haploid human genome (Lund and Dahlberg 1984; Van Arsdell and Weiner 1984; Domitrovich and Kunkel 2003). Major clusters of the human U1 and U2 genes are present on chromosomes 1 and 17, respectively (Lund and Dahlberg 1984; Van Arsdell and Weiner 1984); whereas, the U6 genes are scattered throughout the genome (Domitrovich and Kunkel 2003). It is clear that the assembly of Sm snRNPs and the U6 snRNP follow two distinct pathways. The Sm snRNAs are transcribed

as 3′-extended (2- to 10-nt longer) precursors by RNAPII, and like all other RNAPII transcripts, they co-transcriptionally acquire a 5′- to 5′-linked N7-methyl guanosine (m_1G) cap. The 3′-ends of the pre-snRNAs are generated by an RNA processing event that is coupled to the proximal sequence element-directed transcription. A conserved 3′-box ($GTTTN_{0–3}AAAPuNNAGA$, where Pu = purine and N = any nucleotide) marks the cleavage site which resides ~10-nt upstream and a heterododecameric metallo β-lactamase complex (Integrator) contains the enzymatic activity for 3′-end formation (Baillat et al. 2005). The newly transcribed pre-snRNAs must be transported to the cytoplasm to continue their maturation, necessitating the assembly of an export competent complex. First of all, the nuclear CBC, consisting of CBP20 and CBP80, first associates with the m_1G cap of the RNA (Izaurralde et al. 1992; Izaurralde et al. 1995). Next, the phosphorylated adaptor for RNA export (PHAX) binds the CBC–RNA complex. The export receptor, Exportin 1/Chromosome Region Maintenance 1 (Xpo1/CRM1), recognizes the export adaptor, PHAX, in its phosphorylated form bound to its CBC/pre-snRNA cargo and binds to this complex together with RanGTP (Segref et al. 2001). While all of the above interactions are individually quite weak, cooperative binding ensures the formation of a stable export complex. It was shown recently that the assembly of an export competent mRNA begins at the transcriptional unit (Kohler and Hurt 2007; Patel et al. 2007). After assembly, the entire complex translocates through the nuclear pore complex (NPC). The transcription of nascent snRNAs, 3′-end formation, assembly into an export complex, putative transit through CBs and translocation through the NPC are rapid processes, as revealed by radiolabeled experiments, requiring only 4 minutes (Eliceiri and Sayavedra 1976; Eliceiri and Gurney 1978). Upon entry in the cytoplasm, PHAX is dephosphorylated by protein phosphatase 2A, but remains associated with the CBC/pre-snRNA complex, presumably until the m_1G cap is hypermethylated (Ohno et al. 2000;). CBC and PHAX are subsequently recycled to the nucleus, where the latter is phosphorylated by casein kinase 2 (CK2) to initiate another round of pre-snRNA export. The cytoplasmic phase of snRNP maturation is orchestrated by a large 20S assemblysome called the survival of motor neuron protein (SMN) complex, which consists of the SMN, seven distinct Gemin proteins (Gemin 2-8), and several other protein factors (Battle et al. 2006; Kolb et al. 2007). The SMN complex participates in all three snRNP maturation events in the cytoplasm: (i) the assembly of an Sm ring onto the Sm site; (ii) the hypermethylation of the m_1G cap; and (iii) the trimming of the pre-snRNA's 3′-end. Indeed, the SMN complex associates with a distinct set of snRNP populations, each

representing different stages in their cytoplasmic maturation: a disassembled export complex, the core Sm snRNP and an import complex (Massenet et al. 2002) (Figure I.3).

(i) The assembly of the core snRNP begins with the formation of the Sm ring around the Sm site. Although the Sm proteins do not form rings in the absence of the snRNA, they exist as dimers (B/B′-D3, D1-D2) or trimers (E-F-G) (Raker et al. 1996). First, the SMN complex facilitates the formation of a semi-stable open ring complex consisting of D1-D2-E-F-G proteins around the Sm site of pre-U snRNAs. Then, the SMN complex completes the formation of a 7-membered ring (-D3-B/B′-D1-D2-E-F-G-) upon integration of the B/B′-D3 heterodimer (Raker et al. 1996). The SMN complex likely serves as a specificity factor—in addition to an assembly factor—that ensures the assembly of the Sm ring only on RNAs with the appropriate snRNP code (Pellizzoni et al. 2002). The WD repeat containing subunit of the SMN complex, Gemin 5, recognizes this code on the snRNA, which consists of the Sm site and parts of the adjacent stem-loop structure(s) (Kasim et al. 2006). The U1 snRNA is distinct in that its code consists of stem-loop I (SL1) (Yong et al. 2002). SL1, however, is not a strict requirement as an SL1-deleted U1 snRNA still acquires its Sm complement and is recruited to the nucleus.

(ii) The hypermethylation of the m_1G cap occurs after the assembly of the Sm ring. Trimethyl guanosine synthase 1 (Tgs1), an SMN complex-associated methyltransferase, recognizes SmB/B′ in the context of an Sm core as well as the m_1G cap on the snRNA and subsequently transfers two methyl groups to position 2 of the m_1G cap forming the m_3G cap (Plessel et al. 1994; Mouaikel et al. 2003). Since the addition of B/B′-D3 heterodimer completes the assembly of the Sm ring, the association of Tgs1 to B/B′ ensures that only snRNAs with fully assembled Sm rings are hypermethylated.

(iii) Nucleolytic trimming of the 3′-end of the pre-snRNA generates the mature length snRNA.

While the molecular mechanisms regulating the cytoplasmic maturation events of snRNAs have been extensively studied, the spatial arrangement of these events within the cytoplasm remains poorly understood. A recent study suggests that snRNP maturation might partly occur in discrete cytoplasmic bodies (Liu and Gall 2007). These organelles were named the `U bodies' because they contain the major U snRNPs.

The core snRNP must be brought into the nucleus to continue its maturation and, afterwards, participate in pre-mRNA splicing. The requirements for nuclear import vary depending on the particular snRNP and on the cell system. In general, the m_3G cap and the Sm core are considered to be nuclear localization signals (NLS) that utilize the same import receptor importin β (Imp β) but distinct import adaptors, snurportin-1 (SPN1) and possibly SMN, respectively. The nuclear phase is the least understood part of the entire snRNP biogenesis pathway, and it involves a multiplicity of processes and factors, as well as trafficking to several subnuclear domains. In particular, extensive internal modifications of the U snRNAs by 2′-*O*-methylation and pseudouridylation represent a critical step in the making of a fully functional snRNP.

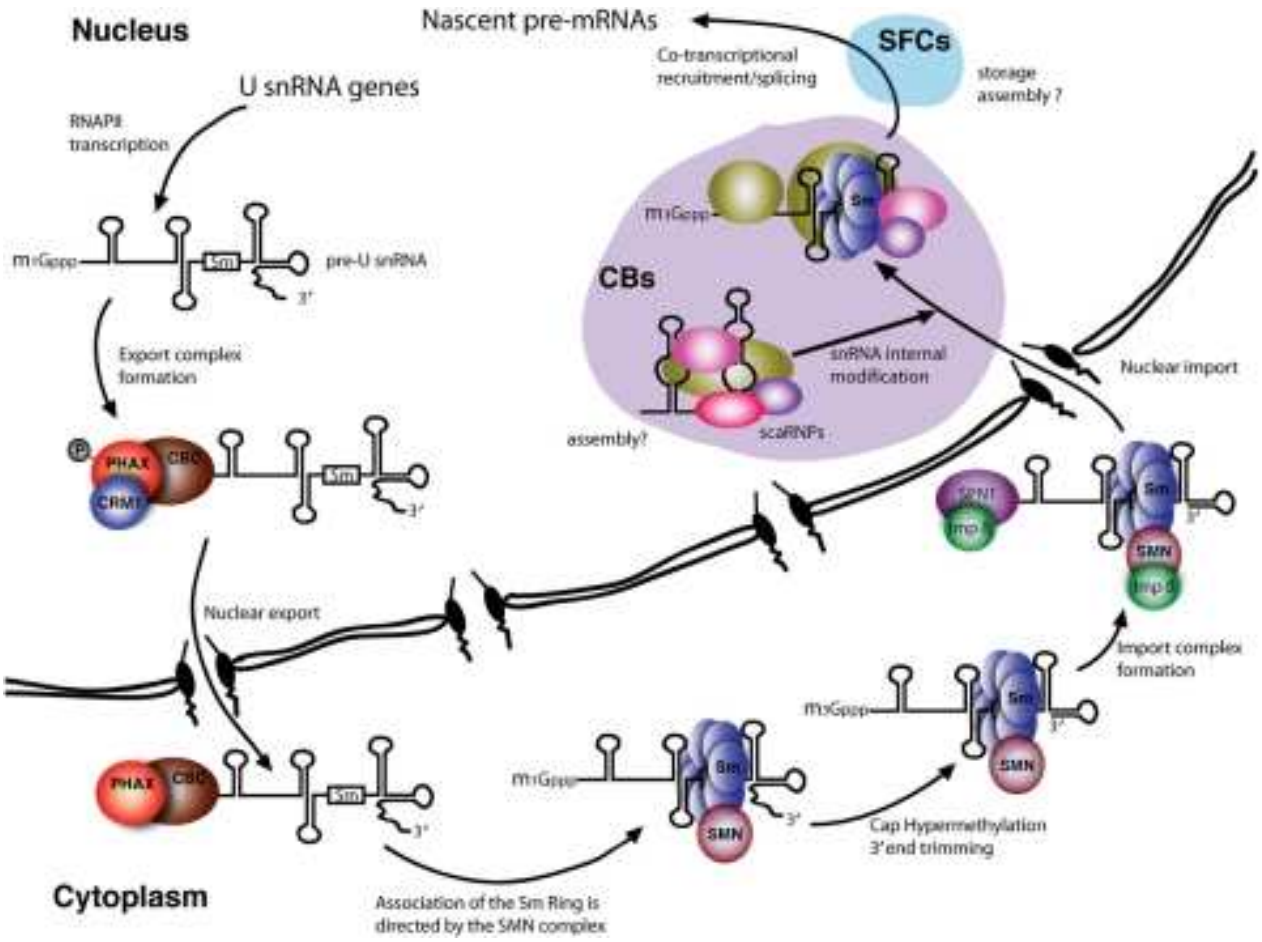


Figure I.3 The Sm snRNP assembly and maturation pathway. The U2 snRNA was used here as a representative member of the Sm snRNAs. While discussed in the text, the U bodies are not shown here because of a lack of understanding of their role(s) in snRNP maturation, if any. scaRNP = small Cajal body specific ribonucleoprotein particle.

Splicing reaction occur in a stepwise manner

Spliceosomal snRNPs have a critical role in the recognition of correct splice sites within a multitude of similar sequences. The production of a spliced, mature mRNA requires extensive specific and dynamic interactions of different nature, such as RNA-RNA base pairing, RNA-protein and protein-protein binding and many structural changes. Through biochemical assays, distinct intermediate complexes of the splicing reaction have been detected and thoroughly studied. Thus it became evident that the spliceosome assembly occurs in a stepwise manner, involving assembly/disassembly of different snRNP particles and non-snRNP splicing factor on the pre-mRNA (Bringmann and Luhrmann 1986; Bindereif and Green 1987; Jamison and Garcia-Blanco 1992; Hong et al. 1997; Das and Reed 1999; Kent et al. 2005; Tardiff and Rosbash 2006). The assembly of the spliceosome (E –*early*- complex or commitment complex) begins with the recognition of the 5' ss sequence by the U1 snRNP through its 5'-tail (Rossi et al. 1996; Will et al. 1996). U1 snRNP-associated proteins U1-70k and U1C stabilize this transient interaction. Another important step following the U1 snRNP-5' ss recognition is the recognition of the 3' ss: the U2 Auxiliary Factor (which is an heterodimer made of U2AF65/35) identifies the AG dinucleotide at the intron/exon junction together with the PPT and SF1/mBBP protein binds at the BP site. Mutual stabilization of contacts with the U2AF bound to the 3' ss and the downstream U1 snRNP at the 5' ss can be mediated by members of the serine/arginine-rich (SR) protein family. The establishment of multiple weak interactions from the 3' ss to the 5' ss defines an exon, and constitutes the commitment step towards the splicing pathway (Robberson et al. 1990; Berget 1995). Subsequent to E complex formation, the A (prespliceosome) complex is built. The recruitment of U2 snRNP to the BP site, in an ATP-dependent fashion, with the concomitant displacement of SF1/mBBP from the BP site (Hong, Bennett et al. 1997). U2 snRNP base pairing to the BP sequence is facilitated by the U2AF65 subunit bound at the PPT (Ruskin et al. 1988). This base-pair interaction is further stabilized by heteromeric complexes of the U2 snRNP, namely SF3a and SF3b (Gozani et al. 1996). The transition from A to B complex are marked by the ATP-dependent addition of the U4/U6 and U5 snRNP, preassembled in U4/U6.U5 tri-snRNP. At this level all snRNPs are present, but the spliceosome is catalytically inactive and requires a conformational and compositional rearrangement to became active and promoting the first transesterification step of splicing.

During spliceosome activation, U1 and U4 are destabilized or removed, leading to a B* complex (B activated complex) (Turner et al. 2004). Eight evolutionarily conserved DExD/H-type RNA-dependent ATPase/helicases act at specific steps of the splicing cycle to catalyzed RNA-RNA rearrangements and RNP remodelling events (Valadkhan et al. 2009).

The C complex is then formed, and the spliceosome undergoes the first catalytic step. Subsequently additional rearrangements in RNPs network are necessary prior to undergone to the second transesterification reaction (Wahl et al. 2009). When also the second catalytic reaction has occurred, U2, U5 and U6 are released and these snRNPs are recycled for additional rounds of splicing. Spliceosome assembly and function also appear to be regulated by protein kinases and phosphatases. Phosphorylation/desphosphorylation cycles of constitutive and alternative splicing factors have been observed during assembly and catalytic steps (Tazi et al. 1993; Shi and Manley 2007; Stamm 2008; Heyd and Lynch 2010; Tripathi et al. 2010) (Figure I.4).

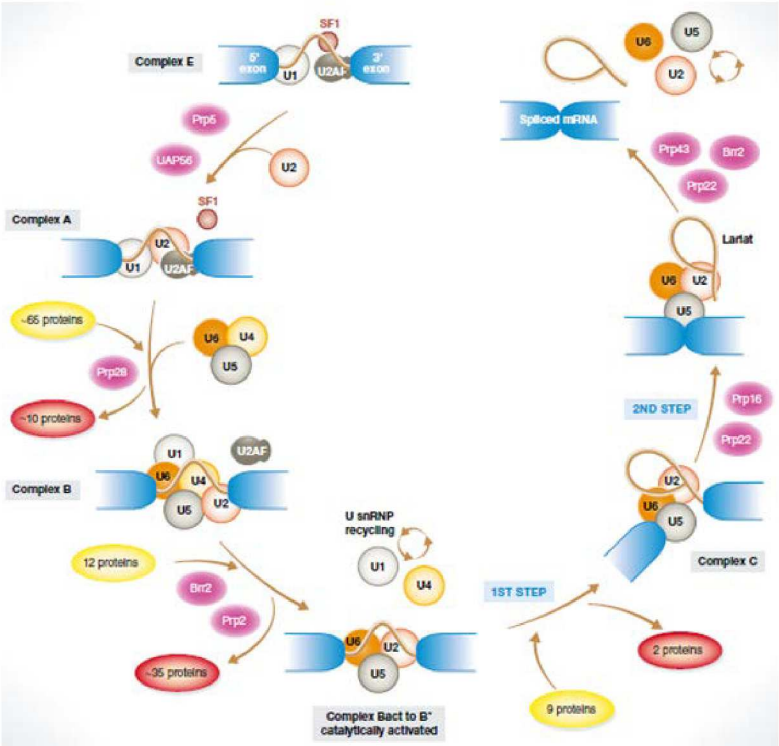


Figure I.4 A schematic view of spliceosome assembly

U1 small nuclear RNA (U1 snRNA)

In this thesis we have used U1 small nuclear RNA (U1 snRNA) as a key player in the RNA base approaches. Human cells contain about one million copies of this molecule. This 164-nucleotide long RNA is synthesized by RNA polymerase II and contains an unusual trimethylguanosine cap at the 5' end but no poly(A) at the 3' end. Human U1 RNA is encoded by a multigene family. By now, at least seven different U1 RNA genes have been cloned from human cell DNA, and hybridization analysis indicates that the total number of genes could range around 30 copies of true genes/haploid genome and of 125 copies of pseugenes/haploid genome equivalent. On the assumption that there are 30 true genes for U1 RNA synthesis/human haploid genome equivalent and that 10^6 molecules of U1 RNA are made in each 16-h generation of a rapidly dividing cell, there must be an initiation event on each of these genes approximately once every 4 s.

The true genes of U1snRNA, considering the high level of transcription, are found to be not methylated to any appreciable extent, while pseudogenes tend to be more highly methylated than true genes. The human U1 snRNA gene (RNU1-1) is repeated many times in the human genome as clustered repeat units of 45 Kb on chromosome 1 (Lund and Dahlberg 1984; Bernstein et al. 1985). These genes have a short TATA-less promoter, composed by a distal sequence element (DSE) that serves as a transcription enhancer and a proximal sequence element (PSE), located in the core promoter region upstream from the transcription start site. The PSE sequence is common to all human snRNA genes whereas the absence of a TATA box specifies the recruitment of RNA pol II and its associated transcription apparatus. The PSE is recognized by the snRNA activating protein complex, SNAPc (snRNA activator protein complex)(Sadowski et al. 1993), which serves as a target for transcription activators and repressors, such as Oct-1 ,p53 and RB (Ford et al. 1998; Hirsch et al. 2000; Gridasova and Henry 2005). SNAPc binding to the U1 PSE is necessary for the recruitment of general transcription factors such as TBP (transcription factor IIB), TFIIA, TFIIB, TFIIE and TFIIF (Kuhlman et al. 1999)(Figure I.5). The DSE is typically located 200 bp upstream of the transcription start site, and seems to be necessary for high-level expression of snRNA. It is a compound element, having an octamer motif recognized by the activators Oct-1 and Staf/SBF (Carbon et al. 1987; Schaub et al. 1997).

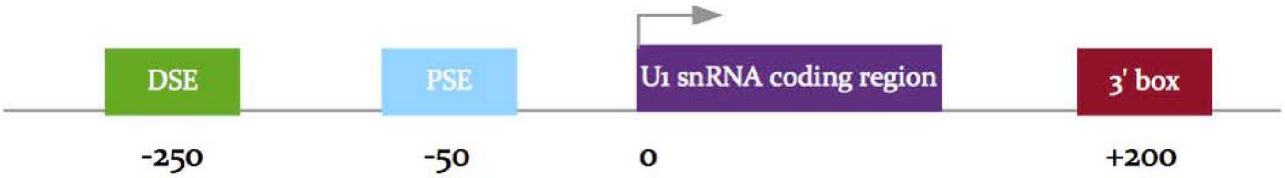


Figure I.5 The structure of human U1 snRNA gene transcribed by RNA pol II.
The diagram shows the DSE and PSE cis--acting promoter elements and the 3'box cis--acting RNA-processing element of pol II--transcribed U1--snRNA gene boxed, with their position relative to the transcription start site noted below. The start site of transcription is marked with an arrow above the line.

The transcriptional stimulatory activity of these factors relies on multiple interactions with the general transcription machinery components associated to the SNAPc (Mittal et al. 1996; Ford et al. 1998). U1 snRNA gene transcript is not spliced and the 3' end is not polyadenilated: probably this feature prevent the association with the translation machinery (Hernandez 2001). Nevertheless the snRNA gene-specific 3'-box (9-19 bp downstream the RNA-encoding region) is required for correct 3'-end formation of U1 snRNA (Egloff et al. 2008). The 3'-end formation occurs in a step manner: first there is the recognition of the *cis*-acting 3'-box. The 3'-box is a 13-16 nucleotide long element that directs the production of a 3'-extended pre-snRNA which is subsequently processed, leading to a formation of the mature 3'end after transport to the cytoplasm (Huang et al. 1997; Kiss 2004). Moreover, recently a large complex termed Integrator (constituted of Int11 and Int9) has been shown to play a role in pre-snRNA 3'-end formation (Baillat et al. 2005). Thus, these findings indicate that the 3'-box is an RNA-processing element analogous to the polyadenilation signal commonly found in protein coding genes (Uguen and Murphy 2003; Egloff et al. 2008). Recent works has demonstrated that the phosphorylation of the CTD (C- terminal domain) of the large subunit of RNA polymerase II is necessary for the 3'-box-dependent RNA 3'-end formation *in vivo*, indicating that processing occurs co-transcriptionally (Medlin et al. 2003; Jacobs et al. 2004). In particular it has been demonstrated that the CTD phosphorylation is fundamental for recruiting Integrator complex which binding is crucial for a correct 3'-end processing (Egloff et al. 2007).

U1 small nuclear ribonucleoprotein particle (U1 snRNP)

Mammalian U1 snRNP consists of the 164 bp long U1 snRN and ten different associated proteins: seven Sm proteins (B, D1, D2, D3, E, F and G) and three U1 snRNP-specific proteins: U1-70K, U1-A and U1-C (Stark et al. 2001). The U1 snRNA possess a characteristic cloverleaf secondary structure, where four stem-loops are evident, in addition

to the 5'-tail (Pomeranz Krummel et al. 2009). The association of the two largest U1-specific proteins U1-70K and U1-A depends on the presence of stem-loop I and II, respectively. Deletion or sequence alteration of these stem-loops alter significantly binding of U1-70k and U1-A and may negative consequences on U1 snRNP assembly and function (Surowy et al. 1989; Yuo and Weiner 1989; Hamm et al. 1990). The smaller U1-C protein is probably attached by protein-protein interaction with the first 97 residues of U1-70k (Nelissen et al. 1994). The characteristic Sm site, to which the Sm proteins bind (sequence *AAUUUGUGG*) is located between stem-loop III and IV (Raker et al. 1999). U1 snRNP guides the formation of the E complex, by recognizing and binding to the 5' ss on the pre-mRNA substrate (Mount et al. 1983). This interaction occurs via short RNA-RNA interactions between the consensus 5' ss at the exon/intron boundary and the 5'-tail of the U1 snRNP. The role of U1 snRNP in 5' ss recognition has been established long ago through different experimental approaches including targeting U1 snRNP 5'-tail with RNA oligos or its degradation by ribonucleases, all of which inhibit splicing processing (Mount et al. 1983; Rinke et al. 1984; Black et al. 1985). Furthermore mutations at the 5' ss can be suppressed through the introduction of compensatory changes into the U1 snRNA 5'-tail, further highlighting the role of the U1 snRNP in 5' ss recognition (Zhuang and Weiner 1986). Although 5' ss recognition can occur in absence of a U1 snRNP 5'-tail (Du and Rosbash 2002), the splicing efficiency is compromised. U1C protein contributes to 5' ss reconigition and its stabilizes the RNA duplex between the 5' ss and the U1 snRNP 5'-tail (Du and Rosbash 2002; Pomeranz Krummel et al. 2009), enhancing the formation of E complexes (Will et al. 1996). U1A seems not to be required for splicing *in vitro*, as deletion of the stem-loop II or U1A depletion from nuclear extracts do not impact on U1 snRNP activity (Heinrichs et al. 1990; Will, Rumpler et al. 1996)(Figure I.6 , I.7 and I.8). In addition to U1C, U1-70K plays an important role in U1 snRNP stabilization and promotion of E complex formation. U1-70k is known to interact with several SR proteins such as ASF/SF2, SC35 and ZRANB2 (Wu and Maniatis 1993; Jamison et al. 1995; Cao and Garcia-Blanco 1998; Wang et al. 1998; Adams et al. 2001) through its RS domain. These interactions are expected to facilitate U1snRNP recruitment to the 5' ss and subsequent formation of the E complex (Cho, Hoang et al. 2011). The importance of U1-70k behind U1 snRNP activity is further evidenced since it is the only U1 snRNP-associated protein which is regulated by phosphorylation (Tazi, Kornstadt et al. 1993). Interestingly, U1-70k phosphorylation do not impede spliceosome formation but blocks splicing catalysis (Mermoud et al. 1994) (Figure I.6 , I.7 and I.8). Even though

complementarity between the 5' ss and U1 snRNP 5'-tail is a major determinant in identification of the 5' ss, but base-pairing alone is not sufficient to specify the site of nucleophilic attack (Zhuang and Weiner 1986; Seraphin et al. 1988; Siliciano and Guthrie 1988; Liao et al. 1990). Increased base pairing to a mutant 5' ss can nevertheless maintain usage of cryptic splice sites located nearby (Siliciano and Guthrie 1988; Cohen et al. 1994). Furthermore, targeting U1 snRNPs to downstream regions of the 5' ss resulted in increased correct 5' ss usage and protein synthesis (Cohen et al. 1994). Recently a group of natural 5' ss with a shifted complementarity to U1 snRNP 5'-tail has been described, supporting the notion that U1 snRNP is required for splicing commitment of a 5' ss but do not specify itself 5' ss activation and cleavage site. U5 and U6 snRNP contribute greatly to ensure proper 5' ss selection, as they interact directly with the donor site sequence after U1 snRNP release from the pre-mRNA. In addition, some cases of U1- independent splicing have been reported, in which an abundance of SR proteins compensate the absence of a functional U1 snRNP 5'-tail (Tarn and Steitz 1994). However, the SR proteins themselves cannot discriminate real from cryptic splice sites as the successive interactions of U1, U5 and U6 snRNPs on the pre-mRNA do (Kandels-Lewis and Seraphin 1993; Lesser and Guthrie 1993; Tarn and Steitz 1994).

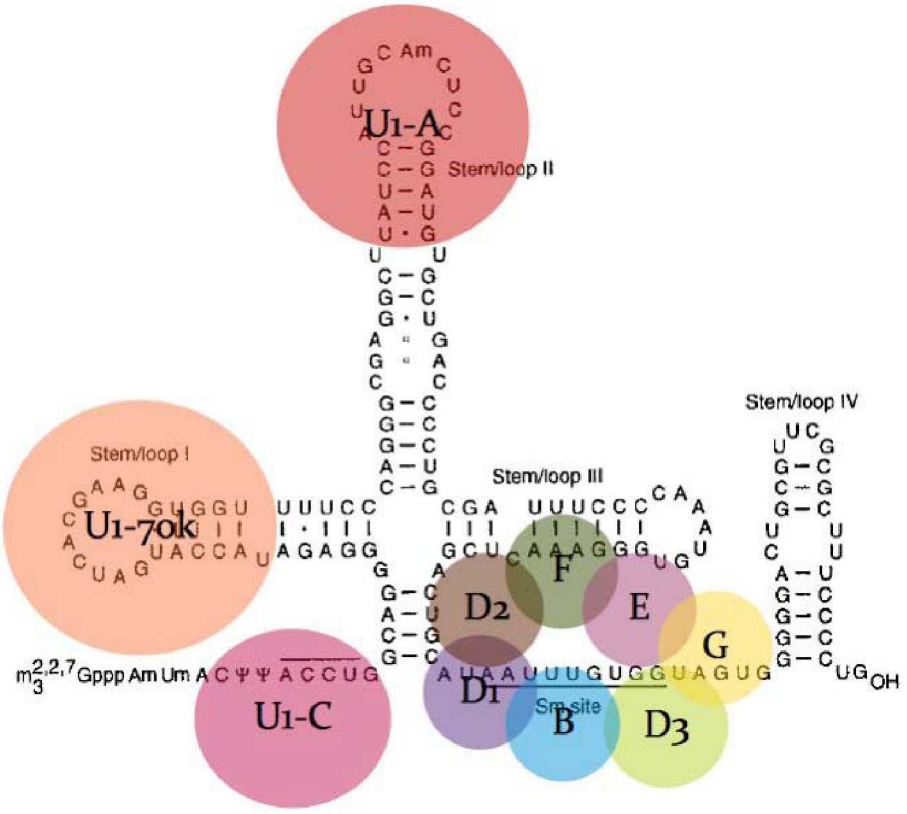


Figure I.6 Schematic structure of U1 snRNP

The secondary structure of the U1 snRNA and the sites at which the proteins U1-70K and U1-A bind the RNA are represented. The Sm site is represented with the Sm proteins. U1-70k physically interacts with U1-C and Sm-D1/D2.

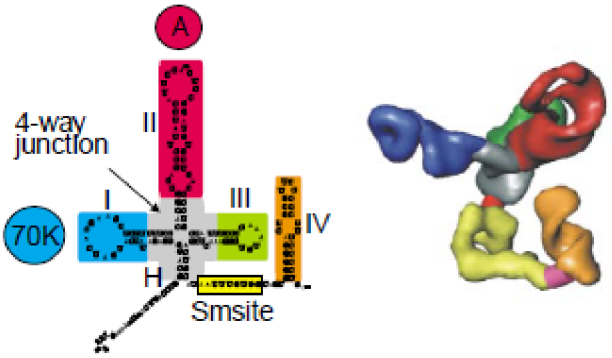


Figure I.7 Secondary and 3D structural models of the U1 snRNA.
The four stem-loops, four-way junction and Sm site are color coded.

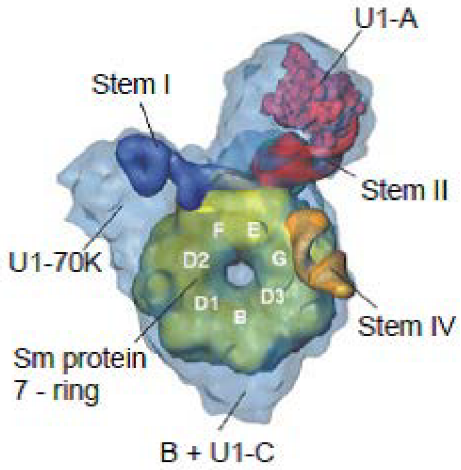


Figure I.8 Model of the 3D arrangement of the RNA and proteins in the human U1 snRNP.
A surface representation of the Sm protein ring was fitted into the 3D structure.The precise positions of stem-loops I and II of the U1 snRNA are not known

I.3 ALTERNATIVE SPLICING and SPLICING REGULATORY ELEMENTS

The use of different pairs of splice sites generates distinct mRNA isoforms from a single gene, a phenomena known as alternative splicing. It has been recently estimated that about 95% of the human genes show alternative splicing, with about 80% of these events resulting in significant levels of different protein sequences (Wang et al. 2008; Pan et al. 2008). Consequently alternative splicing generates protein isoforms with different biological functions, displaying distinct protein domains, subcellular localization or catalytic ability. This fact highlights the well-known relevance of alternative splicing in many cellular processes such as sex determination, cell differentiation, cell transformation or apoptosis (Smith and Valcàrcel 2000; Black DL 2003; Tazi et al. 2009; Nilsen et al. 2010; Pajares et al. 2007). Therefore, alternative splicing is a key regulator of gene expression that also contributes to proteome complexity.

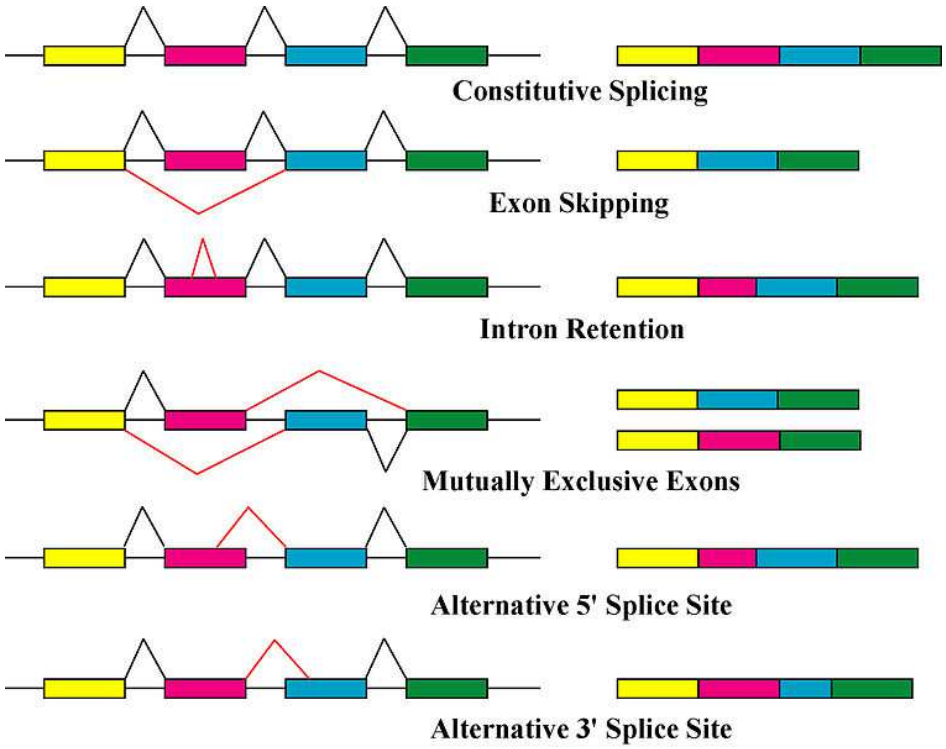


Figure I.9 Patterns of alternative splicing

A regulated exon can sometimes be included and sometimes exclude (Figure I.9) to produce a final transcript with an alternative coding sequence. In addition, introns that are normally excised can be retained in the mRNA or the position of either 5' or 3' ss can be shifted to make exons longer/shorter. A regulated exon that is sometimes included and sometimes excluded from the mRNA is usually referred to as a "cassette" exon. In some cases, multiple cassette exons are mutually exclusive producing mRNAs that always include one of few possible exon choices. All these individual patterns can be combined in a single transcription unit to produce a complex array of splice isoforms (Smith and Valcàrcel 2000; Black 2003). Alternative inclusion or skipping of cassette exons are the most common events in U2-type introns, while alternative 5' and 3' ss choices are the most observed alternative splicing events in U12-introns, due to the incompatibility of chimeric U12-U2 spliceosomes (Chang et al. 2008). The mechanisms that determine which splice sites are utilized and how this selection is regulated in different cell types or developmental stages have been heavily studied in recent years. Most alternatively spliced exons are thought to be controlled by multiple auxiliary cis-acting splicing sequences, whose activity depends on their location relative to the canonical splice sites and the trans-acting factors that recognized them (Cartegni et al. 2002; Black 2003). lndeed much effort has been made in identifying the "combinatorial code" guiding alternative splicing choices, and recently a splicing code has been proposed, based on genome- wide data generated by splice-junction microarrays or RNA-seq studies (Castle et al. 2008; Barash et al. 2010). Furthermore, the complexity of the spliceosome assembly onto the pre-mRNA and consequent catalysis reactions, together with the enormous number of factors involved, allows for fine regulation at every stage of the reaction (Park et al. 2004; Pleiss et al. 2007; Saltzman et al. 2011). In addition, alternative splicing decisions are tightly coupled to epigenetics factors such as RNA pol II elongation rate, nucleosome positioning and chromatin remodellers (Kadener et al. 2002; Batsché et al. 2006; Tilgner et al. 2009). All these events are coupled not only in time but also functionally as they influence each other and these interactions have an impact on exon definition and fate (Cramer et al. 1999; Das et al. 2006; Das et al. 2007; Luco et al. 2011). Moreover, environmental signals such as external stimuli or DNA damage modulate alternative splicing either through posttranslational modifications on trans-acting factors or alteration of epigenetic signals at the chromatin level (Blaustein et al. 2005; Luco et al. 2010;).

The degenerate and composite nature of the 5' and 3' ss allows for the existence of a wide range of natural splice sites with varying levels of strength, that is "likeness" to the consensus sequences. This intrinsic feature of the splice sites allows for a great degree of regulation of the splicing process in the form of alternative splicing. It has also an undesired consequence ; pseudo splice sites largely outnumber real splice sites and in many cases their strength can surpass that of correct splice sites (Sun et al. 2000; Buratti et al. 2007). Nevertheless, the splicing machinery is able to recognize the real splice sites, although they are weak, distant from each other and surrounded by several pseudo 5' and 3' ss. The canonical signals reported above are not sufficient to define exon/intron junctions and recognition of correct splice sites in vivo is the result of a combinatorial regulatory mechanism that involved other cis-acting regulatory elements (Smith and Valcàrcel 2000; Bruce et al. 2001). These cis-acting elements can stimulate or inhibit the use of specific splice sites: depending on their position and function, these elements are identified as exonic splicing enhancers (ESE) and silencer (ESS) or intronic splicing enhancer (ISE) and silencer (ISS).

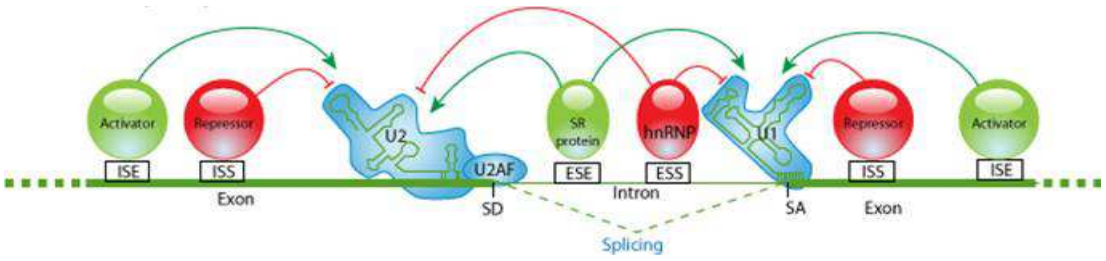
Exonic splicing enhancers. The ESEs were identified and extensively studied as regulators of alternative splicing (Black 2003) but they are also implicated in constitutive splicing events (Lavigueur et al. 1993; Schaal et al. 1999). Studies demonstrated that most of these regulatory sequences are recognized by members of serine-arginine-rich protein family (SR) (Graveley 2000) and strongly contribute to the exon definition by recruiting splicing factors and/or by interacting negatively on the action of nearby splicing silencer elements (Cartegni, et al. 2002). The majority of ESEs are located within 100 nucleotides of the splice sites and may loose their activity when located further away (Tian et al. 1994). lndeed proximity to the splice sites is used as a measure of their strength: when an ESE has better functionality from a greater distance from the splice sites, is referred as strong ESE (Graveley et al. 1998). The initial classification of ESEs was based on the type of nucleotides present in sequence cluster (Lavigueur et al. 1993; Sun Q. et al. 1993; Roscigno et al. 1995; Zuo et al. 1996; Graveley 1998). However, an exon sequence having one or more SR binding sites does not necessarily function as an ESE (Zheng et al. 1999, Zheng et al. 1998) since SR proteins also bind splking suppressors (ESSs) (Zheng et al. 1998; Mayeda et al. 1999). The fact of that there is not a single consensus sequence that can describe all the ESEs, making their identification difficult through sequence comparison or even by their interacting factors.

Exonic splicing silencers. The ESS elements are less well characterized than ESEs. They can be purine or pyrimidine-rich and bind a diverse array of proteins. In general, splicing silencers mediate exon skipping by binding to trans-acting factors that interfere with spliceosome activity mostly belonging to the hnRNP family (Fairbrother et al. 2000). Some studies have suggested exon splicing silencers to have a fundamental role in preventing pseudoexon inclusion in mature transcripts (Sironi et al. 2004). Furthermore, a specific subset of ESSs were also suggested to have distinct effects on the regulation of intron retention events in alternative splicing (Wang et al. 2004).

Intronic splicing enhancers and silencers. Fewer large-scale screens have been conducted for intronic elements and many more intronic elements are expected to be identified in future studies. One of the best characterized is represented by G triplets (GGG) or G runs (G3), that acts as ESE elements to enhance recognition of adjacent 5' or 3' ss (McCullough et al. 1997). The most studied intronic enhancer proteins are Fox-1 and Fox-2 that act at UGCAUG motifs of the brain-enriched exons (Fagnani et al. 2007; Minovitsky et al. 2005). Fox-1 proteins regulated splicing by antagonizing the repressive effect of hnRNP proteins or by regulation of the pre spliceosomal complex formation (Zhou et al. 2008) .

Composite exonic regulatory elements of splicing, CERES. The presence of the previously mentioned cis elements (enhancer and silencer) can certainly explain most of the impressive flexibility widely displayed by the splicing system. Mostly these pure ESEs/ESSs are defined by protein based score matrices and followed by in silico prediction or in an in vitro system, which hardly reflect the original environment. Moreover, the pure ESEs/ESSs most of the time behaves pretty much like the original context even if it is in a heterologous context. As a result, it is very hard to explain the effect of overlapping ESEs/ESSs sequence on splicing regulation from a classical point of view of ESEs/ESSs. Accordingly, these new types of regulatory sequences has been rena med with the acronym of CERES (composite Exonic Regulatory element of splicing). In literature, nomenclature CERES has been used so far in human CFTR exon 9 and 12 (Pagani et al. 2003a, Pagani et al. 2003b) and exon 10 of Luteinizing hormone receptor type 1 and 2 although this kind of dynamic splicing regulatory sequences or similar are found in many other exons, as SMN2 exon 7 (Singh et al 2004a). Site directed mutations in CERES elements show that they are like an overlapping enhancer and silencer, rather than

individual ESEs or ESSs (Pagani et al. 2005; Raponi et al. 2007). Moreover, these elements are also not predictable in computer-assisted systems and its functionality is not reproducible in a heterologous context (Raponi et al. 2007).



Fiure I.10 Regulation of alternative splicing

Trans-acting splicing factors

The non-snRNP splicing factors share similar structural features such as a different number of RNA binding domains and/or protein binding domains. These proteins usually target short sequence elements adjacent to sites of regulation. The non-snRNP RNA binding proteins can be involved in general or tissue-specific splicing events (David et al. 2008). In particular, two families of RNA binding proteins, hnRNP and SR proteins, have been found as components of distinct regulatory complexes with functional specificity in splicing (David et al. 2008; Chaudhury et al. 2010; Shepard et al. 2009). The SR proteins are highly conserved and play multiple roles in splicing and in general in RNA metabolism (B R Graveley 2000; Y. Huang et al. 2005; Shepard et al. 2009). All SR proteins have a modular organization and contain a N-terminal RNA-binding domain that interacts with the pre-mRNA and a C- terminal RS domain rich in arginines and serines involved mainly in protein - protein interaction; moreover the RNA binding and RS domains are modular structures and they can be exchanged between different SR proteins (Cartegni et al. 2002; Shepard et al. 2009).

SR proteins are required both for constitutive and alternative splicing events . Two non-exclusive models have been proposed to explain the role of SR protein in pre-mRNA splicing. One model is based on the ability of this splicing factors to bind ESEs and through their RS domain to recruit and stabilize U1 snRNP and U2AF binding to the 5' and 3' ss respectively. The second model proposes that a SR protein, bound to an ESE, can antagonize the negative effect of a nearby silencer element (Cartegni et al. 2002; J R Sanford et al. 2005). The activity of SR proteins is regulated through phosphorylation/dephosphorylation cycles this post-translational modification appears to

influence the subcellular localization of the protein and its ability to bind the RNA. The RS domain phosphorylation is required for the translocation of SR proteins from the cytoplasm to the nucleus (Misteli et al. 1997;; Sanford et al. 2005; Bourgeois et al. 2004; Blaustein et al. 2005; Stefan Stamm 2008).

The hnRNP proteins family is a class of several RNA-binding proteins that associate with nascent pre-mRNA: these factors remain associated with pre-mRNA until its processing is completed and with mRNA during its export from nucleus to cytoplasm (Kohler et al. 2007; Hocine et al. 2010). The hnRNP proteins frequently mediate splicing repression, particularly through binding to exonic splicing silencer (ESS) elements or by sterical interference with other protein interaction (J. Zhu et al. 2001; Cartegni et al. 2002; Fisette et al. 2010).

Although many of the hnRNPs are localized in the nucleus, a subset of these proteins shuttles continuously between nucleus and cytoplasm: this indicates a role of these proteins in nuclear export and in other cytoplasm process (Pinol- Roma et al. 1993; Martinez-Contreras et al. 2006). The role of hnRNPs in splicing is usually associated with the recognition of splicing silencers antagonizing directly or indirectly SR proteins.

I.4 DEFECTIVE SPLICING AND DISEASE

Considering the complexity of the pre-mRNA splicing process, it is not surprising that many gene mutations are directly linked to aberrant splicing processes. Hence, the study of the network of interactions between defective splicing and disease became a central issue in the medical research field (Faustino et al. 2003; Pagani et al. 2004; Garcia-blanco et al. 2004; Wessagowit et al. 2005; Emanuele Buratti et al. 2006; D.Baralle et al. 2009; Ward et al. 2010; Berasain 2010; Wang et al. 2007; Evsyukova et al. 2010; C. J. David and J. L. Manley 2010).

It is in fact now clear that substitutions that had for a long time been regarded as harmless synonymous changes in protein coding regions may have some very severe consequences on splicing process, and thus on the appearance of disease (Pagani et al. 2004; Emanuele Buratti et al. 2006). Although the frequency of splicing mutations varies considerably between individual genes, it is estimated that about one third of pathogenic mutations cause disease through the defect they introduce in the splicing mechanism. In addition, it has been demonstrated that the pathological consequences of some nonsense mutations are not due to its predicted aminoacid change but actually to their impact on splicing

(Vankeerberghen et al. 1998; K Ohno et al. 2001; Aznarez et al. 2003). Thus the primary mechanism of disease behind most pathological exonic mutations is a catastrophic splicing abnormality rather than a direct effect on coding potential (Lopez-Bigas et al. 2005). Nevertheless, the reason that the role of splicing mutations has been realized relatively late is because it has been difficult to show a clear correlation between the suspected mutation and the disease (G.-shin Wang et al. 2007; D. Baralle et al. 2005). As research ha.s progressed, it has became clear that genomic variants even if found in intronic region- should be considered as a potential disease-causing mutation affecting splicing (D. Baralle et al. 2009).

Numerous methodological developments have also aided researchers in the task of building connections between splicing and disease. For example, the refinement of minigene-based technologies for alternative splicing analysis initially described about 25 yea rs ago (Vibe-pedersen et al. 1984) has allowed a relatively fast approach to identify splicing spoilers and to study their underlying functional mechanism (D. Baralle et al. 2005; T. a Cooper 2005).

I.5 SPLICING MUTATIONS

Mutation or natural variation in pre-mRNA sequences, as well a spliceosomal components and regulatory factors, has been implicated in the etiology and progression of numerous pathologies. Splicing signals are frequent targets of mutations in genetic diseases and cancer. Most of them are single point mutations occurring in one of the first two bases (GT) in intron, immediately downstream a 5' ss, as well as AG in intron immediately upstream a 3'ss completely abolish splicing, followed by mutations at position +5 (Pohlenz et al. 2002; Krawczak et al. 2007, 1992).

Mutations at these positions are thought to reduce the complementarity between the donor splice site and the U1 snRNA 5'-tail, which is one of the first step in the complex process of pre-mR NA splicing. This results usually in exon skipping (Krawczak et al. 2007) although additional events can take place, ranging from weak splice site recognition, cryptic splice site activation, full intron inclusion or modification in mRNA secondary structure. It is important to highlight that despite weak splice site recognition leads to a decreased exon inclusion yet some mRNA is produced, and protein product is functional. On other cases, usually mature mRNA is not produced or the protein product is nonfunctional. Therefore it is of crucial importance to test each mutation to assess their

effect on splicing processing and develop better diagnostic tools and therapeutic approaches (Spurdle et al. 2008; Tournier et al. 2008; Houdayer et al. 2008; Hartmann et al. 2008). Pathological PPT mutations are less abundant than those affecting splice sites and show a homogeneous distribution (Krawczak et al. 2007). Nucleotide substitutions that occur within cis-acting elements can be classified as loss orgain of function splicing mutations if the splicing element considered is destroyed/weakened or created/enhanced, respectively (Faustino et al. 2003; Garcia-blanco et al. 2004) .These mutations in general can affect ESE, ESS or CERES or intronic regulatory elements and have been identified for several gene models (Pagani et al. 2003; Hovhannisyan et al. 2005; Petkovic et al. 2007; Skoko et al. 2008; Vidal et al. 2009; Fukao et al. 2010; Whiley et al. 2010; Drogemuller et al. 2011; Covaciu et al 2011).

A well-studied example in which the effect of a splicing mutation on an exonic regulatory sequence has been analyzed is represented by the spinal muscular atrophy (SMA), a disease in which the severity of the pathology corresponds 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 G substitution at position 6 in the pseudogene SMN2 exon 7. This mutation, though it does not change the aminoacid coding, significantly alter the splicing of pattern of the SMN2 pre-mRNA, causing frequent skipping of the exon 7 that produces an inactive and unsuitable protein lacking the last 16 aminoacids (Monani et al. 1999; Lorson et al. 1999; T. a Cooper et al. 2009). Two models have been proposed to explain exon 7 skipping in SMN2: one is that the mutation disrupt an ESE, which the splicing activator ASF2/SF2 binds (Luca Cartegni and Adrian R Krainer 2002), and the other is that it creates an ESS to which the splicing suppressor hnRNP Ai binds (Kashima et al. 2007).

Another example of mutation involving non canonical splicing regulatory sequences is represented by the dystrophin gene. It has 78 introns and mutations involving this gene are responsible of Duchenne muscular dystrophy (DMD) disease. DMD is caused by loss of function mutations, and while more than 6fAi of DMD mutations are genomic deletions, a large number of exonic and intronic point mutations cause disease through aberrant splicing (T. a Cooper et al. 2009). It is interesting to note that a particular a T>A substitution in exon 31 not only creates a premature termination codon, but also introduce an ESS that binds to hnRNP A1, resulting in partial exon skipping: the mRNA lacking this exon lose coding for one spectrin-like repeat but retain the correct reading frame to

produce a partially functional protein, giving a reason of the wilder for of the pathology of patients with this particular mutation (Disset et al. 2006).

I.6 CORRECTION APPROACHES FOR SPLICING DEFECTS

Because splicing is a prevalent cause of disease, significant efforts have been made to modulate splicing in a transcript-specific manner. For that purpose, several strategies have been tested in order to reverse disease-causing splicing defects and alleviate its symptoms.

Antisense oligonucleotides and small molecules

Antisense oligonucleotides (AONs) and small molecules are two of the most promising approaches (T. a Cooper et al. 2009; Hammond et al. 2011). AONs are designed to target and block pre-mRNA sequences that are functionally important for a disease-causing alternative splicing event. The use of AON as therapeutic molecules is providing encouraging results for the treatment of SMA. AONs designed to target an ISS in intron 7 of SMN2 stimulates exon inclusion, and consequently have been shown to significantly delay and reduce the phenotypic effects on a mild SMA mouse model (Hua et al. 2008, 2010). These AON display surprisingly persistent effect an, strikingly, systemic administration by subcutaneous injection in even more effective than intracervroventricular administration, arguing that SMN function in peripheral tissues strongly contributes to disease progression. Such approaches are currently under clinical trials for treatment of SMA and Duchenne muscular dystrophy (Benchauir et al. 2015). The rationale for the treatment of DMD is that the effect of disease-causing mutations in the *dystrophin* gene can be overcome by inducing skipping of the exon containing the mutation (or additional exon to preserve the reading frame). Other strategies involve antisense oligonucleotides blocking or degrading the CUG repeat expansion in DMD (Wheeler TM et al. 2009-2012). ASOs show high specificity, reducing the probability of causing side-effects, but delivery presents significant hurdles. In contrast, the use of small molecules, even broadly used drugs, has emerged as another tool to modify alternative splicing. Several small molecules have been described to stall spliceosome assembly at different stages of the pathway (Kuhn et al. 2009), to act as inhibitors of SR proteins (Soret et al. 2005; Bakkour et al. 2007; Keriel et al. 2009) and SR protein kinases (SRPKs and Clks) (Muraki et al. 2004; Fukuhara et al. 2006) or to bind core components of the spliceosome (Kaida et al. 2007; Kotake et al.

2007). The use of these molecules as therapeutic tools depends entirely on the side effects that they might cause.

U snRNP-based therapies for splicing modulation.

The use of modified U snRNPs represents an alternative to ASO-based therapies for several reasons. First, U snRNAs are naturally exported to the nucleus to target pre-mRNAs, ensuring higher efficient delivery with respect to ASOs. Second, delivery of U snRNAs in a proper delivery vector (i.e., adeno-associated virus, AAV) would require few, if not only one, administrations to prospective patients. Third, the antisense sequence is protected from degradation since it is integrated within the U snRNA gene.

The U7 snRNP belongs to the family of small nuclear ribonucleoprotein particles like U1 snRNP but it is not part of the spliceosomal complex. U7 snRNP is a key player of the histone 3' end pre-mRNA processing (B. Miller et al. 1997). The normal U7 snRNP has a low affinity Sm site which is responsible for its low levels in the nucleus. Therefore in order to be used for splicing modulation, the Sm site was replaced for an optimal site. This particle, named U7 SmOPT exhibits improved nuclear accumulation levels and is no longer functional for histone RNA processing (Grimm et al. 1993; Stefanovic et al. 1995). The natural anti-histone 3' end motif can be easily replaced for the antisense target sequence of interest. Modified versions of U7 SmOPT have been previously used to redirect alternative splicing events in several gene models (Gorman et al. 1998; Goyenvalle et al. 2004; Madocsai et al. 2005; Uchikawa et al. 2007; Asparuhova et al. 2007)

The 5'-tail of U1 snRNA have been also altered in order to deliver antisense sequences. A modified U1 snRNA (with a 54 bp modified 5'-tail) targeting mouse DMD gene exon 23 3' and 5' ss was systemically delivered using AAV vectors to the dystrophin-deficient mouse model of DMD, mdx. Body-wide dystrophin restoration was observed in treated mice, although heterogeneous throughout the skeletal muscles (Denti et al. 2006). More recently, investigators used U1 snRNA to correct splicing through exon skipping in huma n DMD pre-mRNA in primary patient fibroblasts . The most efficient exon skipping was achieved from targeting splice site as well as exonic regulatory regions of exon 51. Similarly, U7 SmOPT snRNA targeting SMN2 exon 7 in SMA patient-derived fibroblasts was able to induce exon 7 inclusion in up to 80% of transcripts (Geib et al. 2009).

The modified U1 snRNA can be exploited to suppress the impact of disease-associated splice site mutations, complementing the inability of the normal U1snRNP activity. In these cases, the modified U1 snRNAs have few nucleotide changes in comparison to the

wt sequences and base pair exactly to the mutant donor sites. These approaches are based on the previous notion that compensatory changes suppress 5' ss mutants (Zhuang and Weiner, 1986). These approaches have been proposed for the correction of splicing defects in human malignant infantile osteoporosis (arOP)(Susani et al. 2004), coagulation factor VII and IX (FVII-FIX) deficiency (Pinotti et al. 2008, 2009; Fernandez Alanis 2012; Balestra et al. 2014, Balestra et al. 2015), ATP8B1 deficiency (van der Woerd WL et al. 2015), Netherton syndrome (NS) (Dal Mas A. et al. 2015), retinitis pigmentosa (Tanner et al. 2009; Glaus et al. 2011), Fanconi anemia (Hartmann et al. 2010) and Bardet-Biedl syndrome (Schmid et al. 2011). For several of these models, in addition to correct pre-mRNA splicing processing, protein biosynthesis and function was restored upon treatment with modified U1snRNPs (Pinotti et al 2009; Hartmann et al. 2010).

U1 snRNAs complementary to the 5'ss might target by complementarity similar donor sites and thus potentially interfere with splicing of other pre-mRNAs with toxic effects. The use of Exon Specific U1snRNAs (ExSpeU1) recently identified in our group (Fernandez et al. 2012) represents a novel promising strategy to overcome the above mentioned limitations. ExSpeU1s are designed to bind by complementarity to intronic sequences downstream of the exon and rescue different types of splicing defects associated to exon skipping. Binding of the ExSpeU1s to non-conserved intronic sequences will significantly reduce the possibility of off-target events. ExSpeU1s are active on several 5'ss mutations, on mutations in the polypyrimidine tract and on defective splicing caused by exonic variants (Fernandez Alanis 2012).

Trans-splicing

Trans-splicing is a natural process involving splice sites in two different pre-mRNA transcripts and occurs in a variety of organisms, including protozoa, trypanosomes and nematodes. It has also been observed in *Drosophila* and mammalian cells, linked to apoptosis, axon guidance, and maintenance of cell pluripotency (Gingeras TR 2009; Wu CS et al. 2014). Pre-mRNA *Trans*-splicing, *Trans*-splicing ribozymes, and t-RNA splicing endonucleases have been proposed as RNA repair strategies of potential therapeutic value (Yang Y et al. 2005). Spliceosome-mediated RNA *Trans*-splicing (SMaRT) approaches have been anticipated for the treatment of several disease, including cystic fibrosis (Lui X et al. 2002), SMA (Coady TH et al. 2007), DMD (Lorain S et al. 2013), and RP (Berger A et al. 2015). The general strategy is to introduce a pre-*trans*-splicing molecules (PTM) containing the sequence to be replaced, preceded by a targeting sequence complementary

to an intron in the target RNA and containing also a 3' ss. Splicing between the 5' ss of the target intron and the 3' ss of the PTM leads to chimeric transcripts that restore correct mRNA expression. The main hurdle for the therapeutic application of these technologies remains to enhance the limited in vivo efficacy of the *trans*-splicing process.

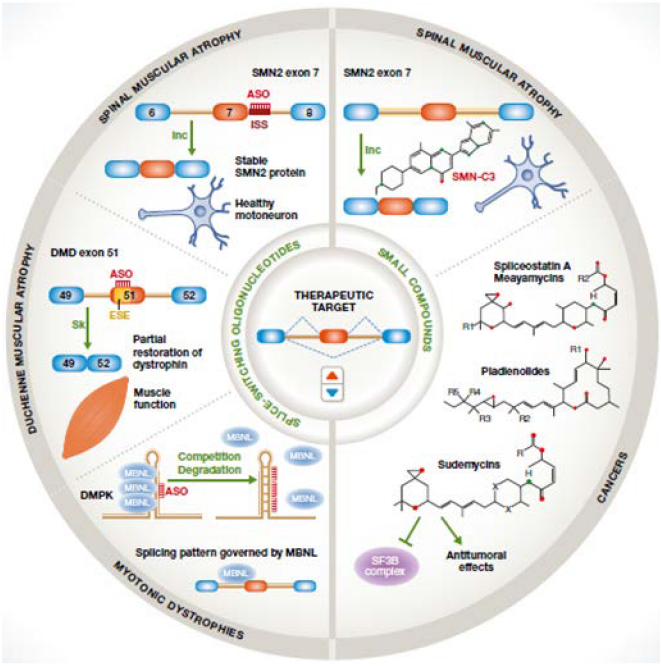


Figure I.11 Summary of therapeutic approaches based upon splicing modulation. Two main strategies are outlined. Splice-switching antisense oligonucleotides target splice sites or splicing regulatory sequences to prevent the binding of cognate factor to modulate splice site selection.Examples include blocking of an ISS element that promote exon 7 inclusion in SMN2 gene; induction of skipping of a mutation containing exon in DMD gene; prevention of sequestration of the MBNL splicing regulation in CUG repeat expansion in DMPK transcripts. Small molecules modulate alternative splicing of cell cycle control genes and display anti-tumoral proprieties.

REFERENCES

-Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. EMBO J. 1999 Oct 15;18(20):5789-802.

-Adams DJ, van der Weyden L, Mayeda A, Stamm S, Morris BJ, Rasko JE. ZNF265--a novel spliceosomal protein able to induce alternative splicing. J Cell Biol. 2001 Jul 9;154(1):25-32.

-Aebi M, Hornig H, Weissmann C. 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. Cell. 1987 Jul 17;50(2):237-46.

-Asparuhova MB, Marti G, Liu S, Serhan F, Trono D, Schümperli D. Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing Tat and Rev exon skipping. J Gene Med. 2007 May;9(5):323-34.

-Ast G. How did alternative splicing evolve? Nat Rev Genet. 2004 Oct;5(10):773-82. Review.

-Aznarez I, Chan EM, Zielenski J, Blencowe BJ, Tsui LC. Characterization of disease-associated mutations affecting an exonic splicing enhancer and two cryptic splice sites in exon 13 of the cystic fibrosis transmembrane conductance

-Babich A, Nevins JR, Darnell JE Jr. Early capping of transcripts from the adenovirus major late transcription unit. Nature. 1980 Sep 18;287(5779):246-8.

-Baillat D, Hakimi MA, Näär AM, Shilatifard A, Cooch N, Shiekhattar R. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell. 2005 Oct 21;123(2):265-76.

-Bakkour N, Lin YL, Maire S, Ayadi L, Mahuteau-Betzer F, Nguyen CH, Mettling C, Portales P, Grierson D, Chabot B, Jeanteur P, Branlant C, Corbeau P, Tazi J. Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. PLoS Pathog. 2007 Oct 26;3(10):1530-9.

-Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M. An engineered U1 small nuclear RNA rescues splicing defective coagulation F7 gene expression in mice. J Thromb Haemost. 2014 Feb;12(2):177-85.

-Balestra D, Barbon E, Scalet D, Cavallari N, Perrone D, Zanibellato S, Bernardi F, Pinotti M. Regulation of a strong F9 cryptic 5'ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides. Hum Mol Genet. 2015 Sep 1;24(17):4809-16. doi: 10.1093/hmg/ddv205. Epub 2015 Jun 10.

-Baralle D, Lucassen A, Buratti E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. EMBO Rep. 2009 Aug;10(8):810-6. doi:10.1038/embor.2009.170. Review.

-Barash Y, Calarco JA, Gao W, Pan Q, Wang X, Shai O, Blencowe BJ, Frey BJ. Deciphering the splicing code. Nature. 2010 May 6;465(7294):53-9. doi: 10.1038/nature09000.

-Batsché E, Yaniv M, Muchardt C. The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat Struct Mol Biol. 2006 Jan;13(1):22-9. Epub 2005 Dec 11.

-Battle DJ, Kasim M, Yong J, Lotti F, Lau CK, Mouaikel J, Zhang Z, Han K, Wan L, Dreyfuss G. The SMN complex: an assembly machine for RNPs. Cold Spring Harb Symp Quant Biol. 2006;71:313-20. Review.

-Behrens SE, Lührmann R. Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. Genes Dev. 1991 Aug;5(8):1439-52.

-Benchaouir R, Robin V, Goyenvalle A. Gene and splicing therapies for neuromuscular diseases. Front Biosci (Landmark Ed). 2015 Jun 1;20:1190-233.

-Berasain C, Goñi S, Castillo J, Latasa MU, Prieto J, Avila MA. Impairment of pre-mRNA splicing in liver disease: mechanisms and consequences. World J Gastroenterol. 2010 Jul 7;16(25):3091-102. Review.

-Berger A, Lorain S, Joséphine C, Desrosiers M, Peccate C, Voit T, Garcia L, Sahel JA, Bemelmans AP. Repair of rhodopsin mRNA by spliceosome-mediated RNA trans-splicing: a new approach for autosomal dominant retinitis pigmentosa. Mol Ther. 2015 May;23(5):918-30. doi: 10.1038/mt.2015.11. Epub 2015 Jan 26.

-Berget SM. Exon recognition in vertebrate splicing. J Biol Chem. 1995 Feb 10;270(6):2411-4. Review.

-Berglund JA, Chua K, Abovich N, Reed R, Rosbash M. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell. 1997 May 30;89(5):781-7.

-Bernstein LB, Manser T, Weiner AM. Human U1 small nuclear RNA genes: extensive conservation of flanking sequences suggests cycles of gene amplification and transposition. Mol Cell Biol. 1985 Sep;5(9):2159-71.

-Black DL, Chabot B, Steitz JA. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell. 1985 Oct;42(3):737-50.

-Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem. 2003;72:291-336. Epub 2003 Feb 27. Review.

-Blaustein M, Pelisch F, Tanos T, Muñoz MJ, Wengier D, Quadrana L, Sanford JR, Muschietti JP, Kornblihtt AR, Cáceres JF, Coso OA, Srebrow A. Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT. Nat Struct Mol Biol. 2005 Dec;12(12):1037-44. Epub 2005 Nov 20.

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

-Branlant C, Krol A, Ebel JP, Lazar E, Haendler B, Jacob M. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. EMBO J. 1982;1(10):1259-65.

-Bringmann P, Lührmann R. Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. EMBO J. 1986 Dec 20;5(13):3509-16

-Bringmann P, Lührmann R. Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. EMBO J. 1986 Dec 20;5(13):3509-16.

-Bruce SR, Peterson ML. Multiple features contribute to efficient constitutive splicing of an unusually large exon. Nucleic Acids Res. 2001 Jun 1;29(11):2292-302.

-Buratti E, Baralle FE. Influence of RNA secondary structure on the pre-mRNA splicing process. Mol Cell Biol. 2004 Dec;24(24):10505-14. Review.

-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(13):4250-63. Epub 2007 Jun 18.

-Cao W, Garcia-Blanco MA. A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding. J Biol Chem. 1998 Aug 7;273(32):20629-35.

-Carbon P, Murgo S, Ebel JP, Krol A, Tebb G, Mattaj LW. A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase III and U2 snRNA by RNA polymerase II. Cell. 1987 Oct 9;51(1):71-9.

-Carmel I, Tal S, Vig I, Ast G. Comparative analysis detects dependencies among the 5' splice-site positions. RNA. 2004 May;10(5):828-40.

-Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet. 2002 Apr;3(4):285-98. Review.

-Cartegni L, Krainer AR. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat Genet. 2002 Apr;30(4):377-84. Epub 2002 Mar 4.

-Castle JC, Zhang C, Shah JK, Kulkarni AV, Kalsotra A, Cooper TA, Johnson JM. Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. Nat Genet. 2008 Dec;40(12):1416-25. doi: 10.1038/ng.264. Epub 2008 Nov 2.

-Chanfreau G, Jacquier A. Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron. EMBO J. 1993 Dec 15;12(13):5173-80.

-Chang WC, Chen HH, Tarn WY. Alternative splicing of U12-type introns. Front Biosci. 2008 Jan 1;13:1681-90. Review.

-Chaudhury A, Chander P, Howe PH. Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. RNA. 2010 Aug;16(8):1449-62. doi: 10.1261/rna.2254110. Epub 2010 Jun 28. Review.

-Cho S, Hoang A, Sinha R, Zhong XY, Fu XD, Krainer AR, Ghosh G. Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. Proc Natl Acad Sci U S A. 2011 May 17;108(20):8233-8. doi: 10.1073/pnas.1017700108. Epub 2011 May 2.

-Coady TH, Shababi M, Tullis GE, Lorson CL. Restoration of SMN function:delivery of a trans-splicing RNA re-directs SMN2 pre-mRNA splicing. Mol Ther. 2007 Aug;15(8):1471-8. Epub 2007 Jun 5.

-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 U S A. 1994 Oct 25;91(22):10470-4.

-Cooper TA. Use of minigene systems to dissect alternative splicing elements. Methods. 2005 Dec;37(4):331-40.

-Cooper TA, Wan L, Dreyfuss G. RNA and disease. Cell. 2009 Feb 20;136(4):777-93. doi: 10.1016/j.cell.2009.02.011. Review.

-Coppola JA, Field AS, Luse DS. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. Proc Natl Acad Sci U S A. 1983 Mar;80(5):1251-5.

-Covaciu C, Grosso F, Pisaneschi E, Zambruno G, Gregersen PA, Sommerlund M, Hertz JM, Castiglia D. 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. 2011 Sep;165(3):678-82. doi: 10.1111/j.1365-2133.2011.10414.x. Epub 2011 Jul 28.

-Cramer P, Cáceres JF, Cazalla D, Kadener S, Muro AF, Baralle FE, Kornblihtt AR. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol Cell. 1999 Aug;4(2):251-8.

-Dal Mas A, Fortugno P, Donadon I, Levati L, Castiglia D, Pagani F. Exon-Specific U1s Correct SPINK5 Exon 11 Skipping Caused by a Synonymous Substitution that Affects a Bifunctional Splicing Regulatory Element. Hum Mutat. 2015 May;36(5):504-12. doi: 10.1002/humu.22762. Epub 2015 Mar 19.

-Das R, Dufu K, Romney B, Feldt M, Elenko M, Reed R. Functional coupling of RNAP II transcription to spliceosome assembly. Genes Dev. 2006 May 1;20(9):1100-9.

-Das R, Reed R. Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels. RNA. 1999 Nov;5(11):1504-8.

-Das R, Yu J, Zhang Z, Gygi MP, Krainer AR, Gygi SP, Reed R. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. Mol Cell. 2007 Jun 22;26(6):867-81.

-David CJ, Manley JL. The search for alternative splicing regulators: new approaches offer a path to a splicing code. Genes Dev. 2008 Feb 1;22(3):279-85. doi: 10.1101/gad.1643108.

-David CJ, Manley JL. Alternative pre-mRNA splicing regulation in cancer:pathways and programs unhinged. Genes Dev. 2010 Nov 1;24(21):2343-64. doi:10.1101/gad.1973010. Review.

-Denti MA, Rosa A, D'Antona G, Sthandier O, De Angelis FG, Nicoletti C, Allocca M, Pansarasa O, Parente V, Musarò A, Auricchio A, Bottinelli R, Bozzoni I. Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model. Proc Natl Acad Sci U S A. 2006 Mar 7;103(10):3758-63. Epub 2006 Feb 24.

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

-Domitrovich AM, Kunkel GR. Multiple, dispersed human U6 small nuclear RNA genes with varied transcriptional efficiencies. Nucleic Acids Res. 2003 May 1;31(9):2344-52.

-Drögemüller C, Reichart U, Seuberlich T, Oevermann A, Baumgartner M, Kühni Boghenbor K, Stoffel MH, Syring C, Meylan M, Müller S, Müller M, Gredler B, Sölkner J, Leeb T. An unusual splice defect in the mitofusin 2 gene (MFN2) is associated with degenerative axonopathy in Tyrolean Grey cattle. PLoS One. 2011 Apr 15;6(4):e18931. doi: 10.1371/journal.pone.0018931.

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

-Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. Science. 2007 Dec 14;318(5857):1777-9.

-Egloff S, O'Reilly D, Murphy S. Expression of human snRNA genes from beginning to end. Biochem Soc Trans. 2008 Aug;36(Pt 4):590-4. doi: 10.1042/BST0360590. Review.

-Eliceiri GL, Gurney T Jr. Subcellular location of precursors to small nuclear RNA species C and D and of newly synthesized 5 S RNA in HeLa cells. Biochem Biophys Res Commun. 1978 Apr 14;81(3):915-9.

-Eliceiri GL, Sayavedra MS. Small RNAs in the nucleus and cytoplasm of HeLa cells. Biochem Biophys Res Commun. 1976 Sep 20;72(2):507-12.

-Evsyukova I, Somarelli JA, Gregory SG, Garcia-Blanco MA. Alternative splicing in multiple sclerosis and other autoimmune diseases. RNA Biol. 2010 Jul-Aug;7(4):462-73. Epub 2010 Jul 1. Review.

-Fabrizio P, Esser S, Kastner B, Lührmann R. Isolation of S. cerevisiae snRNPs: comparison of U1 and U4/U6.U5 to their human counterparts. Science. 1994 Apr 8;264(5156):261-5.

-Fagnani M, Barash Y, Ip JY, Misquitta C, Pan Q, Saltzman AL, Shai O, Lee L, Rozenhek A, Mohammad N, Willaime-Morawek S, Babak T, Zhang W, Hughes TR, van der Kooy D, Frey BJ, Blencowe BJ. Functional coordination of alternative splicing in the mammalian central nervous system. Genome Biol. 2007;8(6):R108.

-Fairbrother WG, Chasin LA. Human genomic sequences that inhibit splicing. Mol Cell Biol. 2000 Sep;20(18):6816-25.

-Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. Genes Dev. 2003 Feb 15;17(4):419-37. Review.

-Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. Hum Mol Genet. 2012 Jun 1;21(11):2389-98. doi:10.1093/hmg/dds045. Epub 2012 Feb 23.

-Fisette JF, Toutant J, Dugré-Brisson S, Desgroseillers L, Chabot B. hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection. RNA. 2010 Jan;16(1):228-38. doi: 10.1261/rna.1890310. Epub 2009 Nov 19.

-Ford E, Strubin M, Hernandez N. The Oct-1 POU domain activates snRNA gene transcription by contacting a region in the SNAPc largest subunit that bears sequence similarities to the Oct-1 coactivator OBF-1. Genes Dev. 1998 Nov 15;12(22):3528-40.

-Fukao T, Horikawa R, Naiki Y, Tanaka T, Takayanagi M, Yamaguchi S, Kondo N. A novel mutation (c.951C>T) in an exonic splicing enhancer results in exon 10 skipping in the human mitochondrial acetoacetyl-CoA thiolase gene. Mol Genet Metab. 2010 Aug;100(4):339-44. doi: 10.1016/j.ymgme.2010.03.012. Epub 2010 Mar 19.

-Fukuhara T, Hosoya T, Shimizu S, Sumi K, Oshiro T, Yoshinaka Y, Suzuki M, Yamamoto N, Herzenberg LA, Herzenberg LA, Hagiwara M. Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. Proc Natl Acad Sci U S A. 2006 Jul 25;103(30):11329-33. Epub 2006 Jul 13.

-Garcia-Blanco MA, Baraniak AP, Lasda EL. Alternative splicing in disease and therapy. Nat Biotechnol. 2004 May;22(5):535-46.

-Geib T, Hertel KJ. Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs. PLoS One. 2009 Dec 8;4(12):e8204. doi: 10.1371/journal.pone.0008204.

-Gingeras TR. Implications of chimaeric non-co-linear transcripts. Nature. 2009 Sep 10;461(7261):206-11. doi: 10.1038/nature08452. Review.

-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 May;19(5):936-41. doi: 10.1038/mt.2011.7. Epub 2011 Feb 15.

-Gooding C, Clark F, Wollerton MC, Grellscheid SN, Groom H, Smith CW. A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones. Genome Biol. 2006;7(1):R1. Epub 2006 Jan 13.

-Gorman L, Suter D, Emerick V, Schümperli D, Kole R. Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. Proc Natl Acad Sci U S A. 1998 Apr 28;95(9):4929-34.

-Goyenvalle A, Vulin A, Fougerousse F, Leturcq F, Kaplan JC, Garcia L, Danos O. Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. Science. 2004 Dec 3;306(5702):1796-9. Epub 2004 Nov 4.

-Gozani O, Feld R, Reed R. 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. 1996 Jan 15;10(2):233-43.

-Graveley BR, Hertel KJ, Maniatis T. A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. EMBO J. 1998 Nov 16;17(22):6747-56.

-Graveley BR. Sorting out the complexity of SR protein functions. RNA. 2000 Sep;6(9):1197-211. Review.

-Gridasova AA, Henry RW. The p53 tumor suppressor protein represses human snRNA gene transcription by RNA polymerases II and III independently of sequence-specific DNA binding. Mol Cell Biol. 2005 Apr;25(8):3247-60.

-Grimm C, Stefanovic B, Schümperli D. The low abundance of U7 snRNA is partly determined by its Sm binding site. EMBO J. 1993 Mar;12(3):1229-38.

-Hamm J, Dathan NA, Scherly D, Mattaj IW. Multiple domains of U1 snRNA, including U1 specific protein binding sites, are required for splicing. EMBO J. 1990 Apr;9(4):1237-44.

-Hammond SM, Wood MJ. Genetic therapies for RNA mis-splicing diseases. Trends Genet. 2011 May;27(5):196-205. doi: 10.1016/j.tig.2011.02.004. Epub 2011 Apr 15. Review.

-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 Oct 8;87(4):480-93. doi: 10.1016/j.ajhg.2010.08.016.

-Hartmann L, Theiss S, Niederacher D, Schaal H. Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases? Front Biosci. 2008 May 1;13:3252-72. Review.

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

-Hernandez N. Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. J Biol Chem. 2001 Jul 20;276(29):26733-6. Epub 2001 Jun 4. Review.

-Heyd F, Lynch KW. Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing. Mol Cell. 2010 Oct 8;40(1):126-37. doi: 10.1016/j.molcel.2010.09.013.

-Hirsch HA, Gu L, Henry RW. The retinoblastoma tumor suppressor protein targets distinct general transcription factors to regulate RNA polymerase III gene expression. Mol Cell Biol. 2000 Dec;20(24):9182-91.

-Hocine S, Singer RH, Grünwald D. RNA processing and export. Cold Spring Harb Perspect Biol. 2010 Dec;2(12):a000752. doi: 10.1101/cshperspect.a000752. Epub 2010 Oct 20. Review.

-Hong W, Bennett M, Xiao Y, Feld Kramer R, Wang C, Reed R. Association of U2 snRNP with the spliceosomal complex E. Nucleic Acids Res. 1997 Jan 15;25(2):354-61.

-Horowitz DS, Krainer AR. Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing. Trends Genet. 1994 Mar;10(3):100-6. Review.

-Houdayer C, Dehainault C, Mattler C, Michaux D, Caux-Moncoutier V, Pagès-Berhouet S, d'Enghien CD, Laugé A, Castera L, Gauthier-Villars M, Stoppa-Lyonnet D. Evaluation of in silico splice tools for decision-making in molecular diagnosis. Hum Mutat. 2008 Jul;29(7):975-82. doi: 10.1002/humu.20765.

-Hovhannisyan RH, Carstens RP. A novel intronic cis element, ISE/ISS-3, regulates rat fibroblast growth factor receptor 2 splicing through activation of an upstream exon and repression of a downstream exon containing a noncanonical branch point sequence. Mol Cell Biol. 2005 Jan;25(1):250-63.

-Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. Am J Hum Genet. 2008 Apr;82(4):834-48. doi: 10.1016/j.ajhg.2008.01.014. Epub 2008 Mar 27.

-Huang Q, Jacobson MR, Pederson T. 3' processing of human pre-U2 small nuclear RNA: a base-pairing interaction between the 3' extension of the precursor and an internal region. Mol Cell Biol. 1997 Dec;17(12):7178-85.

-Huang Y, Steitz JA. SRprises along a messenger's journey. Mol Cell. 2005 Mar 4;17(5):613-5. Review.

-Izaurralde E, Stepinski J, Darzynkiewicz E, Mattaj IW. A cap binding protein that may mediate nuclear export of RNA polymerase II-transcribed RNAs. J Cell Biol. 1992 Sep;118(6):1287-95.

-Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj IW. A cap-binding protein complex mediating U snRNA export. Nature. 1995 Aug 24;376(6542):709-12.

-Jacobs EY, Ogiwara I, Weiner AM. Role of the C-terminal domain of RNA polymerase II in U2 snRNA transcription and 3' processing. Mol Cell Biol. 2004 Jan;24(2):846-55.

-Jamison SF, Garcia-Blanco MA. An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. Proc Natl Acad Sci U S A. 1992 Jun 15;89(12):5482-6.

-Jamison SF, Pasman Z, Wang J, Will C, Lührmann R, Manley JL, Garcia-Blanco MA. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. Nucleic Acids Res. 1995 Aug 25;23(16):3260-7.

-Jurica MS, Moore MJ. Pre-mRNA splicing: awash in a sea of proteins. Mol Cell. 2003 Jul;12(1):5-14. Review.

-Kadener S, Fededa JP, Rosbash M, Kornblihtt AR. Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. Proc Natl Acad Sci U S A. 2002 Jun 11;99(12):8185-90.

-Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H, Tani T, Horinouchi S, Yoshida M. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat Chem Biol. 2007 Sep;3(9):576-83. Epub 2007 Jul 22.

-Kandels-Lewis S, Séraphin B. Involvement of U6 snRNA in 5' splice site selection. Science. 1993 Dec 24;262(5142):2035-9. Erratum in: Science 1994 Feb 11;263(5148):739.

-Kashima T, Rao N, David CJ, Manley JL. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. Hum Mol Genet. 2007 Dec 15;16(24):3149-59. Epub 2007 Sep 19.

-Kent OA, Ritchie DB, Macmillan AM. Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway. Mol Cell Biol. 2005 Jan;25(1):233-40.

-Keriel A, Mahuteau-Betzer F, Jacquet C, Plays M, Grierson D, Sitbon M, Tazi J. Protection against retrovirus pathogenesis by SR protein inhibitors. PLoS One. 2009;4(2):e4533. doi: 10.1371/journal.pone.0004533. Epub 2009 Feb 19.

-Kiss T. Biogenesis of small nuclear RNPs. J Cell Sci. 2004 Dec 1;117(Pt 25):5949-51. Review.

-Köhler A, Hurt E. Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol. 2007 Oct;8(10):761-73. Review.

-Kolb SJ, Battle DJ, Dreyfuss G. Molecular functions of the SMN complex. J Child Neurol. 2007 Aug;22(8):990-4. Review.

-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 Feb;28(2):150-8.

-Kuhlman TC, Cho H, Reinberg D, Hernandez N. 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. 1999 Mar;19(3):2130-41.

-Lamond AI, Konarska MM, Sharp PA. A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. Genes Dev. 1987 Aug;1(6):532-43.

-Lander ES, Linton LM,et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860-921.

-Lavigueur A, La Branche H, Kornblihtt AR, Chabot B. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. Genes Dev. 1993 Dec;7(12A):2405-17.

-Lesser CF, Guthrie C. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science. 1993 Dec 24;262(5142):1982-8.

-Liao XL, Kretzner L, Seraphin B, Rosbash M. Universally conserved and yeast-specific U1 snRNA sequences are important but not essential for U1 snRNP function. Genes Dev. 1990 Oct;4(10):1766-74.

-Liu JL, Gall JG. U bodies are cytoplasmic structures that contain uridine-rich small nuclear ribonucleoproteins and associate with P bodies. Proc Natl Acad Sci U S A. 2007 Jul 10;104(28):11655-9. Epub 2007 Jun 26.

-Liu S, Rauhut R, Vornlocher HP, Lührmann R. The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. RNA. 2006 Jul;12(7):1418-30. Epub 2006 May 24.

-López-Bigas N, Audit B, Ouzounis C, Parra G, Guigó R. Are splicing mutations the most frequent cause of hereditary disease? FEBS Lett. 2005 Mar 28;579(9):1900-3.

-Lorain S, Peccate C, Le Hir M, Griffith G, Philippi S, Précigout G, Mamchaoui K, Jollet A, Voit T, Garcia L. Dystrophin rescue by trans-splicing: a strategy for DMD genotypes not eligible for exon skipping approaches. Nucleic Acids Res. 2013 Sep;41(17):8391-402. doi: 10.1093/nar/gkt621. Epub 2013 Jul 16.

-Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc Natl Acad Sci U S A. 1999 May 25;96(11):6307-11.

-Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of alternative splicing by histone modifications. Science. 2010 Feb 19;327(5968):996-1000. doi: 10.1126/science.1184208. Epub 2010 Feb 4.

-Luco RF, Misteli T. More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. Curr Opin Genet Dev. 2011 Aug;21(4):366-72. doi: 10.1016/j.gde.2011.03.004. Epub 2011 Apr 15.Review.

-Lund E, Dahlberg JE. True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation. J Biol Chem. 1984 Feb 10;259(3):2013-21.

-Madocsai C, Lim SR, Geib T, Lam BJ, Hertel KJ. Correction of SMN2 Pre-mRNA splicing by antisense U7 small nuclear RNAs. Mol Ther. 2005 Dec;12(6):1013-22. Epub 2005 Oct 14.

-Malca H, Shomron N, Ast G. The U1 snRNP base pairs with the 5' splice site within a penta-snRNP complex. Mol Cell Biol. 2003 May;23(10):3442-55.

-Maniatis T, Reed R. An extensive network of coupling among gene expression machines. Nature. 2002 Apr 4;416(6880):499-506. Review.

-Martinez-Contreras R, Fisette JF, Nasim FU, Madden R, Cordeau M, Chabot B. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. PLoS Biol. 2006 Feb;4(2):e21. Epub 2006 Jan 10.

-Massenet S, Pellizzoni L, Paushkin S, Mattaj IW, Dreyfuss G. The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. Mol Cell Biol. 2002 Sep;22(18):6533-41.

-Mayeda A, Screaton GR, Chandler SD, Fu XD, Krainer AR. Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. Mol Cell Biol. 1999 Mar;19(3):1853-63.

-Mayes AE, Verdone L, Legrain P, Beggs JD. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. EMBO J. 1999 Aug 2;18(15):4321-31.

-McCullough AJ, Berget SM. G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. Mol Cell Biol. 1997 Aug;17(8):4562-71.

-Medlin JE, Uguen P, Taylor A, Bentley DL, Murphy S. The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA. EMBO J. 2003 Feb 17;22(4):925-34.

-Mermoud JE, Cohen PT, Lamond AI. Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. EMBO J. 1994 Dec 1;13(23):5679-88.

-Minovitsky S, Gee SL, Schokrpur S, Dubchak I, Conboy JG. The splicing regulatory element, UGCAUG, is phylogenetically and spatially conserved in introns that flank tissue-specific alternative exons. Nucleic Acids Res. 2005 Feb 3;33(2):714-24. Print 2005.

-Misteli T, Spector DL. Protein phosphorylation and the nuclear organization of pre-mRNA splicing. Trends Cell Biol. 1997 Apr;7(4):135-8.

-Mittal V, Cleary MA, Herr W, Hernandez N. The Oct-1 POU-specific domain can stimulate small nuclear RNA gene transcription by stabilizing the basal transcription complex SNAPc. Mol Cell Biol. 1996 May;16(5):1955-65.

-Monani UR, Lorson CL, Parsons DW, Prior TW, Androphy EJ, Burghes AH, McPherson JD. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. Hum Mol Genet. 1999 Jul;8(7):1177-83.

-Mouaikel J, Narayanan U, Verheggen C, Matera AG, Bertrand E, Tazi J, Bordonné R. Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. EMBO Rep. 2003 Jun;4(6):616-22.

-Mount SM, Pettersson I, Hinterberger M, Karmas A, Steitz JA. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell. 1983 Jun;33(2):509-18.

-Muraki M, Ohkawara B, Hosoya T, Onogi H, Koizumi J, Koizumi T, Sumi K, Yomoda J, Murray MV, Kimura H, Furuichi K, Shibuya H, Krainer AR, Suzuki M, Hagiwara M. Manipulation of alternative splicing by a newly developed inhibitor of Clks. J Biol Chem. 2004 Jun 4;279(23):24246-54. Epub 2004 Mar 8.

-Nagai K, Muto Y, Pomeranz Krummel DA, Kambach C, Ignjatovic T, Walke S, Kuglstatter A. Structure and assembly of the spliceosomal snRNPs. Novartis Medal Lecture. Biochem Soc Trans. 2001 May;29(Pt 2):15-26.

-Nelissen RL, Will CL, van Venrooij WJ, Lührmann R. The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common UsnRNP proteins. EMBO J. 1994 Sep 1;13(17):4113-25.

-Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010 Jan 28;463(7280):457-63. doi: 10.1038/nature08909. Review.

-Nilsen TW. The spliceosome: no assembly required? Mol Cell. 2002 Jan;9(1):8-9.

-Ohno K, Tsujino A, Brengman JM, Harper CM, Bajzer Z, Udd B, Beyring R, Robb S, Kirkham FJ, Engel AG. Choline acetyltransferase mutations cause myasthenic syndrome associated with episodic apnea in humans. Proc Natl Acad Sci U S A. 2001 Feb 13;98(4):2017-22.

-Ohno M, Segref A, Bachi A, Wilm M, Mattaj IW. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. Cell. 2000 Apr 14;101(2):187-98.

-Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet. 2004 May;5(5):389-96. Review.

-Pagani F, Raponi M, Baralle FE. Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. Proc Natl Acad Sci U S A. 2005 May 3;102(18):6368-72. Epub 2005 Apr 19.

-Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, Baralle FE. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum Mol Genet. 2003 May 15;12(10):1111-20.

-Pagani F, Stuani C, Zuccato E, Kornblihtt AR, Baralle FE. Promoter architecture modulates CFTR exon 9 skipping. J Biol Chem. 2003 Jan 17;278(3):1511-7. Epub 2002 Nov 5.

-Pajares MJ, Ezponda T, Catena R, Calvo A, Pio R, Montuenga LM. Alternative splicing: an emerging topic in molecular and clinical oncology. Lancet Oncol. 2007 Apr;8(4):349-57. Review.

-Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 2008 Dec;40(12):1413-5. doi: 10.1038/ng.259. Epub 2008 Nov 2. Erratum in:Nat Genet. 2009 Jun;41(6):762.

-Park, J.W., Parisky, K., Celotto, A.M., Rennan, R.a, and Graveley, Brenton R (2004). Identification of alternative splicing regulators by RNA interference in Drosophila. Proceedings of the National Academy of Sciences of the Unite States of America 101,15974-9.

-Patel AA, Steitz JA. Splicing double: insights from the second spliceosome. Nat Rev Mol Cell Biol. 2003 Dec;4(12):960-70. Review.

-Patel SB, Novikova N, Bellini M. Splicing-independent recruitment of spliceosomal small nuclear RNPs to nascent RNA polymerase II transcripts. J Cell Biol. 2007 Sep 10;178(6):937-49.

-Pellizzoni L, Yong J, Dreyfuss G. Essential role for the SMN complex in the specificity of snRNP assembly. Science. 2002 Nov 29;298(5599):1775-9.

-Petkovic V, Lochmatter D, Turton J, Clayton PE, Trainer PJ, Dattani MT, Eblé A, Robinson IC, Flück CE, Mullis PE. Exon splice enhancer mutation (GH-E32A) causes autosomal dominant growth hormone deficiency. J Clin Endocrinol Metab. 2007 Nov;92(11):4427-35. Epub 2007 Aug 28.

-Piñol-Roma S, Dreyfuss G. hnRNP proteins: localization and transport between the nucleus and the cytoplasm. Trends Cell Biol. 1993 May;3(5):151-5.

-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 Jun 18;113(25):6461-4. doi: 10.1182/blood-2009-03-207613. Epub 2009 Apr 22.

-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 Mar 1;111(5):2681-4. Epub 2007 Dec 21.

-Pleiss JA, Whitworth GB, Bergkessel M, Guthrie C. Rapid, transcript-specific changes in splicing in response to environmental stress. Mol Cell. 2007 Sep 21;27(6):928-37.

-Plessel G, Fischer U, Lührmann R. m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. Mol Cell Biol. 1994 Jun;14(6):4160-72.

-Pohlenz J, Dumitrescu A, Aumann U, Koch G, Melchior R, Prawitt D, Refetoff S. Congenital secondary hypothyroidism caused by exon skipping due to a homozygous donor splice site mutation in the TSHbeta-subunit gene. J Clin Endocrinol Metab. 2002 Jan;87(1):336-9.

-Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature. 2009 Mar 26;458(7237):475-80. doi: 10.1038/nature07851.

-Raker VA, Hartmuth K, Kastner B, Lührmann R. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. Mol Cell Biol. 1999 Oct;19(10):6554-65.

-Raker VA, Plessel G, Lührmann R. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. EMBO J. 1996 May 1;15(9):2256-69.

-Raker VA, Plessel G, Lührmann R. The snRNP core assembly pathway:identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. EMBO J. 1996 May 1;15(9):2256-69.

-Raponi M, Baralle FE, Pagani F. 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.2007;35(2):606-13. Epub 2006 Dec 15.

-Reed R, Maniatis T. Intron sequences involved in lariat formation during pre-mRNA splicing. Cell. 1985 May;41(1):95-105.

-Reed R. The organization of 3' splice-site sequences in mammalian introns. Genes Dev. 1989 Dec;3(12B):2113-23. PubMed PMID: 2628164. regulator gene. Hum Mol Genet. 2003 Aug 15;12(16):2031-40. PubMed PMID: 12913074.Review.

-Rinke J, Appel B, Blöcker H, Frank R, Lührmann R. 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. 1984 May 25;12(10):4111-26.

-Robberson BL, Cote GJ, Berget SM. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol Cell Biol. 1990 Jan;10(1):84-94.

-Roca X, Olson AJ, Rao AR, Enerly E, Kristensen VN, Børresen-Dale AL, Andresen BS, Krainer AR, Sachidanandam R. Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics. Genome Res. 2008 Jan;18(1):77-87. Epub 2007 Nov 21.

-Roscigno RF, Garcia-Blanco MA. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA. 1995 Sep;1(7):692-706.

-Roscigno RF, Weiner M, Garcia-Blanco MA. A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. J Biol Chem. 1993 May 25;268(15):11222-9.

-Rossi F, Forné T, Antoine E, Tazi J, Brunel C, Cathala G. Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5' splice site-U1 snRNP interaction. J Biol Chem. 1996 Sep 27;271(39):23985-91.

-Ruskin B, Krainer AR, Maniatis T, Green MR. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell. 1984 Aug;38(1):317-31.

-Ruskin B, Zamore PD, Green MR. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell. 1988 Jan 29;52(2):207-19.

-Sadowski CL, Henry RW, Lobo SM, Hernandez N. Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. Genes Dev. 1993 Aug;7(8):1535-48.

-Sahashi K, Masuda A, Matsuura T, Shinmi J, Zhang Z, Takeshima Y, Matsuo M, Sobue G, Ohno K. In vitro and in silico analysis reveals an efficient algorithm to predict the splicing consequences of mutations at the 5' splice sites. Nucleic Acids Res. 2007;35(18):5995-6003. Epub 2007 Aug 28.

-Saltzman AL, Pan Q, Blencowe BJ. Regulation of alternative splicing by the core spliceosomal machinery. Genes Dev. 2011 Feb 15;25(4):373-84. doi: 10.1101/gad.2004811.

-Sanford JR, Ellis J, Cáceres JF. Multiple roles of arginine/serine-rich splicing factors in RNA processing. Biochem Soc Trans. 2005 Jun;33(Pt 3):443-6.

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

-Schaub M, Myslinski E, Schuster C, Krol A, Carbon P. Staf, a promiscuous activator for enhanced transcription by RNA polymerases II and III. EMBO J. 1997 Jan 2;16(1):173-81.

-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 Jul;32(7):815-24. doi: 10.1002/humu.21509.

-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 Jul;32(7):815-24. doi: 10.1002/humu.21509.

-Schneider C, Will CL, Makarova OV, Makarov EM, Lührmann R. Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. Mol Cell Biol. 2002 May;22(10):3219-29.
-Segref A, Mattaj IW, Ohno M. The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export. RNA. 2001 Mar;7(3):351-60.

-Séraphin B, Kretzner L, Rosbash M. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. EMBO J. 1988 Aug;7(8):2533-8.

-Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. 1987 Sep 11;15(17):7155-74.

-Shatkin AJ. Capping of eucaryotic mRNAs. Cell. 1976 Dec;9(4 PT 2):645-53.

-Shepard PJ, Hertel KJ. The SR protein family. Genome Biol. 2009;10(10):242. doi: 10.1186/gb-2009-10-10-242. Epub 2009 Oct 27. Review.

-Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res. 2006;34(14):3955-67. Epub 2006 Aug 12.

-Shi Y, Manley JL. A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock. Mol Cell. 2007 Oct 12;28(1):79-90.

-Siliciano PG, Guthrie C. 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev. 1988 Oct;2(10):1258-67.

-Singh NN, Androphy EJ, Singh RN. An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy. Biochem Biophys Res Commun. 2004 Mar 5;315(2):381-8.

-Singh R, Reddy R. Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. Proc Natl Acad Sci U S A. 1989 Nov;86(21):8280-3.

-Sironi M, Menozzi G, Riva L, Cagliani R, Comi GP, Bresolin N, Giorda R, Pozzoli U. Silencer elements as possible inhibitors of pseudoexon splicing. Nucleic Acids Res. 2004 Mar 19;32(5):1783-91. Print 2004.

-Skoko N, Baralle M, Buratti E, Baralle FE. The pathological splicing mutation c.6792C>G in NF1 exon 37 causes a change of tenancy between antagonistic splicing factors. FEBS Lett. 2008 Jun 25;582(15):2231-6. doi:10.1016/j.febslet.2008.05.018. Epub 2008 May 27.

-Smith CW, Valcárcel J. Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci. 2000 Aug;25(8):381-8. Review.

-Soret J, Bakkour N, Maire S, Durand S, Zekri L, Gabut M, Fic W, Divita G, Rivalle C, Dauzonne D, Nguyen CH, Jeanteur P, Tazi J. Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. Proc Natl Acad Sci U S A. 2005 Jun 14;102(24):8764-9. Epub 2005 Jun 6.

-Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM; IARC Unclassified Genetic Variants Working Group. Prediction and assessment of splicing alterations: implications for clinical testing. Hum Mutat. 2008 Nov;29(11):1304-13. doi: 10.1002/humu.20901.

-Stamm S. Regulation of alternative splicing by reversible protein phosphorylation. J Biol Chem. 2008 Jan 18;283(3):1223-7. Epub 2007 Nov 16. Review.

-Stanek D, Neugebauer KM. The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. Chromosoma. 2006 Oct;115(5):343-54. Epub 2006 Mar 31. Review.

-Stark H, Dube P, Lührmann R, Kastner B. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. Nature. 2001 Jan 25;409(6819):539-42.

-Stefanovic B, Hackl W, Lührmann R, Schümperli D. Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into Xenopus oocytes. Nucleic Acids Res. 1995 Aug 25;23(16):3141-51.
-Stevens SW, Ryan DE, Ge HY, Moore RE, Young MK, Lee TD, Abelson J. Composition and functional characterization of the yeast spliceosomal penta-snRNP. Mol Cell. 2002 Jan;9(1):31-44.

-Sun H, Chasin LA. Multiple splicing defects in an intronic false exon. Mol Cell Biol. 2000 Sep;20(17):6414-25.

-Sun Q, Mayeda A, Hampson RK, Krainer AR, Rottman FM. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev. 1993 Dec;7(12B):2598-608.

-Surowy CS, van Santen VL, Scheib-Wixted SM, Spritz RA. Direct, sequence-specific binding of the human U1-70K ribonucleoprotein antigen protein to loop I of U1 small nuclear RNA. Mol Cell Biol. 1989 Oct;9(10):4179-86.

-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 Sep;24(3):225-35.

-Tardiff DF, Rosbash M. Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. RNA. 2006 Jun;12(6):968-79. Epub 2006 Apr 17.

-Tarn WY, Steitz JA. SR proteins can compensate for the loss of U1 snRNP functions in vitro. Genes Dev. 1994 Nov 15;8(22):2704-17.

-Tazi J, Bakkour N, Stamm S. Alternative splicing and disease. Biochim Biophys Acta. 2009 Jan;1792(1):14-26. doi: 10.1016/j.bbadis.2008.09.017. Epub 2008 Oct 17. Review.

-Tazi J, Kornstädt U, Rossi F, Jeanteur P, Cathala G, Brunel C, Lührmann R. Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. Nature. 1993 May 20;363(6426):283-6.

-Tian M, Maniatis T. A splicing enhancer exhibits both constitutive and regulated activities. Genes Dev. 1994 Jul 15;8(14):1703-12.

-Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcárcel J, Guigó R. Nucleosome positioning as a determinant of exon recognition. Nat Struct Mol Biol. 2009 Sep;16(9):996-1001. doi: 10.1038/nsmb.1658.

-Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat. 2008 Dec;29(12):1412-24. doi:10.1002/humu.20796.

-Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG, Prasanth KV. The nuclear- retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010 Sep 24;39(6):925-38. doi: 10.1016/j.molcel.2010.08.011.

-Turner IA, Norman CM, Churcher MJ, Newman AJ. Roles of the U5 snRNP in spliceosome dynamics and catalysis. Biochem Soc Trans. 2004 Dec;32(Pt 6):928-31. Review.

-Uchikawa H, Fujii K, Kohno Y, Katsumata N, Nagao K, Yamada M, Miyashita T. U7 snRNA-mediated correction of aberrant splicing caused by activation of cryptic splice sites. J Hum Genet. 2007;52(11):891-7. Epub 2007 Sep 13.

-Uguen P, Murphy S. The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. EMBO J. 2003 Sep 1;22(17):4544-54.

-Valadkhan S, Mohammadi A, Jaladat Y, Geisler S. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. Proc Natl Acad Sci U S A. 2009 Jul 21;106(29):11901-6. doi: 10.1073/pnas.0902020106. Epub 2009 Jun 22.

-Van Arsdell SW, Weiner AM. Human genes for U2 small nuclear RNA are tandemly repeated. Mol Cell Biol. 1984 Mar;4(3):492-9.

-van der Woerd WL, Mulder J, Pagani F, Beuers U, Houwen RH, van de Graaf SF. Analysis of aberrant pre-messenger RNA splicing resulting from mutations in ATP8B1 and efficient in vitro rescue by adapted U1 small nuclear RNA. Hepatology. 2015 Apr;61(4):1382-91. doi: 10.1002/hep.27620. Epub 2015 Feb 23.

-Vankeerberghen A, Wei L, Jaspers M, Cassiman JJ, Nilius B, Cuppens H. Characterization of 19 disease-associated missense mutations in the regulatory domain of the cystic fibrosis transmembrane conductance regulator. Hum Mol Genet. 1998 Oct;7(11):1761-9.

-Vidal VP, Verdone L, Mayes AE, Beggs JD. Characterization of U6 snRNA-protein interactions. RNA. 1999 Nov;5(11):1470-81.

-Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009 Feb 20;136(4):701-18. doi: 10.1016/j.cell.2009.02.009. Review.

-Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008 Nov 27;456(7221):470-6. doi: 10.1038/nature07509.

-Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008 Nov 27;456(7221):470-6. doi: 10.1038/nature07509.

-Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet. 2007 Oct;8(10):749-61. Epub 2007 Aug 29. Review.

-Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet. 2007 Oct;8(10):749-61. Epub 2007 Aug 29. Review.

-Wang HY, Lin W, Dyck JA, Yeakley JM, Songyang Z, Cantley LC, Fu XD. SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. J Cell Biol. 1998 Feb 23;140(4):737-50.

-Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB. Systematic identification and analysis of exonic splicing silencers. Cell. 2004 Dec 17;119(6):831-45.

-Ward AJ, Cooper TA. The pathobiology of splicing. J Pathol. 2010 Jan;220(2):152-63. doi: 10.1002/path.2649. Review.

-Wessagowit V, Nalla VK, Rogan PK, McGrath JA. Normal and abnormal mechanisms of gene splicing and relevance to inherited skin diseases. J Dermatol Sci. 2005 Nov;40(2):73-84. Epub 2005 Jul 27. Review.

-Wheeler TM, Leger AJ, Pandey SK, MacLeod AR, Nakamori M, Cheng SH, Wentworth BM, Bennett CF, Thornton CA. Targeting nuclear RNA for in vivo correction of myotonic dystrophy. Nature. 2012 Aug 2;488(7409):111-5. doi: 10.1038/nature11362.

-Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. Science. 2009 Jul 17;325(5938):336-9. doi: 10.1126/science.1173110.

-Whiley PJ, Pettigrew CA, Brewster BL, Walker LC; kConFab Investigators, Spurdle AB, Brown MA. Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts. BMC Med Genet. 2010 May 28;11:80. doi: 10.1186/1471-2350-11-80.

-Will CL, Rümpler S, Klein Gunnewiek J, van Venrooij WJ, Lührmann R. 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. 1996 Dec 1;24(23):4614-23.

-Wu CS, Yu CY, Chuang CY, Hsiao M, Kao CF, Kuo HC, Chuang TJ. Integrative transcriptome sequencing identifies trans-splicing events with important roles in human embryonic stem cell pluripotency. Genome Res. 2014 Jan;24(1):25-36. doi:10.1101/gr.159483.113. Epub 2013 Oct 16.

-Wu JY, Maniatis T. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell. 1993 Dec 17;75(6):1061-70.

-Wu S, Romfo CM, Nilsen TW, Green MR. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature. 1999 Dec 16;402(6763):832-5.

-Yuo CY, Weiner AM. 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. 1989 May;3(5):697-707.

-Zheng ZM, He PJ, Baker CC. Function of a bovine papillomavirus type 1 exonic splicing suppressor requires a suboptimal upstream 3' splice site. J Virol. 1999 Jan;73(1):29-36.

-Zheng ZM, Huynen M, Baker CC. A pyrimidine-rich exonic splicing suppressor binds multiple RNA splicing factors and inhibits spliceosome assembly. Proc Natl Acad Sci U S A. 1998 Nov 24;95(24):14088-93.

-Zhou HL, Lou H. Repression of prespliceosome complex formation at two distinct steps by Fox-1/Fox-2 proteins. Mol Cell Biol. 2008 Sep;28(17):5507-16. doi:10.1128/MCB.00530-08. Epub 2008 Jun 23.

-Zhu J, Mayeda A, Krainer AR. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol Cell. 2001 Dec;8(6):1351-61.
-Zhuang Y, Weiner AM. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell. 1986 Sep 12;46(6):827-35.

-Zuo P, Maniatis T. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes Dev. 1996 Jun 1;10(11):1356-68.

II

Aberrant splicing as cause of Hemophilia B and Seckel syndrome

II.1 COAGULATION FACTOR IX AND HEMOPHILIA B

II.1.1 Blood coagulation and factor IX

The blood-clotting cascade

Blood coagulation is a host defence system that helps protecting the integrity of the closed mammalian circulatory system after blood vessel injury in parallel with the inflammatory and repair responses (Furie and Furie 1992). In invertebrates, the clotting reaction is primarily due to cell aggregation and agglutination. In higher organisms, however, the vascular pressures are high and this increases the risk of bleeding, thus the mechanisms for initiating and regulating blood coagulation in humans are far more complex (Davie 2003). The response to vascular damage initiates with the immediate contraction of blood vessels at the site of injury and culminates in the formation of a platelet plug, stabilized by the generation of a fibrin clot, the deposition of white cells in the area of tissue injury and the beginning of inflammation and repair. This system is normally quiescent because all the proteins and cellular components involved in the process exist under normal physiological conditions in an inactive form, but becomes active within seconds after damage.

The coagulation process was defined in 1964 by Davie and Ratnoff as a "waterfall" or a "cascade" by MacFarlane because from one stage to the next a greater amount of plasma proenzymes is activated to its enzyme form, leading to a final explosion (Davie and Ratnoff 1964; Macfarlane 1964) (Figure II.1).

Blood clotting may be initiated through either the intrinsic pathway, where all of the protein components are present in blood, or the extrinsic pathway, where the cell cell-membrane protein Tissue Factor plays a critical role.

The beginning of the intrinsic pathway involves the activation of Factor XII (FXII) to Factor XIIa (FXIIa), a reaction that is promoted by certain surfaces such as glass or collagen. Although kallicrein is capable of FXII activation, the particular protease responsible for physiologically activation of FXII is unknown. The collagen that becomes exposed in the sub-endothelium after vessel damage may provide the negatively charged surface required for this reaction *in vivo*. FXIIa, in association with its cofactor High Molecular Weight Kininogen (HMWK), converts Factor XI (FXI) to its activated form Factor XIa (FXIa), an unusual serine protease that contains two catalytic sites. In the presence of Ca^{2+} ions, FXIa activates Factor IX (FIX) to Factor IXa (FIXa), which in

complex with Factor VIIIa (FVIIIa) on membrane surfaces catalyses the activation of Factor X (FX) to Factor Xa (FXa). In the presence of Calcium ions and Factor Va (FVa), bound to membrane surfaces, FXa activates prothrombin (PT) to thrombin, that converts fibrinogen to fibrin by cleavage of two peptide bonds, thereby releasing two small amino-terminal peptides, fibrinopeptide A and fibrinopeptide B, and giving rise to fibrin monomers that polymerise spontaneously. The insoluble fibrin network is subsequently stabilized by the transglutaminase action of FXIIIa, activated by thrombin in the presence of Ca^{2+} ions (Furie and Furie 1988).

The extrinsic pathway requires tissue factor (TF), which is located in the tissue adventitia and comes in contact with blood only after vascular injury. TF has a high affinity for Factor VII (FVII) and, in the presence of calcium ions, the two proteins forma one-to-one complex that helps the conversion of FVII to a serine protease (Factor VIIa, FVIIa) by minor proteolysis. This Ca^{2+}-dependent reaction is catalysed by a trace amount of a protease circulating in blood (such as FXa, thrombin, FIXa or FVIIa itself) or by some unidentified plasma or cellular enzyme. The FVIIa-TF complex then converts the membrane-bound FX to FXa by the cleavage of a single peptide bond in the amino-terminal end of the heavy chain, releasing a small activation peptide. The newly generated FXa, in the presence of calcium ions and phospholipid, combines with FVa on the membrane of activated platelets to form a macromolecular complex, also referred to as prothrombinase, which converts PT to thrombin. FVa increases the V_{max} of this reaction about 1000-fold. FXa, in addition, amplifies the procoagulant signal by the feedback activation of TF-FVII and the conversion of profactors FV and FVIII to their active forms. FVIIIa acts as a cofactor of FIXa in the activation of FX within the intrinsic tenase complex (Davie, Fujikawa et al. 1991).

Newly formed thrombin greatly enhances the production of more thrombin by stimulating upstream steps of the coagulation cascade via positive feedback loops. Thrombin is responsible for the massive activation of FV and FVIII. Moreover, in the presence of negatively charged surfaces, such as sulfatide, heparin or dextran sulphate, thrombin activates FXI, which in turn activates FIX in the presence of calcium ions as well as FIX itself. These reactions lead to the late recruitment of the intrinsic pathway, which becomes the quantitatively predominant mechanism for thrombin generation when the extrinsic pathway, which is short-lived, is inhibited by the Tissue Factor Pathway Inhibitor (TFPI). Surprisingly, TFPI does not block the activation of FIX by the TF-FVII complex, thus the activation of FX may continue in part via the intrinsic pathway by the formation of FIXa.

Moreover, when the amount of TF is limiting, the ability of the TF-FVII complex to activate FIX may play a significant role in the initiation of the extrinsic pathway of coagulation because in these conditions FIX appears to be a better substrate than FX.

The intrinsic pathway is clearly important in the clotting of blood *in vitro*, but its physiologic importance has been questioned by the absence of bleeding tendency in patients deficient in FXII, prekallicrein and HMWK. Therefore, the activation of this pathway may be limited to non-physiological conditions, such as exposure to glass or kaolin. The role of FXI is uncertain, as patients with Factor XI deficiency have an extremely variable bleeding tendency.

Although the precise mechanism for the initiation of blood coagulation remains still unclear, it seems probable that the extrinsic pathway plays a dominant physiological role, but the lower part of the intrinsic pathway also contributes to thrombin generation, at least at a later stage (the so called "propagation phase"). The historical dualism of the coagulation process has thus been replaced by an integrated view of the two pathways (Mann 1999; Dahlback 2000).

Platelets, anucleate cells that circulate in the blood in a resting form, adhere at the site of tissue injury upon stimulation induced by vascular damage. Platelets activation includes a series of morphological and biochemical modification promoted by thrombin, thromboxane A_2, ADP, epinephrine and platelet-activating factor (PAF). As a consequence of activation, platelets release the content of two types of granules: α-granules, which contain thrombospondin, fibrinogen, von Willebrand Factor (vWF), FV, growth factors and other proteins involved in haemostasis; and δ-granules, which are rich in Ca^{2+}, Mg^{2+}, ADP, serotonin, histamine. Platelets aggregate to form a plug that reduces or temporarily blocks the loss of blood. The activation of platelets also releases numerous proteins and small molecules that accelerate and increase platelet plug formation and begin the process of tissue repair. Plasma proteins such as vWF play an important role in platelet adhesion, acting as a bridge between the sub-endothelium collagen and specific receptors on activated platelets (i.e. glycoprotein Ib) (Lopez et al. 1988). In a similar manner, fibrinogen forms a bridge between activated platelets by binding to the surface receptors (glycoprotein IIb/IIIa) (Bennett et al. 1982), leading to platelet aggregation and plug formation. These reactions also set the stage for the coagulation cascade and fibrin formation by making available negatively charged phospolipids, such as phosphatidylserine, on the surface of the activated platelets or damaged cell membranes. A series of reactions are then triggered,

leading to fibrin formation and the generation of an insoluble fibrin clot that strengthens the platelet plug.

Vascular damage is always accompanied by inflammation and repair reactions. Thrombin plays a key role in these processes by chemotactically drawing leukocytes to the site of injury and by stimulating tissue remodelling and mitogenesis. P-selectin expressed on the platelet membrane in the haemostatic plug acts as a receptor for monocytes and neutrophils which, in addition to providing ideal membrane surfaces for blood coagulation, sustain the inflammatory response. During wound healing, the fibrin clot in the vessel is degraded by plasmin, a serine protease generated form the plasma zymogen plasminogen, in a process known as fibrinolysis. The activation of plasminogen is catalysed by several plasminogen activators, including two serine proteases, tissue plasminogen activator (t-PA) and urokinase (UK). Their activity is in turn regulated by two plasma protease inhibitors, plasminogen activator inhibitor and α_2-antiplasmin (Furie and Furie 1988).

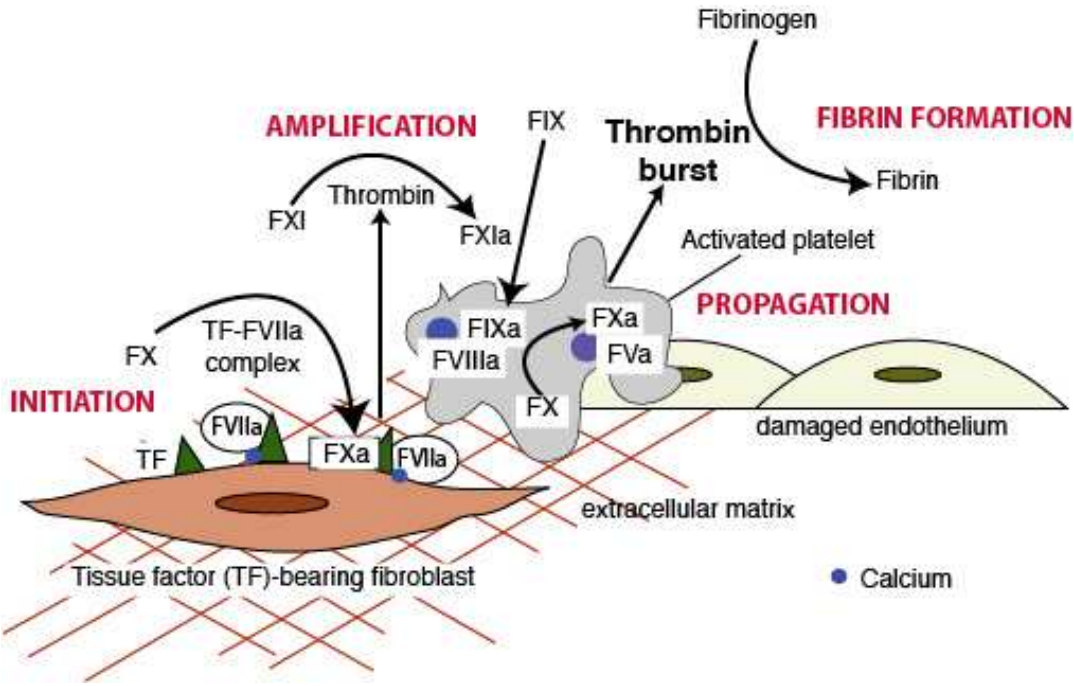


Figure II.1 Cell-based model of coagulation.
A deficiency of coagulation factor IX compromises the coagulation cascade activation , leading to inefficient clot formation and resulting in prolonged bleeding time (haemophilia B).

Blood clotting proteins

Vascular injury triggers the sequential activation of several plasma proteins (coagulation factors), eventually leading to fibrin generation. Coagulation factors comprise enzymes, non-enzymatic cofactors (plasmatic as Factor VIII and Factor V, cellular as Tissue Factor and thrombomodulin) and structural proteins (fibrinogen) (Furie and Furie 1988).

All enzymatic proteins involved in coagulation are vitamin-K dependent serine proteases, which circulate in plasma in a zymogen form (inactive precursors) and are activated by limited proteolysis. These plasma glycoproteins include Factor VII (FVII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), Factor XII (FXII) and prothrombin (PT). In addition to their highly homologous catalytic domains, they share a number of conserved structural motifs that mediate interactions with other proteins and with membranes within macromolecular complexes: a highly conserved Gla-domain required for calcium binding and conformational transitions; one or more EGF domains responsible for interaction with cell surfaces and receptors on other proteins; Kringle structures of uncertain role, probably containing recognition elements important for macromolecular assembly (Furie and Furie 1988; Mann et al. 1990).

The structure and organization of the genes coding for the blood coagulation proteins emphasize that the evolution of new protein function occurs via gene duplication, gene modification and exon shuffling (Gilbert 1978; Patthy 1985). Each of the exons may be considered a module coding for a homologous domain in each protein. The three-dimensional structures of the polypeptide backbones of these homologous domains are likely to be nearly identical, but substitution of amino acid side chains on the protein surface gives definition to unique properties of substrate recognition, cofactor binding, or membrane interaction. The blood coagulation proteins remain a primary example of the development of a family of protein with diverse functional properties but common, unified structural elements.

Coagulation factor V (FV) and factor VIII (FVIII), once activated, function as cofactors of the serine proteases FXa and FIXa, respectively. They do not possess catalytic activity *per se*, but considerably enhance activation reactions by organizing macromolecular enzyme complexes on the surface of phospholipid membranes. FV and FVIII show marked structural and functional similarities (Kane and Davie 1988).

Tissue Factor (TF) and thrombomodulin are integral membrane proteins that do not require proteolytic activation and do not share significant amino acid sequence homology. They only show similarity at the level of the gross structural organization (Mann 1999).

Fibrinogen is composed of two identical monomers, each comprising three polypeptide chains: Aα, Bβ and γ; its structure is unique because each of the three subunits is encoded by a separate gene. During fibrinogen activation, thrombin removes fibrinopeptide A from the Aα chain and fibrinopepitde B from the Bβ chain, unmasking sites for the spontaneous polymerization of the fibrin monomers.

Focus on FIX: gene and protein features

Coagulation factor IX gene (F9) is located on the X chromosome, on the long arm, more towards the centromere at Xq28. The gene is approximately 34 kb in length and contains only eight exons, the largest of which is only 1935 bp. The transcript is 2803 bases in length and comprises a short 5' UTR (29 bp), an open reading frame plus stop codon (1383 bp) and a 3' UTR (1390 bp). The open reading frame encodes a pre-pro-protein in which the signal sequence (pre-sequence) directs factor IX for secretion, the pro-sequence provides a binding domain for vitamin K dependent carboxylase, which carboxylates certain glutamic acid residues in the adjacent Gla domain, and the remainder represents the factor IX zymogen (Anson et al. 1984). Activation of factor IX involves cleavage of two peptide bonds, one on the C-terminal side of arginine 145 (α-cleavage) the other on the C-terminal side of arginine 180 (β-cleavage). These cleavages are caused by activated factor XI generated through the intrinsic pathway or via tissue factor/activated factor VII complex of the extrinsic pathway. The activation cleavages generate an N-terminal light chain and a C-terminal heavy chain, held together by a disulphide bond between cystein residues 132 and 279 (Bowen 2002) (Figure II.2).

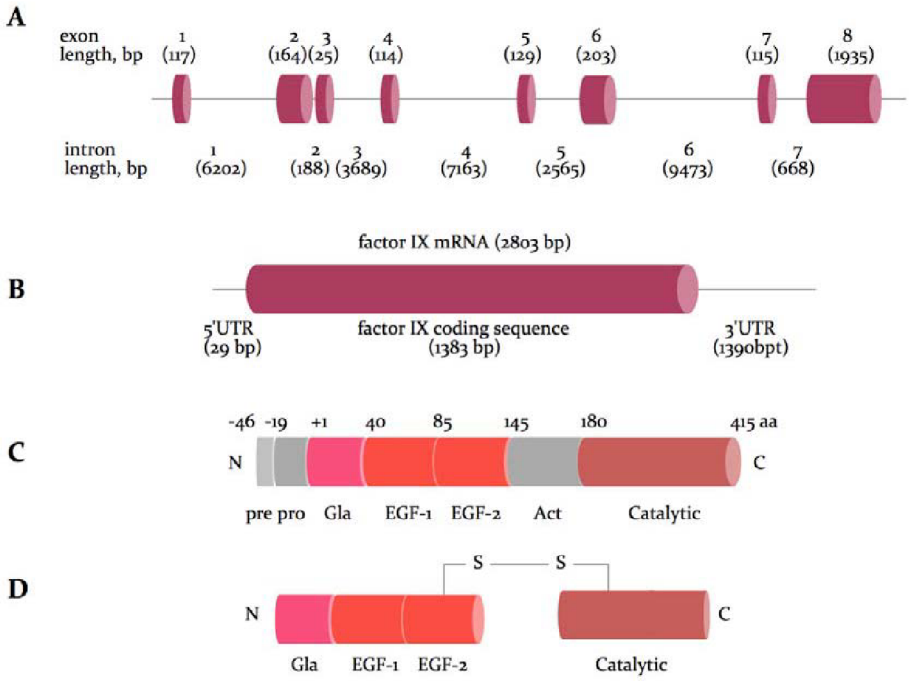


Figure II.2 Schematic representation of F9 gene, mRNA and protein.
A, genomic organization of F9 gene. Exon and intron number and size are reported above and below, respectively.
B, FIX mRNA showing the relative size and location of the ORF
C, the synthetized FIX protein comprising a pre-pro signal sequence and a mature peptide of 45 aminoacids.
D, activated Factor IX comprising a N-terminal light chain and a C-terminal heavy chain held together by a disulphide bond between cysteine residue 132 and 279. Gla, Gla domain; EGF, epidermal growth factor-like domain; Act, activation peptide released after proteolitic cleavage; catalytic, the serine protease domain.

Role of coagulation factor IX in the coagulation cascade

Coagulation factors VIII and IX, whose deficiency are known to cause haemophilia A and B respectively, circulate as inactive precursors that are activated at the time of haemostatic challenge, via the intrinsic or extrinsic pathways (Zdziarska et al. 2009). Factor VIII is a cofactor with no enzymatic activity per se; factor IX is a serine protease with an absolute requirement for factor VIII as cofactor. Upon activation, and in presence of calcium ions and phospholipid surfaces, factor VIII and factor IX form an active complex, which activates factor X. Subsequent stages of the cascade then proceed, culminating in the deposition of fibrin, the structural polymer of the blood clot (Bowen 2002).

II.1.2 Hemophilia B

Hemophilia B (or Christmas morbus) is a coagulopathy X-linked caused by mutations in the F9 gene (incidence of 1:35000 live male births). Based on FIX levels (antigen and/or protein activity), patients experience severe hemorrhagic symptoms, not rarely life-threatening (central nervous system and gastrointestinal bleeds), or causing substantial handicap (hemarthrosis, muscle hematoma). Based on FIX levels on plasma, the phonotype of patients is classified as mild, moderate or severe:

- Mild , when the FIX level is included between 6 or 25% (0,05-0,40 IU/mL)
- Moderate, if FIX level is in 1-5% range (or 0,01-0,05 IU/mL)
- Severe , when the FIX level is below 1% or <0,01 IU/mL

Generally mild and moderate patients do not suffer of spontaneous hemorrhage, even if they can experience life threatening hemorrhage during surgery if not properly treated. The mutations causing haemophilia B have been localized and characterized in several hundreds of patients. Based on the enormous number of mutations that have been elucidated it is clear now that the molecular basis of haemophilia are extremely different. Among all mutations, missense mutations are the most common (68%), followed by non-sense mutations (14%). Mutations altering splicing sequences have been found in 9% of all patients, with frame-shift mutation, promoter located and in frame deletion mutations ranging in 5%, 3% and 1% respectively. Point mutations (single nucleotide substitutions) are the most common gene defect and are present in approximately 90% of patients. Deletions are the second most common gene defects are present in approximately 5-10% of patients. Insertions and other rearrangements are quite rare within the haemophilia B population (Bowen 2002). The point mutations that occur in haemophilia B comprise missense mutation (these change a codon so that a different aminoacid is encoded), nonsense point mutation (these change an aminoacid codon into a translation stop codon), and mRNA splice site point mutations (these corrupt a true mRNA splice site, or create a novel one) (Koeberl et al. 1990; Ketterling, Drost et al. 1999).

In particular mutations that destroy or create mRNA splice sites are associated with variable severity of haemophilia: this depends on whether some correct transcripts can be processed (mild to moderate disease) or whether there is a complete loss of correct mRNA processing (severe disease). Exon skipping is a possible consequence of a mutation affecting splicing: the outcome depends on whether the skip is in frame or results in a frame shift (Tavassoli et al. 1998).

In haemophilia approximately 30% of mutations involves a CpG site; the remaining 70% of distinct point mutations do not occur a CpG sites and may arise, for example, as a result of nucleotide misincorporation during DNA replication (Bowen 2002).

In general nonsense mutations are associated with severe forms of haemophilia; exon skipping is a further possibility arising from a nonsense mutation and it is also extremely detrimental: an in frame skip will result in a protein lacking the aminoacids encoded by skipped exon, an out of frame skip will result in a frame shift (Dietz et al. 1993; Ketterling et al. 1999)(Figure II.3).

Deletions of F9 gene include whole gene deletions, partial gene deletions at 5' or 3' end or within the gene, and microdeletions of one to several base pairs. A deletion, in general, has a high probability of destroying genetic function, removing domains of a protein, or introducing a frame shift, all of which are extremely detrimental. Therefore is not surprisingly that deletion are associated with severe forms of the disease (Cooper 1991; Giannelli and Green 1996).

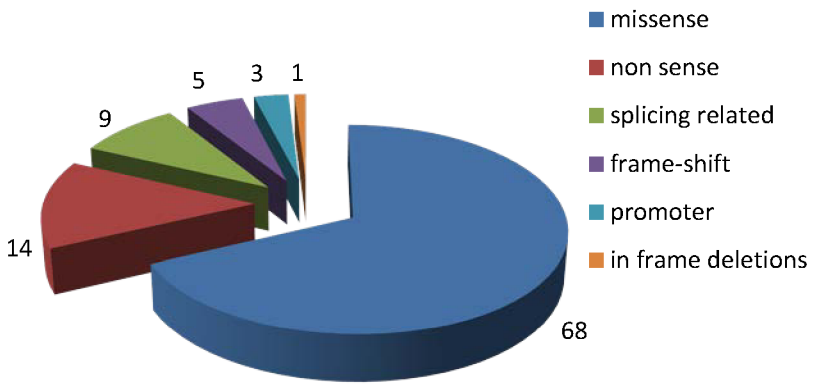


Figure II.3 Pie chart showing the type of mutations reported in the Hemophilia B Mutation Database

Treatments for Haemophilia B

The administration of FIX, either recombinant or plasma derived, represents the current therapy for hemophilia B. Various drawbacks have been associated with prolonged infusion of FIX concentrates: disseminated intravascular coagulation, pulmonary embolus, and deep venous thrombosis. Moreover, the administration of a protein that the body is not able to synthetize by itself results, in a long term therapy, in the development of neutralizing FIX antibodies. The costs of substitutive treatments (~50000 euro/year/person with severe disease) are prohibitive for the majority of the world hemophilia populations, so the demand for alternative therapies is growing up.

In this context, several innovative approaches are being pursued to extend the half-life of factor VIIa, factor VIII and factor IX, utilizing technologies such as Fc fusion, recombinant albumin fusion and addition of polyethyleneglycol (PEG) (PEGylation). These methods prolong the half-life of factor infused and by reducing degradation and clearance from circulation (Mannucci PM 2015), thus reducing the number of infusions and the cost of the therapy.

Other strategies under-development for Hemophilia B

Enormous efforts have been pushed on the replacement gene therapy that consists in the 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 (Murphy and High 2008; Petrus et al. 2010).

Hemophilia B is regarded as an ideal target disease for a gene therapy strategy since the limited FIX coding size makes it packeageble in virtually any viral vectors, and a modest increase (up to 5%) would convert the disease phenotype from severe to mild.

The first gene therapy trial for Hemophilia B was carried out in China using an onco-retroviral vector to transduce autologous human skin fibroblasts with a FIX construct, which was subsequently injected into patients (Qui X et al. 1996). An increase to 2% of normal FIX activity was observed over a year after the injections, and partial correction of the bleeding phenotype was seen. However, concerns regarding the mutagenic insertion of retroviral vector limit the applicability of such approach.

Based on successful phenotypic correction of a canine Hemophilia B model with an adeno-associated vector (AAV)(Herzog RW et al. 1999), new gene replacement therapy trials have been started. In 2011, the results from a dose-escalation trial, in which therapeutic

levels of FIX were achieved from a single intravenous administration of an AAV8 vector, were reported. This was the first example of sustained therapeutic expression of FIX in human patients (Nathwani AC et al. 2011). A follow-up study showed that these patients, along with a new cohort of injected patients, maintained circulating FIX levels at >3 years post-gene transfer (Nathwani AC et al. 2014). It is now almost five years since these patients received this gene therapy treatment, and all have retained circulating FIX levels of between 1 and 5% of normal.

Monogenic diseases, including Hemophilia, represent also ideal targets for genome-editing approaches aimed at correcting a defective gene. Recently, the delivery of zinc finger nucleases (ZFNs) by AAV enables the production of high levels of human factor IX in a murine model of Hemophilia B (Anguela XM et al. 2013).

II. ATR AND SECKEL SYNDROME

II.1.1 The DNA damage response and ATR

The DNA damage response

Genome maintenance is critical to prevent disease. Challenges to genome integrity come from environment mutagens, by products of cellular respirations, and errors during nucleic acid metabolism including DNA replication. Cells have DNA damage response (DDR) activities that continually monitor the integrity of the DNA and function to prevent the occurrence of deleterious mutations and rearrangements. The DDR is regulated by the phosphoinositide-3-kinase-related protein kinases (PIKKs). The PIKKs primarily responsible for signaling the presence of DNA damage include ATM, ATR and DNA-PKcs. These PIKKs phosphorylate hundreds of protein that maintain genome integrity through regulation of cell cycle progression, DNA repair apoptosis, and cellular senescence.

ATR and ATM have a crucial role in DDR

The maintenance of genomic integrity is crucial for the survival of all organisms. In humans, compromised genomic integrity contributes to genetic disorders, aging and cancers. The task of safeguarding the genome is accomplished by the concerted action of a number of cellular processes, including DNA replication, DNA repair, senescence and apoptosis. Many, if not all, of these processes are regulated and coordinated by the DNA-damage checkpoint, which is a complex signaling network that is triggered by DNA damage or genomic instability. Two phospho-inositide 3-kinase-like protein kinases (PIKKs) – ataxia telangiectasia mutated (ATM), and ATM- and Rad3-related (ATR) – are master regulators of two major checkpoint pathways (Harper and Elledge, 2007). ATM and ATR are expressed in most tissues, and mutations in the encoding genes result in the autosomal recessive disorders ataxia telangiectasia (Lavin, 2008) and Seckel syndrome (O'Driscoll et al., 2003), respectively. The clinical manifestations of ataxia telangiectasia include progressive ataxia, telangiectasia, immune defects, genome instability and malignancy; Seckel syndrome is characterized by postnatal dwarfism, microcephaly, intrauterine growth defects and mental retardation. ATM and ATR belong to the class-IV phosphoinositide 3-kinase (PI3K)-related kinase (PIKK) family, along with mammalian target of rapamycin (mTOR) and DNA-dependent protein kinase (DNA-PK). ATM and ATR nonetheless lack lipid kinase activity; instead, they phosphorylate proteins that

contain Ser or Thr residues that are followed by Gln (SQ or TQ motifs) (Bakkenist and Kastan, 2004). DDR is mediated by ATM and ATR, as well as by two downstream kinases, checkpoint kinases 1 and 2 (Chk1 and Chk2; encoded by CHEK1 and CHEK2, respectively). After their activation, both ATM and ATR upregulate cell cycle checkpoint pathways, inducing cell cycle arrest and DNA repair. ATM and ATR respond to different types of DNA lesions, to which they are recruited through specific co-factors; ATM responds primarily to DNA double-strand breaks (DSBs) (Paull, 2015), whereas ATR protects the integrity of replicating chromosomes (Branzei and Foiani, 2008). ATR is also activated by DSBs, however, through a mechanism that depends on ATM and the MRE11–RAD50–NBS1 (MRN) complex (Jazayeri et al. 2006; Doksani et al. 2009). Recent studies suggest that ATM and ATR also act in response to other cellular stresses, and that they might control cell pathways that do not converge on DNA repair mechanisms but that instead maintain cell homeostasis.

ATM and ATR are members of the PIKK family

ATM and ATR belong to the PIKK family together with mTOR, human suppressor of morphogenesis in genitalia-1 (SMG-1), DNA-PK catalytic subunit (DNAPKcs; encoded by PRKDC), and transformation/transcription-associated protein (TRRAP). PIKK-family members share similarity in their kinase domains with the catalytic loops of class-I PI3K and are therefore categorized separately from the classic protein kinases. In addition, PIKK enzymes have a conserved FRAP-ATM-TRRAP (FAT), PIKK-regulatory domain (PRD) and FAT carboxy-terminal (FATC) domains, as well as an N-terminus that bears numerous α-helical Huntingtin, elongation factor 3, protein phosphatase 2A, TOR1 (HEAT) repeat motifs (Lovejoy and Cortez, 2009). The FAT domain is at the N-terminus of the kinase domain, whereas the FATC domain lies in its C-terminus. The PRD, FAT and FATC regions are crucial for ATR and ATM activation, and mutations in these regions can hamper their kinase activity (Bakkenist and Kastan, 2003; Liu et al. 2011; Mordes and Cortez, 2008; Nam and Cortez, 2011; Sun et al. 2005).

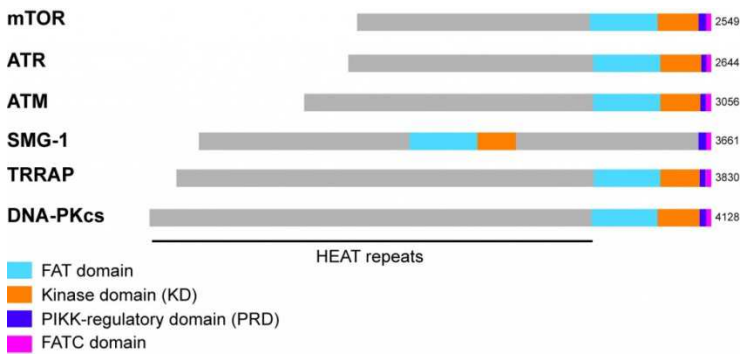
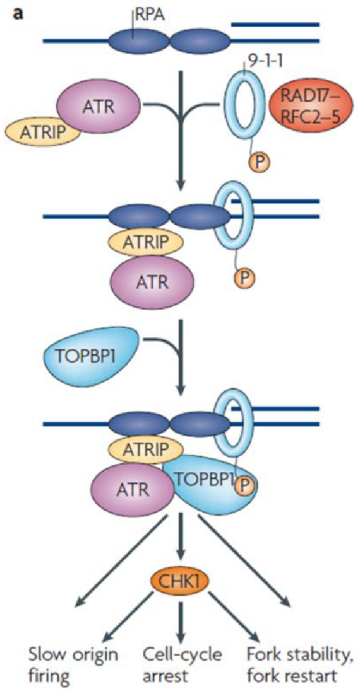


Figure II.4 Member of the PIKK family

The N-terminal HEAT repeats act as a scaffold; they mediate the interactions of ATM and ATR with proteins that regulate their catalytic activity and have an important role in their stability (Perry and Kleckner, 2003). The HEAT repeats can also act as elastic connectors that undergo deformation following mechanical stimulation and regulate protein activity (Grinthal et al. 2010). As is the case for other PI3K-family members, PIKKs exist as homo or heterodimers, and dimerization influences their stability and kinase activity. Accordingly, ATM forms dimers or oligomers under non-stress conditions and it is released in monomeric form following induction of stress (Bakkenist and Kastan, 2003). ATR forms a heterodimer with its obligatory partner ATR-interacting protein (ATRIP); this heterodimerization stabilizes ATR, although the interaction with ATRIP does not appear to be very strong.

ATR activation

ATR is activated in response to a variety of DNA lesions that induce the formation of single-strand (ss)DNA (Cimprich and Cortez, 2008; Zou and Elledge, 2003). ATR activation is a multistep process, because the ATR–ATRIP heterodimer is unable to interact with DNA directly, and depends on nucleofilaments that are formed between the replication protein A heterotrimer (RPA) and ssDNA for DNA binding.



FigureII.5. ATR pathway

ATR associated ATRIP interacts directly with ssDNA-bound RPA and so promotes ATR localization to sites of replication stress and DNA damage (Zou and Elledge, 2003). It has recently been suggested that pre-mRNA processing factor 19 (PRP19)-assisted ubiquitylation of RPA facilitates ATRIP binding to damaged DNA (Maréchal et al. 2014), which leads to partial ATR activation. The RPA-assisted ATRIP–ATR complex then interacts with the DNA-damage-specific RAD9– RAD1–HUS1 clamp (also called 9-1-1) that is bound at junctions between ssDNA and double-strand (ds)DNA (Ellison and Stillman, 2003). 9-1-1 itself is loaded onto the ssDNA–dsDNA junction by the clamp loader complex, RAD17–RFC, which is facilitated by RPA (Ellison and Stillman, 2003; Zou et al. 2003). This is followed by phosphorylation of the 9-1-1 subunit RAD9 on residue S387, which enables the association of DNA topoisomerase 2-binding protein 1 (TopBP1) with the FATC domain of ATR, leading to ATR activation (Choi et al. 2010). Furthermore, TopBP1 can interact with both phosphorylated or unphosphorylated 9-1-1 complexes, and, moreover, its interaction with 9-1-1 can be mediated by other proteins such as RHINO (Rad9, Rad1, Hus1 interacting nuclear orphan, encoded by RHNO1) (Cotta-Ramusino et al. 2011). A recent study has confirmed the function of RHINO in ATR activation following genotoxic stress, but also showed that RHINO is dispensable for the interaction of TopBP1 with the 9-1-1 complex (Lindsey-Boltz et al. 2015). Importantly, the ATR-activation domain in TopBP1 can also mediate ATR activation in the absence of

DNA damage or other known ATR activators, and so can initiate checkpoint signaling in the absence of DNA damage (Toledo et al. 2008). Although the current notion is that ATR is activated by RPA–ssDNA nucleofilaments, some studies appear to contradict this model because they have found RPA to be dispensable for ATR activation and subsequent Chk1 phosphorylation following genotoxic stress (Ball et al. 2005; Dodson et al. 2004). Moreover, the phenotypes of cells that have been depleted of RAD9 and of HUS1-depleted cells do not fully recapitulate those of cells with defective ATR signaling (Hopkins et al., 2004; Weiss et al., 2000), suggesting that RAD9 and HUS1 might be redundant for some of the ATR functions, or that ATR has other functions that do not require the 9-1-1 complex.

Signaling outputs of ATR and ATM

Phosphorylation of downstream-acting proteins is at the heart of ATR and ATM signaling in response to cell stress events. Both kinases phosphorylate proteins on S/TQ motifs and initiate widespread cell responses through phosphorylation of downstream effector proteins. Much of the knowledge gained with regard to ATR and ATM signaling cascades has been derived from the treatment of various cell types with high doses of DNA-damaging agents, which result in a burst of ATR and ATM activity, and downstream signaling. These studies might, nonetheless, underestimate the impact of ATR and ATM mediated phosphorylation events that are elicited under more physiological conditions, particularly in light of the accumulating evidence that suggests that ATM and ATR mediated pathways are not restricted to nuclear events and DDR. Screens for S/TQ-containing substrates of ATR and ATM following DNA damage have also revealed putative substrates outside of the nucleus, as well as factors that are not directly linked to DNA repair processes (Matsuoka et al. 2007; Mu et al. 2007; Paulsen et al. 2009), although the functional relevance of many of these phosphorylation events remains to be determined. Intriguingly, DNA-PK has also been located at the Golgi, reinforcing the possibility that other PIKKs might also function outside of the nucleus (Farber-Katz et al. 2014).

ATR in replication stress

Activation of the ATR–ATRIP complex (which is loaded onto DNA together with TopBP1) initiates a signaling cascade that coordinates cell cycle progression with DNA metabolic processes. ATR activity is necessary both for the stabilization of stalled

replication forks and for fork restart following replication stress. On one hand, ATR inhibition results in the increased firing of origins in the absence of DNA damage, but on the other hand, under replication stress, ATR-dependent phosphorylation of `Fanconi anemia, complementation group I' (FANC1) inhibits the firing of dormant origins (Chen et al. 2015b). In addition, ATR deficiency in aphidicholine-treated cells causes incomplete replication of regions with fragile sites (Casper et al. 2002; Paulsen and Cimprich, 2007). When a replication fork encounters a gene that is being transcribed, the Mec1/ATR pathway phosphorylates nucleoporin components to release the transcribed chromatin, which is attached to the nuclear envelope (Bermejo et al. 2011), thereby preventing torsional-stress induced fork reversal. Hence, the ATR response ensures robust replication through different means – by stabilizing replication forks, preventing fragile-site expression, influencing replication origin firing (Chen et al. 2015b; Shechter et al. 2004) and also coordinating replication with transcription (Bermejo et al. 2011), as well as by triggering the replication stress response (Flynn and Zou, 2011). Under conditions of replication stress, a balance between the amounts of RPA and ssDNA appears to be crucial for the stability of the replication fork. Accordingly, ssDNA uncoating owing to RPA paucity results in collapse of the replication fork and the generation of DSBs (Toledo et al. 2013).

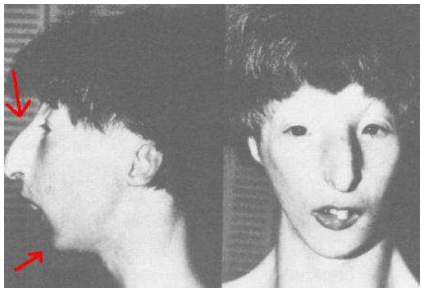
ATR in cell cycle regulation

ATR also influences DNA repair processes, such as the repair of DSBs, nucleotide excision repair (NER) and inter-strand crosslink repair (ICL). The best characterized ATR effector, to date, is Chk1, which is activated through ATR-mediated phosphorylation at residues S317 and S34 in a reaction that is stimulated by claspin binding to, by the 9-1-1 complex and RHINO, as well as by other factors. Chk1 activation stabilizes the chromatin-bound CDC7–DBF4 apoptosis signal regulating kinase (ASK) complex, which assists in replication and origin firing (Yamada et al. 2013). Chk1 affects progression through S phase at the level of origin firing, replication elongation and fork integrity (Brown and Baltimore, 2003; Heffernan et al. 2007; Petermann et al. 2006, 2010; Segurado and Diffley, 2008; Zhao et al. 2002). In addition, activated Chk1 also has an effect on S-phase progression through phosphorylation of CDC25 and regulation of cyclin-dependent kinases (CDKs) (Sanchez et al. 1997). The ATR-mediated Chk1 pathway also has a central role in preventing cells from entering into mitosis with unreplicated or damaged DNA (Brown and Baltimore, 2003). Chk1-dependent sequestration of CDC25C into the cytoplasm and

degradation of CDC25A maintains CDK1 in its inactive state, resulting in G2/M arrest (Mailand et al. 2002; Nghiem et al. 2001; Peng et al. 1997; Sanchez et al. 1997). Cells that express kinase-dead ATR mutants undergo premature chromosome condensation and entry into mitosis (Nghiem et al. 2001). Other ATR targets, such as the helicase SMARCAL1, whose fork regression activity is compromised after phosphorylation by ATR, thereby preventing aberrant fork processing, thus have important roles in the response of ATR to replication stress (Couch et al. 2013). ATR is also activated in G1 following γ-irradiation in order to facilitate DNA repair (Gamper et al. 2013). It has also been suggested that ATR has a role in the physical separation of cells during cytokinesis through Chk1-mediated phosphorylation of Aurora B (encoded by AURKB) (Mackay and Ullman 2015), a known mediator of furrow cleavage that promotes cytokinesis (Marumoto et al. 2005). Recently it has been shown that ATR responds to osmotic and mechanical stress, and that cells derived from individuals with Seckel syndrome fail to coordinate chromatin condensation with nuclear envelope breakdown (Kumar et al. 2014). This process does not appear to involve DNA-damage sensing and might be mediated through the elastic properties of the N-terminal ATR HEAT repeats, which could act as mechanosensors (Grinthal et al. 2010).

II.2.2 SECKEL SYNDROME

Seckel syndrome is an extremely rare inherited disorder characterized by growth delays prior to birth (intrauterine growth retardation) resulting in low birth weight. Growth delays continue after birth (postnatal), resulting in short stature (dwarfism). Other symptoms and physical features associated with Seckel syndrome include an abnormally small head (microcephaly), varying degrees of mental retardation and/or unusual characteristic facial features including “beak-like” protrusion of the nose.



FigureII.6 Seckel Syndrome features (E. Thompson 1985)

SS is due to a defective DNA damage response (Kerzendorfer and O' Driscoll 2009) that mainly affects neuronal development. SS is genetically heterogeneous with different independent loci identified (O'Driscoll et al. 2003; Qvist et al. 2011; Kalay et al. 2011; Ogi et al. 2012). Three variants of Seckel syndrome involve disruptions or changes (mutations) of genes on three different chromosomes. The gene map locations are: Seckel syndrome 1, on chromosome 3 (3q22-q24); Seckel syndrome 2, on chromosome 18 (18p11.31-q11) and Seckel syndrome 3, on chromosome 14 (14q21-q22). Seckel syndrome-1 (SS-1) is caused by mutations in the gene for the ataxia-telangiectasia and RAD3-related protein (ATR; 601215)(O'Driscoll et a. 2003; Ogi et al. 2012). The genes involved in Seckel syndrome types 2 and 3 are unknown.

The ATR 2101a/g mutation has been the first homozygous mutation identified in SS-1 patients (O'Driscoll et al. 2003). It has been demonstrated that this synonymous Gly674Gly mutation in exon 9 mainly induces exon skipping, thus predicting a frame-shift effect. However, trace levels of correct ATR mRNA were also detected, which explains patients survival. The causative nature of this mutation has been demonstrated in mouse models harboring the humanized ATR gene (Murga et al. 2009).

Recently other mutations, including a 540 kbp deletion encompassing the ATR gene and a missense mutation located in exon 33, were identified in patient with Seckel syndrome.

Affected Populations

Seckel syndrome is an extremely rare inherited disorder that appears to affect males and females in equal numbers. The exact incidence of this disorder is now known. More than 100 cases have been reported in the medical literature since its original description in 1960.

Standard Therapies

With the advent of technically superior ultrasonography, Seckel syndrome may be diagnosed before birth (prenatally). In fetal ultrasonography, reflected sound waves are used to create an image of the developing fetus. After birth, Seckel syndrome may be suspected based upon a thorough clinical evaluation, a detailed patient history, and a variety of specialized tests. Although distinctive craniofacial, skeletal, and/or other physical abnormalities associated with Seckel syndrome may be present at birth (congenital), a diagnosis of Seckel syndrome may not be confirmed, in some cases, until an affected child ages and the full syndrome develops. Short stature associated with Seckel syndrome involves proportional growth of the arms and legs, which allows for differential

diagnosis from syndromes that involve short stature and abnormally small arms and legs (short-limbed dwarfism).

REFERENCES

-Anguela XM, Sharma R, Doyon Y, Miller JC, Li H, Haurigot V, Rohde ME, Wong SY, Davidson RJ, Zhou S, Gregory PD, Holmes MC, High KA. Robust ZFN-mediated genome editing in adult hemophilic mice. Blood. 2013 Nov 7;122(19):3283-7. doi: 10.1182/blood-2013-04-497354. Epub 2013 Oct 1.

-Anson DS, Choo KH, Rees DJ, Giannelli F, Gould K, Huddleston JA, Brownlee GG. The gene structure of human anti-haemophilic factor IX. EMBO J. 1984 May;3(5):1053-60.

-Bakkenist CJ, Kastan MB. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature. 2003 Jan 30;421(6922):499-506.

-Bakkenist CJ, Kastan MB. Initiating cellular stress responses. Cell. 2004 Jul 9;118(1):9-17. Review.

-Ball HL, Myers JS, Cortez D. ATRIP binding to replication protein A-single-stranded DNA promotes ATR-ATRIP localization but is dispensable for Chk1 phosphorylation. Mol Biol Cell. 2005 May;16(5):2372-81. Epub 2005 Mar 2.

-Bennett JS, Vilaire G, Cines DB. Identification of the fibrinogen receptor on human platelets by photoaffinity labeling. J Biol Chem. 1982 Jul 25;257(14):8049-54.

-Bermejo R, Capra T, Jossen R, Colosio A, Frattini C, Carotenuto W, Cocito A, Doksani Y, Klein H, Gómez-González B, Aguilera A, Katou Y, Shirahige K, Foiani M. The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. Cell. 2011 Jul 22;146(2):233-46. doi:10.1016/j.cell.2011.06.033.

-Bowen DJ. Haemophilia A and haemophilia B: molecular insights. Mol Pathol. 2002 Apr;55(2):127-44. Review. Erratum in: Mol Pathol 2002 Jun;55(3):208.

-Branzei D, Foiani M. Regulation of DNA repair throughout the cell cycle. Nat Rev Mol Cell Biol. 2008 Apr;9(4):297-308. doi: 10.1038/nrm2351. Epub 2008 Feb 20. Review.

-Brown EJ, Baltimore D. Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. Genes Dev. 2003 Mar 1;17(5):615-28.

-Casper AM, Nghiem P, Arlt MF, Glover TW. ATR regulates fragile site stability. Cell. 2002 Dec 13;111(6):779-89.

-Chen YH, Jones MJ, Yin Y, Crist SB, Colnaghi L, Sims RJ 3rd, Rothenberg E, Jallepalli PV, Huang TT. ATR-mediated phosphorylation of FANCI regulates dormant origin firing in response to replication stress. Mol Cell. 2015b Apr 16;58(2):323-38. doi: 10.1016/j.molcel.2015.02.031. Epub 2015 Apr 2.

-Choi JH, Lindsey-Boltz LA, Kemp M, Mason AC, Wold MS, Sancar A. Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling. Proc Natl Acad Sci U S A. 2010 Aug 3;107(31):13660-5. doi: 10.1073/pnas.1007856107. Epub 2010 Jun 28.

-Cimprich KA, Cortez D. ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol. 2008 Aug;9(8):616-27. doi: 10.1038/nrm2450. Epub 2008 Jul 2. Review.

-Cooper DN. The molecular genetics of familial venous thrombosis. Blood Rev. 1991 Mar;5(1):55-70. Review.

-Couch FB, Bansbach CE, Driscoll R, Luzwick JW, Glick GG, Bétous R, Carroll CM, Jung SY, Qin J, Cimprich KA, Cortez D. ATR phosphorylates SMARCAL1 to prevent replication fork collapse. Genes Dev. 2013 Jul 15;27(14):1610-23. doi:10.1101/gad.214080.113.

-Dahlbäck B. Blood coagulation. Lancet. 2000 May 6;355(9215):1627-32. Review.

-Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance, and regulation. Biochemistry. 1991 Oct 29;30(43):10363-70. Review.

-DAVIE EW, RATNOFF OD. WATERFALL SEQUENCE FOR INTRINSIC BLOOD CLOTTING. Science. 1964 Sep 18;145(3638):1310-2.

-Davie EW. A brief historical review of the waterfall/cascade of blood coagulation. J Biol Chem. 2003 Dec 19;278(51):50819-32. Epub 2003 Oct 21.

-Dietz HC, Valle D, Francomano CA, Kendzior RJ Jr, Pyeritz RE, Cutting GR. The skipping of constitutive exons in vivo induced by nonsense mutations. Science. 1993 Jan 29;259(5095):680-3.

-Dodson GE, Shi Y, Tibbetts RS. DNA replication defects, spontaneous DNA damage, and ATM-dependent checkpoint activation in replication protein A-deficient cells. J Biol Chem. 2004 Aug 6;279(32):34010-4. Epub 2004 Jun 14.

-Doksani Y, Bermejo R, Fiorani S, Haber JE, Foiani M. Replicon dynamics, dormant origin firing, and terminal fork integrity after double-strand break formation. Cell. 2009 Apr 17;137(2):247-58. doi: 10.1016/j.cell.2009.02.016. Epub 2009 Apr 9.

-Ellison V, Stillman B. Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA. PLoS Biol. 2003 Nov;1(2):E33. Epub 2003 Nov 17.

-Farber-Katz SE, Dippold HC, Buschman MD, Peterman MC, Xing M, Noakes CJ, Tat J, Ng MM, Rahajeng J, Cowan DM, Fuchs GJ, Zhou H, Field SJ. DNA damage triggers Golgi dispersal via DNA-PK and GOLPH3. Cell. 2014 Jan 30;156(3):413-27. doi:10.1016/j.cell.2013.12.023.

-Flynn RL, Zou L. ATR: a master conductor of cellular responses to DNA replication stress. Trends Biochem Sci. 2011 Mar;36(3):133-40. doi:10.1016/j.tibs.2010.09.005. Epub 2010 Oct 12. Review.

-Furie B, Furie BC. Molecular and cellular biology of blood coagulation. N Engl J Med. 1992 Mar 19;326(12):800-6. Review.

-Gamper AM, Rofougaran R, Watkins SC, Greenberger JS, Beumer JH, Bakkenist CJ. ATR kinase activation in G1 phase facilitates the repair of ionizing radiation-induced DNA damage. Nucleic Acids Res. 2013 Dec;41(22):10334-44. doi:10.1093/nar/gkt833. Epub 2013 Sep 14.

-Giannelli F, Green PM. The molecular basis of haemophilia A and B. Baillieres Clin Haematol. 1996 Jun;9(2):211-28. Review.

-Gilbert W. Why genes in pieces? Nature. 1978 Feb 9;271(5645):501.
-Grinthal A, Adamovic I, Weiner B, Karplus M, Kleckner N. PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. Proc Natl Acad Sci U S A. 2010 Feb 9;107(6):2467-72. doi:10.1073/pnas.0914073107. Epub 2010 Jan 25.

-Harper JW, Elledge SJ. The DNA damage response: ten years after. Mol Cell. 2007 Dec 14;28(5):739-45. Review.

-Heffernan TP, Unsal-Kaçmaz K, Heinloth AN, Simpson DA, Paules RS, Sancar A, Cordeiro-Stone M, Kaufmann WK. Cdc7-Dbf4 and the human S checkpoint response to UVC. J Biol Chem. 2007 Mar 30;282(13):9458-68. Epub 2007 Feb 2.

-Herzog RW, Yang EY, Couto LB, Hagstrom JN, Elwell D, Fields PA, Burton M, Bellinger DA, Read MS, Brinkhous KM, Podsakoff GM, Nichols TC, Kurtzman GJ, High KA. Long-term correction of canine hemophilia B by gene transfer of blood coagulation factor IX mediated by adeno-associated viral vector. Nat Med. 1999 Jan;5(1):56-63.

-Hopkins KM, Auerbach W, Wang XY, Hande MP, Hang H, Wolgemuth DJ, Joyner AL, Lieberman HB. Deletion of mouse rad9 causes abnormal cellular responses to DNA damage, genomic instability, and embryonic lethality. Mol Cell Biol. 2004 Aug;24(16):7235-48.

-Jazayeri A, Falck J, Lukas C, Bartek J, Smith GC, Lukas J, Jackson SP. ATM-and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. Nat Cell Biol. 2006 Jan;8(1):37-45. Epub 2005 Dec 4.

-Kalay E, Yigit G, Aslan Y, Brown KE, Pohl E, Bicknell LS, Kayserili H, Li Y, Tüysüz B, Nürnberg G, Kiess W, Koegl M, Baessmann I, Buruk K, Toraman B, Kayipmaz S, Kul S, Ikbal M, Turner DJ, Taylor MS, Aerts J, Scott C, Milstein K, Dollfus H, Wieczorek D, Brunner HG, Hurles M, Jackson AP, Rauch A, Nürnberg P, Karagüzel A, Wollnik B. CEP152 is a genome maintenance protein disrupted in Seckel syndrome. Nat Genet. 2011 Jan;43(1):23-6. doi: 10.1038/ng.725. Epub 2010 Dec 5.

-Kane WH, Davie EW. Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders. Blood. 1988 Mar;71(3):539-55. Review.

-Kerzendorfer C, O'Driscoll M. Human DNA damage response and repair deficiency syndromes: linking genomic instability and cell cycle checkpoint proficiency. DNA Repair (Amst). 2009 Sep 2;8(9):1139-52. doi: 10.1016/j.dnarep.2009.04.018. Epub 2009 May 26. Review.

-Ketterling RP, Drost JB, Scaringe WA, Liao DZ, Liu JZ, Kasper CK, Sommer SS. Reported in vivo splice-site mutations in the factor IX gene: severity of splicing defects and a hypothesis for predicting deleterious splice donor mutations. Hum Mutat. 1999;13(3):221-31.

-Koeberl DD, Bottema CD, Ketterling RP, Bridge PJ, Lillicrap DP, Sommer SS. Mutations causing hemophilia B: direct estimate of the underlying rates of spontaneous germ-line transitions, transversions, and deletions in a human gene. Am J Hum Genet. 1990 Aug;47(2):202-17.

-Kumar A, Mazzanti M, Mistrik M, Kosar M, Beznoussenko GV, Mironov AA, Garrè M, Parazzoli D, Shivashankar GV, Scita G, Bartek J, Foiani M. ATR mediates a checkpoint at the nuclear envelope in response to mechanical stress. Cell. 2014 Jul 31;158(3):633-46. doi: 10.1016/j.cell.2014.05.046.

-Lavin MF. Ataxia-telangiectasia: from a rare disorder to a paradigm for cell signalling and cancer. Nat Rev Mol Cell Biol. 2008 Oct;9(10):759-69. doi:10.1038/nrm2514. Review. Erratum in: Nat Rev Mol Cell Biol. 2008 Dec;9(12). doi:10.1038/nrm2514.

-Lindsey-Boltz LA, Kemp MG, Capp C, Sancar A. RHINO forms a stoichiometric complex with the 9-1-1 checkpoint clamp and mediates ATR-Chk1 signaling. Cell Cycle. 2015;14(1):99-108. doi: 10.4161/15384101.2014.967076.

-Liu S, Shiotani B, Lahiri M, Maréchal A, Tse A, Leung CC, Glover JN, Yang XH, Zou L. ATR autophosphorylation as a molecular switch for checkpoint activation. Mol Cell. 2011 Jul 22;43(2):192-202. doi: 10.1016/j.molcel.2011.06.019.

-Lopez JA, Chung DW, Fujikawa K, Hagen FS, Davie EW, Roth GJ. The alpha and beta chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence. Proc Natl Acad Sci U S A. 1988 Apr;85(7):2135-9.

-Lovejoy CA, Cortez D. Common mechanisms of PIKK regulation. DNA Repair (Amst). 2009 Sep 2;8(9):1004-8. doi: 10.1016/j.dnarep.2009.04.006. Epub 2009 May 21. Review.

-MACFARLANE RG. AN ENZYME CASCADE IN THE BLOOD CLOTTING MECHANISM, AND ITS FUNCTION AS A BIOCHEMICAL AMPLIFIER. Nature. 1964 May 2;202:498-9.

-Mackay DR, Ullman KS. ATR and a Chk1-Aurora B pathway coordinate postmitotic genome surveillance with cytokinetic abscission. Mol Biol Cell. 2015 Jun 15;26(12):2217-26. doi: 10.1091/mbc.E14-11-1563. Epub 2015 Apr 22.

-Mailand N, Podtelejnikov AV, Groth A, Mann M, Bartek J, Lukas J. Regulation of G(2)/M events by Cdc25A through phosphorylation-dependent modulation of its stability. EMBO J. 2002 Nov 1;21(21):5911-20.

-Mann KG, Nesheim ME, Church WR, Haley P, Krishnaswamy S. Surface-dependent reactions of the vitamin K-dependent enzyme complexes. Blood. 1990 Jul 1;76(1):1-16. Review. PubMed PMID: 2194585.

-Mann KG. Biochemistry and physiology of blood coagulation. Thromb Haemost. 1999 Aug;82(2):165-74. Review.

-Maréchal A, Li JM, Ji XY, Wu CS, Yazinski SA, Nguyen HD, Liu S, Jiménez AE, Jin J, Zou L. PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry. Mol Cell. 2014 Jan 23;53(2):235-46. doi: 10.1016/j.molcel.2013.11.002. Epub 2013 Dec 12.

-Marumoto T, Zhang D, Saya H. Aurora-A - a guardian of poles. Nat Rev Cancer. 2005 Jan;5(1):42-50. Review.

-Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER 3rd, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, Shiloh Y, Gygi SP, Elledge SJ. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science. 2007 May 25;316(5828):1160-6.

-Mordes DA, Cortez D. Activation of ATR and related PIKKs. Cell Cycle. 2008 Sep 15;7(18):2809-12. Epub 2008 Sep 30. PubMed PMID: 18769153; PubMed Central PMCID:PMC2672405.

-Murga M, Bunting S, Montaña MF, Soria R, Mulero F, Cañamero M, Lee Y, McKinnon PJ, Nussenzweig A, Fernandez-Capetillo O. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. Nat Genet. 2009 Aug;41(8):891-8. doi: 10.1038/ng.420. Epub 2009 Jul 20.

-Murphy SL, High KA. Gene therapy for haemophilia. Br J Haematol. 2008 Mar;140(5):479-87. doi: 10.1111/j.1365-2141.2007.06942.x. Review.

-Nam EA, Cortez D. ATR signalling: more than meeting at the fork. Biochem J. 2011 Jun 15;436(3):527-36. doi: 10.1042/BJ20102162. Review.

-Nathwani AC, Reiss UM, Tuddenham EG, Rosales C, Chowdary P, McIntosh J, Della Peruta M, Lheriteau E, Patel N, Raj D, Riddell A, Pie J, Rangarajan S, Bevan D, Recht M, Shen YM, Halka KG, Basner-Tschakarjan E, Mingozzi F, High KA, Allay J, Kay MA, Ng CY, Zhou J, Cancio M, Morton CL, Gray JT, Srivastava D, Nienhuis AW, Davidoff AM. Long-term safety and efficacy of factor IX gene therapy in hemophilia B. N Engl J Med. 2014 Nov 20;371(21):1994-2004. doi:10.1056/NEJMoa1407309.

-Nathwani AC, Tuddenham EG, Rangarajan S, Rosales C, McIntosh J, Linch DC, Chowdary P, Riddell A, Pie AJ, Harrington C, O'Beirne J, Smith K, Pasi J, Glader B, Rustagi P, Ng CY, Kay MA, Zhou J, Spence Y, Morton CL, Allay J, Coleman J, Sleep S, Cunningham JM, Srivastava D, Basner-Tschakarjan E, Mingozzi F, High KA, Gray JT, Reiss UM, Nienhuis AW, Davidoff AM. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. N Engl J Med. 2011 Dec 22;365(25):2357-65. doi: 10.1056/NEJMoa1108046. Epub 2011 Dec 10.

-Nghiem P, Park PK, Kim Y, Vaziri C, Schreiber SL. ATR inhibition selectively sensitizes G1 checkpoint-deficient cells to lethal premature chromatin condensation. Proc Natl Acad Sci U S A. 2001 Jul 31;98(16):9092-7.

-O'Driscoll M, Ruiz-Perez VL, Woods CG, Jeggo PA, Goodship JA. A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet. 2003 Apr;33(4):497-501. Epub 2003 Mar 17.

-Ogi T, Walker S, Stiff T, Hobson E, Limsirichaikul S, Carpenter G, Prescott K, Suri M, Byrd PJ, Matsuse M, Mitsutake N, Nakazawa Y, Vasudevan P, Barrow M, Stewart GS, Taylor AM, O'Driscoll M, Jeggo PA. Identification of the first ATRIP-deficient patient and novel mutations in ATR define a clinical spectrum for ATR-ATRIP Seckel Syndrome. PLoS Genet. 2012;8(11):e1002945. doi:10.1371/journal.pgen.1002945. Epub 2012 Nov 8.

-Patthy L. Evolution of the proteases of blood coagulation and fibrinolysis by assembly from modules. Cell. 1985 Jul;41(3):657-63. Review.

-Paull TT. Mechanisms of ATM Activation. Annu Rev Biochem. 2015;84:711-38. doi:10.1146/annurev-biochem-060614-034335. Epub 2015 Jan 12. Review.

-Paulsen RD, Cimprich KA. The ATR pathway: fine-tuning the fork. DNA Repair (Amst). 2007 Jul 1;6(7):953-66. Epub 2007 May 24. Review.

-Paulsen RD, Soni DV, Wollman R, Hahn AT, Yee MC, Guan A, Hesley JA, Miller SC, Cromwell EF, Solow-Cordero DE, Meyer T, Cimprich KA. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. Mol Cell. 2009 Jul 31;35(2):228-39. doi: 10.1016/j.molcel.2009.06.021.

-Peng CY, Graves PR, Thoma RS, Wu Z, Shaw AS, Piwnica-Worms H. Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. Science. 1997 Sep 5;277(5331):1501-5.

-Perry J, Kleckner N. The ATRs, ATMs, and TORs are giant HEAT repeat proteins. Cell. 2003 Jan 24;112(2):151-5.

-Petermann E, Maya-Mendoza A, Zachos G, Gillespie DA, Jackson DA, Caldecott KW. Chk1 requirement for high global rates of replication fork progression during normal vertebrate S phase. Mol Cell Biol. 2006 Apr;26(8):3319-26.

-Petermann E, Woodcock M, Helleday T. Chk1 promotes replication fork progression by controlling replication initiation. Proc Natl Acad Sci U S A. 2010 Sep 14;107(37):16090-5. doi: 10.1073/pnas.1005031107. Epub 2010 Aug 30.

-Petrus I, Chuah M, VandenDriessche T. Gene therapy strategies for hemophilia: benefits versus risks. J Gene Med. 2010 Oct;12(10):797-809. doi:10.1002/jgm.1500. Review.

-Qiu X, Lu D, Zhou J, Wang J, Yang J, Meng P, Hsueh JL. Implantation of autologous skin fibroblast genetically modified to secrete clotting factor IX partially corrects the hemorrhagic tendencies in two hemophilia B patients. Chin Med J (Engl). 1996 Nov;109(11):832-9.

-Qvist P, Huertas P, Jimeno S, Nyegaard M, Hassan MJ, Jackson SP, Børglum AD. CtIP Mutations Cause Seckel and Jawad Syndromes. PLoS Genet. 2011 Oct;7(10):e1002310. doi: 10.1371/journal.pgen.1002310. Epub 2011 Oct 6.

-Sanchez Y, Wong C, Thoma RS, Richman R, Wu Z, Piwnica-Worms H, Elledge SJ. Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. Science. 1997 Sep 5;277(5331):1497-501.

-Segurado M, Diffley JF. Separate roles for the DNA damage checkpoint protein kinases in stabilizing DNA replication forks. Genes Dev. 2008 Jul 1;22(13):1816-27. doi: 10.1101/gad.477208.

-Shechter D, Costanzo V, Gautier J. ATR and ATM regulate the timing of DNA replication origin firing. Nat Cell Biol. 2004 Jul;6(7):648-55. Epub 2004 Jun 27.

-Sun Y, Jiang X, Chen S, Fernandes N, Price BD. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. Proc Natl Acad Sci U S A. 2005 Sep 13;102(37):13182-7. Epub 2005 Sep 2.

-Tavassoli K, Eigel A, Dworniczak B, Valtseva E, Horst J. Identification of four novel mutations in the factor VIII gene: three missense mutations (E1875G, G2088S, I2185T) and a 2-bp deletion (1780delTC). Hum Mutat. 1998;Suppl 1:S260-2.

-Toledo LI, Murga M, Gutierrez-Martinez P, Soria R, Fernandez-Capetillo O. ATR signaling can drive cells into senescence in the absence of DNA breaks. Genes Dev. 2008 Feb 1;22(3):297-302. doi: 10.1101/gad.452308.

-Toledo LI, Altmeyer M, Rask MB, Lukas C, Larsen DH, Povlsen LK, Bekker-Jensen S, Mailand N, Bartek J, Lukas J. ATR prohibits replication catastrophe by preventing global exhaustion of RPA. Cell. 2013 Nov 21;155(5):1088-103. doi:10.1016/j.cell.2013.10.043. Erratum in: Cell. 2014 Jan 16;156(1-2):374.

-Weiss RS, Enoch T, Leder P. Inactivation of mouse Hus1 results in genomic instability and impaired responses to genotoxic stress. Genes Dev. 2000 Aug 1;14(15):1886-98.

-Yamada M, Watanabe K, Mistrik M, Vesela E, Protivankova I, Mailand N, Lee M, Masai H, Lukas J, Bartek J. ATR-Chk1-APC/CCdh1-dependent stabilization of Cdc7-ASK (Dbf4) kinase is required for DNA lesion bypass under replication stress. Genes Dev. 2013 Nov 15;27(22):2459-72. doi: 10.1101/gad.224568.113.

-Zdziarska J, Undas A, Basa J, Iwaniec T, Skotnicki AB, de Moerloose P, Neerman-Arbez M. Severe bleeding and miscarriages in a hypofibrinogenemic woman heterozygous for the gamma Ala82Gly mutation. Blood Coagul Fibrinolysis. 2009 Jul;20(5):374-6. doi: 10.1097/MBC.0b013e328329f27a.

-Zhao H, Watkins JL, Piwnica-Worms H. Disruption of the checkpoint kinase 1/cell division cycle 25A pathway abrogates ionizing radiation-induced S and G2 checkpoints. Proc Natl Acad Sci U S A. 2002 Nov 12;99(23):14795-800. Epub 2002 Oct 24.

-Zou L, Elledge SJ. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science. 2003 Jun 6;300(5625):1542-8.

III

Aim of the thesis

In the last decades, enormous efforts have been pushed toward the development of therapeutic approaches for human genetic diseases, and the research in this area has obtained many achievements and successes. Recently, genome editing technologies based on programmable nucleases such as zinc finger nucleases, transcription activator–like effector nucleases and the clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 are opening up the possibility of achieving therapeutic genome editing in disease cells and tissues, resulting in the removal or correction of deleterious mutations or the insertion of protective mutations.

On the other hand, gene replacement, in which a normal copy of the defective gene is efficiently delivered and expressed into target cells, represents one of the most advanced strategies that produced encouraging results in patients with single gene recessive disorders (i.e. like Hemophilia, Leber Congenital Amaurosis, adenosine-deaminase defect-ADA-SCID, lipoprotein lipase deficiency, Spinal Muscular Atrophy (SMA) etc.).

An emerging area of research is represented by the correction of the gene expression of the mutated gene by modulation of the messenger RNA (mRNA) processing; notably, RNA targeting would permit restoration of gene expression only in physiological tissue in which the gene is expressed. Moreover, it has the potential to circumvent the limitation related to the large size of certain human disease genes, which complicate their packaging in viral vectors, and could be also effective to address dominant-negative disease forms.

This work proposes the molecular characterization of mutations occurring in coagulation *F9* and *ATR*, associated respectively to Hemophilia B (HB) and Seckel Syndrome, and the exploitation of RNA-based approaches to correct them.

In HB cellular models, we have demonstrated that a unique spliceosomal U1snRNA targeting an intronic region downstream of a defective exon (Exon-specific U1snRNA, ExSpeU1) is able to rescue aberrant splicing caused by mutations located in the donor (5'ss) or acceptor (3'ss) splice sites. Here, we explored in mice the ExSpeU1-mediated rescue of two model HB-causing mutations at the 5'ss (c.519A>G) or 3'ss (c.392-8T>G) of *F9* exon 5 and assessed the correction at RNA level and the pro-coagulant activity of the rescued protein.

Pathologic splicing mechanisms might also involve exonic regulatory elements, which influence the splicing output and the design of correction strategies. Here we characterized and explored splicing rescue in a model in which numerous Hemophilia B (HB)-causing mutations, either missense or at the 5'ss of coagulation *F9* exon 2, promoted aberrant splicing by inducing the usage of a strong exonic cryptic 5'ss.

The development of correction strategies can be potentially translated to other human genetic disorder sharing similar aberrant splicing mechanisms. A third main aim of the project has been the characterized of the c.A2101G mutation in the ataxia-telengectasia and RAD3-related protein (ATR), a key player in DNA-damage response, to provide an RNA-based approach for Seckel Sindrome 1 (SS-1), a rare autosomal recessive disorder with dramatic unmet medical needs.

IV

An Exon-Specific U1snRNA Induces A Robust Factor IX Activity In Mice Expressing Multiple Human FIX Splicing Mutants

Dario Balestra[1], Daniela Scalet[1], Franco Pagani[2], Rosella Mari[3], Francesco Bernardi[1], Mirko Pinotti[1]

[1]Department of Life Science and Biotechnology, University of Ferrara and LTTA, Ferrara, Italy, [2] Internation Centre for Genetic Engineering and Biotechnology, Trieste, Italy and [3] Haemostasis & Thrombosis Center, University of Ferrara, Ferrara, Italy

Submitted to

Molecular Therapy - Nucleic Acids, February 2016

A crucial event in the earliest splicing step involves the recognition of the donor splice site (5′ss) by the U1 small nuclear ribonucleoprotein (U1snRNP), which is driven by base-pair complementarity with the 5′ tail of its RNA component (U1snRNA) (Roca et al. 2013). Over years, variants of the U1snRNA have been exploited to manipulate pre-mRNA processing for therapeutic purposes. Modified U1snRNAs have been used to carry hammered rybozymes (Michienzi et. al 1996) or, more frequently, as antisense molecules to induce skipping of pseudo-exons or, particularly, of defective exons, the latter hardly applicable to most human disease genes (Gorman et al. 2000; Cazzella et al. 2012).

Differently, our and other groups extensively exploited the physiological role of the U1snRNA to promote exon inclusion in the presence of exon-skipping mutations (Baralle et al. 2003, Susani et al. 2004; Pinotti et al. 2008; Pinotti et al. 2009; Tanner et al. 2009; Hartmann et al. 2010; Glaus et al. 2011; Sanchez-Alcudia et al. 2011; Schmid et al. 2013; Matos et al. 2014; Balestra et al. 2014; Balestra et al. 2015), a relevant cause of severe forms of human genetic disease (Faustino and Cooper 2003; Buratti et al. 2007; Baralle at al. 2009; Pinotti et al. 2011). The first generation U1snRNA had a modified 5′ tail with increased complementarity to defective 5′ss, and were shown to rescue exon inclusion in several cellular (Baralle et al. 2003, Susani et al. 2004; Pinotti et al. 2008; Pinotti et al. 2009; Tanner et al. 2009; Hartmann et al. 2010; Glaus et al. 2011; Sanchez-Alcudia et al. 2011; Schmid et al. 2013; Matos et al. 2014; Balestra et al. 2015) and also *in vivo* (Balestra et al. 2014) disease models. However, these U1snRNAs are often mutation-specific and have the intrinsic risk of off-target effects by recognizing the partially conserved donor splice site (McManus and Graveley 2011) in other splicing units. For this reason, we recently developed a second-generation U1snRNAs, named ExSpeU1s, that are designed to recognize intronic, often poorly conserved and thus gene specific, regions downstream of the affected exon, and shown that they efficiently rescue exon-skipping in different disease models (Fernandez et al. 2012; Dal Mas et al. 2015a; Dal Mas et al. 2015b;

Nizzardo et al. 2015; van der Woerd et al 2015). Most importantly, in cellular cultures, a unique ExSpeU1 was able to rescue normal splicing from different type of exon-skipping mutations either at the 5′ss or acceptor splice site (3′ss), or within the exon (Fernandez et al. 2012). Although these gene-specificity and activity ExSpeU1 features would significantly extend the applicability of a single therapeutic molecule to panels of mutations and thus cohorts of patients, a key issue when addressing the numerous diseases with heterogeneous mutational patterns (www.hgmd.cf.ac.uk), this effect of ExSpeU1s *in vivo* has not been investigated yet.

In this thesis, we explored in mice the efficacy of ExSpeU1 toward exon-skipping mutations leading to Hemophilia B (HB) that is a rare X-linked hemorrhagic disorder (1/35000 males) associated with reduced levels of factor IX (FIX), a key coagulation protein of liver origin (Bolton-Maggs and Pasi 2003).

HB represents a paradigmatic example of human disease with a heterogeneous mutational pattern comprising several splicing mutations, often associated with severe forms (Giannelli et al. 1998), characterized by strong unmet medical needs. The observation that even small increase of FIX levels (>2%), quantitatively measurable by functional and protein assays in plasma, would significantly ameliorate the clinical phenotype makes HB an ideal model to investigate innovative therapeutic approaches such as ExSpeU1 (Pollak and High 2003).

Through the expression in mice of two natural FIX splicing-defective variants at 5′ss or 3′ss associated with severe HB, we show that the selected ExSpeU1 efficiently rescues human FIX (hFIX) splicing in liver and results in the concurrent robust increase of hFIX protein and coagulant activity in plasma.

RESULTS

The *in vivo* ability of ExSpeU1 to restore splicing impaired by exon-skipping mutations was explored in the challenging context of the FIX exon 5 that, as many other splicing units, is characterized by a very weak donor splice site and is intrinsically poorly defined (Fig. IV.1A), and indeed prone to aberrant splicing. As a matter of fact, exon 5 is partially skipped even in physiological conditions, as demonstrated by the splicing pattern observed in human liver (Fig. IV.1B), and completely excluded from the mature mRNA in the presence of mutations at 5′ss or 3′ss, as indicated by cellular models.[23] Among these severe HB causing mutations, the c.519A>G and c.392-8T>G changes at the 5′ss or 3′ss respectively were selected as models to assess in mice the selected ExSpeU1 (U1fix9), which targets the intronic region starting at position +9 downstream of the 5′ss (Fig. IV.1A).

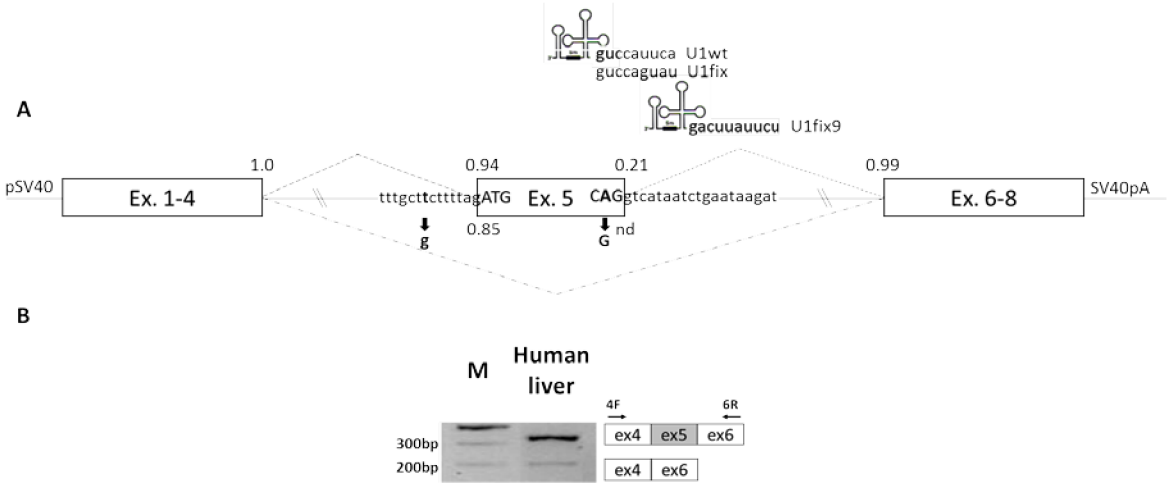


Figure IV.1 Features of the human FIX exon 5 context and FIX mRNA splicing patterns.
A) Schematic representation of the full-length splicing-competent cassette and of the modified U1snRNA (U1fix and U1fix9) with the 5′ tail sequence placed above the corresponding hFIX target. U1fix9 base pairs with the intronic sequence located 9 nt downstream the exon 5 5′ss. The U1wt, used as control, is also shown for the sake of completeness. Exonic and intronic hFIX sequences are represented by boxes and lines, respectively. Upper and lower dotted lines indicate the normal and aberrant splicing patterns of exon 5. The sequences of the 5′ss and 3′ss splicing junctions, with exonic and intronic regions in upper and lower cases, are also reported together with mutations under investigation (arrows). Scores of the 5′ss and 3′ss, calculated using the Splice Site Prediction by Neural Network tool (http://www.fruitfly.org/seq_tools/splice.html), are reported in normal and mutated conditions. "n.d", not determined by the computational program as 5' ss.
pSV40, SV40 promoter; SV40pA, SV40 poly A signal.
B) Evaluation of hFIX alternative splicing patterns in a liver sample from a healthy subject. The schematic representation of the normal and aberrant transcripts, and of primers used for the RT-PCR (arrows), is reported in the right panel. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker.

-The U1fix9 significantly rescues exon 5 inclusion of splicing-defective hFIX variants-

To deliver in mouse liver the plasmids harboring the hFIX splicing-competent (pFIXwt, pFIX-2G$^{5'ss}$, pFIX-8G$^{3'ss}$) or the U1snRNA (pU1fix9, pU1fix, pU1wt) cassettes (Fig. IV.1A) we exploited the hydrodynamic tail-vein injection of a high volume of concentrated DNA. This physical method takes advantage of the high-pressure gradient created in liver capillaries, which enhances endothelial and parenchymal cell permeability to DNA. The injection (1.5 µg/g of mouse body weight) of the pFIXwt, pFIX-2G$^{5'ss}$ and pFIX-8G$^{3'ss}$ variants resulted in mouse liver in appreciable hFIX mRNA expression characterized, for the three expression cassettes, by hFIX transcripts lacking exon 5 (Fig. IV.2, lanes 1, 3, 7).

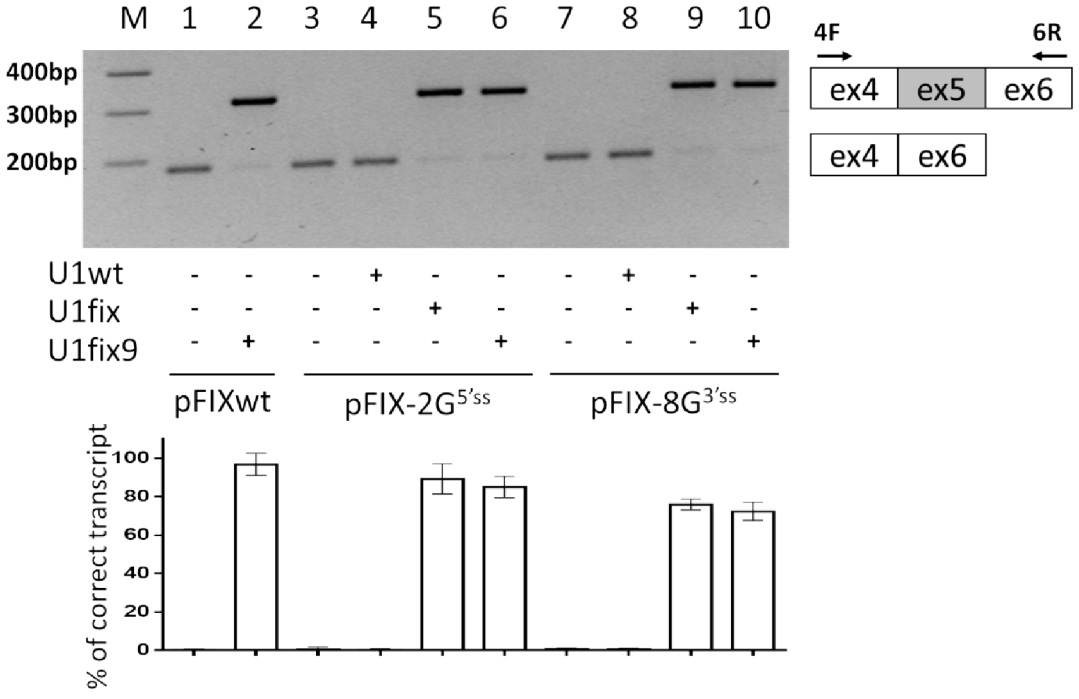


Figure IV.2 Evaluation of the U1fix9-mediated rescue of hFIX mRNA splicing in mice.
Representative example of hFIX alternative splicing patterns in BALB/c mice injected with the pFIX variants alone (-) or in combination (+) with a molar excess (1.5X) of the pU1wt, pU1fix or pU1fix9. The schematic representation of the normal and aberrant transcripts, and of primers used for the RT-PCR (arrows), is reported in the right panel. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker.
The histograms in the lower panel report the relative percentage of correct transcripts measured in liver from mice injected at the different conditions (n=4 for group 1-2, 4-5, 7-10; n=5 and n=7 for groups 3 and 6, respectively), and results are expressed as mean ± standard deviation.

The hFIX mRNA was not detected in other tissues such as heart, kidney and spleen (Fig. IV.3).

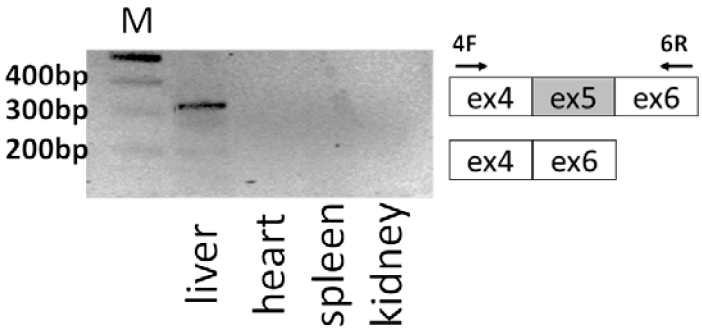


Figure IV.3 Representative example of hFIX splicing patterns in various organs of mice co-injected with pFIX-2G$^{5'ss}$ and pU1fix9.

Co-delivery of the U1fix9 with pFIX variants restored correct splicing of FIXwt (from negligible to 97±6%, mean ± standard deviation) and substantially rescued HB-causing variants with impaired 5′ss or 3′ss (from negligible to 85±5% and 72±4% for pFIX-2G$^{5'ss}$ and pFIX-8G$^{3'ss}$, respectively). Comparable results were obtained with the co-injection of the first generation pU1fix, used as control and having perfect complementarity to the wild type FIX exon 5 5′ss, which efficiently rescued exon inclusion (from negligible to 89±8% and 76±3% for pFIX-2G$^{5'ss}$ and pFIX-8G$^{3'ss}$, respectively)(Fig. IV.2). Conversely, co-injection of the pU1wt, mimicking the endogenous human or mouse U1snRNA, was ineffective.

By exploiting semi-quantitative RT-PCR with U1fix9-specific primers the expression of the U1fix9 in mouse liver upon injection of the pU1fix9 was clearly detectable and accounted for approximately half of those of the endogenous U1snRNA (Fig. IV.4). On the other hand, the endogenous U1snRNA expression showed an apparent increase (140% over untreated mice) in mice injected with the pU1wt.

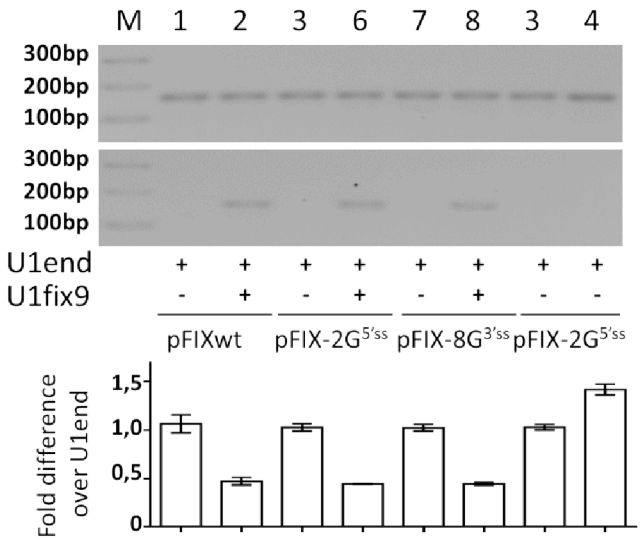


Figure IV.4 Evaluation of the expression of the endogenous (U1end, upper panel) and engineered (U1fix9, lower panel) U1snRNA by semi-quantitative RT- PCRs in mouse livers. The histograms report the relative percentage of U1snRNA expression levels as compared with those of the U1end. Results are expressed as mean ± standard deviation. M, 100 bp molecular weight marker.

-The U1fix9 significantly increases circulating hFIX levels that results in remarkable improvement of FIX coagulant activity-

The hFIX splicing-competent cassettes were designed to incorporate the complete hFIX coding sequence (Fig. IV.1A), which permitted us to evaluate the hFIX protein secreted in mouse plasma. Expression of the pFIXwt, pFIX-2G$^{5'ss}$ and pFIX-8G$^{3'ss}$ produced low hFIX antigen levels (0.33±0.2 µg/ml, 0.12±0.08 µg/ml and 0.18±0.11 µg/ml, respectively)(Fig. IV.5).

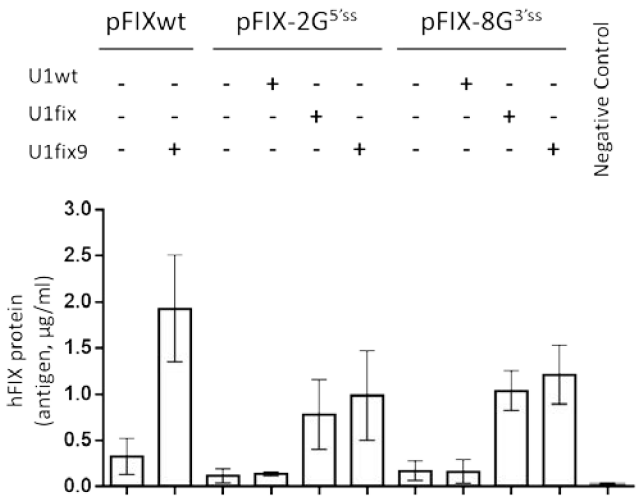


Figure IV.5 Evaluation of U1fix9-mediated rescue of hFIX protein and coagulant properties in mice. Plasma hFIX antigen levels (µg/ml) in mice injected with the pFIX variants alone (-) or in combination (+) with a molar excess (1.5X) of pU1wt, pU1fix or pU1fix9. Negative controls are represented by mice injected with saline solution alone. To evaluate hFIX antigen in mouse plasma, samples were diluted 1:10 or 1:20 and evaluated in duplicate. A standard curve was created by adding known amounts of hFIX to mouse plasma, and the sensitivity threshold was 0.40 ng/mL of hFIX. Histograms reports the results obtained in the different mouse groups (n=4 for group 1-2, 4-5, 7-10; n=5 and n=7 for groups 3 and 6, respectively) and are expressed as mean ± standard deviation.

Western Blotting with two different anti-hFIX antibodies was then exploited to investigate the circulating hFIX isoforms. In our experimental set-up, the anti-hFIX antibody 1 recognized both human and mouse FIX, which however are distinguishable by size (Fig. IV.6a).

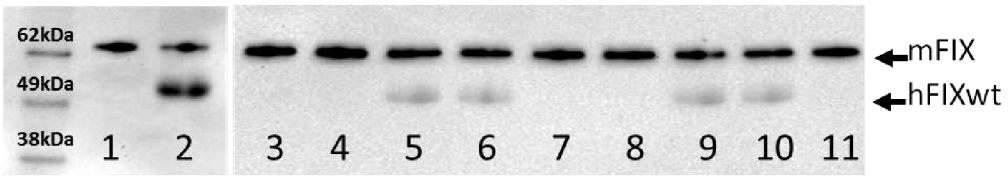


Figure IV.6a Western Blotting analysis of hFIX isoforms in plasma from mice treated as in panel A. mFIX, mouse FIX; hFIXwt, human hFIX.

Investigations in plasma from the pFIXwt, pFIX-2G5′ss and pFIX-8G3′ss injected mice (lanes 1, 3, 7) did not reveal appreciable amount of the normal hFIX form or of that lacking the crucial EGF2 domain and arising from the in-frame exon-skipped FIX

transcript. On the other hand, trace amounts of the deleted hFIX isoform were revealed by Western Blotting with the anti-hFIX antibody 2 (Fig. IV.6b).

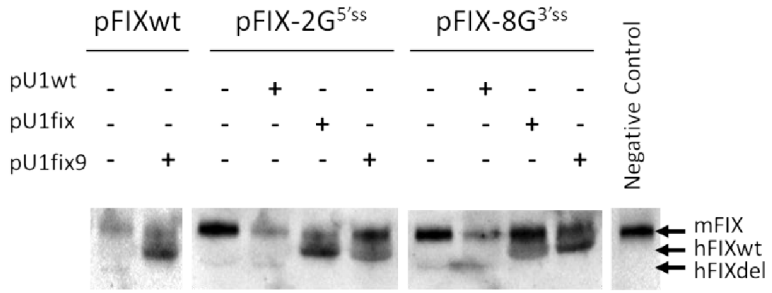


Figure IV.6b Western Blotting analysis of hFIX isoforms in plasma from mice treated as in panel A. mFIX, mouse FIX; hFIXwt, full-length hFIX; hFIXdel, hFIX variant lacking EGF2.

When the FIX coagulant activity was assessed (Fig. IV.7), the deleted variant did not appreciable shorten the FIX-dependent coagulation times (pFIXwt 96±3 s; pFIX-2G$^{5'ss}$, 99±2 s; pFIX-8G$^{3'ss}$, 104±3 s), as compared with control mice (101±2 s).

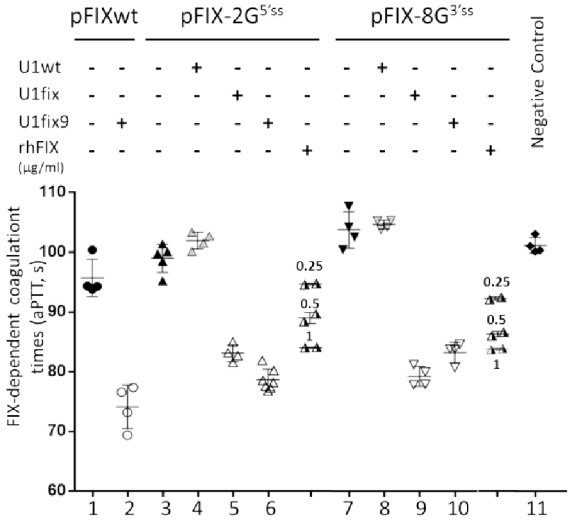


Figure IV.7 FIX-dependent aPTT values (s) in plasma from mice treated as in panel A or in plasma from mice (two independent) expressing the mutated hFIX variants supplemented with known amounts of purified rhFIX (0.25, 0.5 or 1 µg/ml). All samples have been tested in duplicate and the mean reported as single symbol.

Co-injection of the pFIXwt or variants with pU1fix9, but not with pU1wt, resulted in a significant increase of hFIX antigen levels (pFIXwt 1.9±0.58 µg/ml, $p<0.0006$; pFIX-2G$^{5'ss}$ 0.99±0.48 µg/ml, $p<0.001$; pFIX-8G$^{3'ss}$ 1.2±0.32 µg/ml, $p<0.02$) (Fig. IV.5) with the concurrent appearance of the full-length hFIX isoform (Fig. 3B). This led to a remarkable

and statistically significant shortening of FIX-dependent coagulation times ($P<0.0001$), with values in mice expressing the hFIX variants (pFIX-2G$^{5'ss}$, 79±2 s; pFIX-8G$^{3'ss}$, 83±2 s) resembling those for FIXwt (74±4 s) (Fig. IV.7). Similar results were obtained by co-injection of the pU1fix, with a strong increase of hFIX antigen levels (pFIX-2G5′ss, 0.78 ±0.38 µg/ml; pFIX-8G3′ss 1.0±0.22 µg/ml), the appearance of the full-length isoform and a significant ($P<0.0001$) shortening in the coagulation time (pFIX-2G$^{5'ss}$, 83±1 s; pFIX-8G$^{3'ss}$, 79±2 s). Conversely, co-expression of hFIX variants with the U1wt was ineffective. To provide a quantitative evaluation of the contribution of the rescued hFIX to coagulation times, we spiked plasma from mice expressing the FIX-2G$^{5'ss}$ or FIX-8G$^{3'ss}$ variants only with 1, 0.5 or 0.25 µg/ml of recombinant hFIX (Fig. IV7), with the lowest concentration shortening the FIX-dependent coagulation times of 5 seconds in set up experiments. Noticeably, in plasma from mice co-expressing the hFIX variants and the U1fix9 the coagulation times were similar or shorter to those obtained by supplementation with 1 µg/ml of recombinant human FIX (rhFIX).

DISCUSSION

In the present study, by using splicing-defective hFIX expression as model, we evaluated the *in vivo* efficacy of ExSpeU1s toward multiple mutations, a RNA-based approach that could represent a powerful alternative to replacement gene therapy, especially when the regulation and/or the size of the affected gene makes it hardly feasible. The ExSpeU1s, by acting on pre-mRNA splicing, have the invaluable advantage of maintaining the transcriptional gene regulation and of guaranteeing the rescue of gene expression in physiologic tissues only, thus potentially overcoming immunologic concerns. Moreover, the small size of the ExSpeU1 transgene (~600 bp) makes it packageable in virtually any viral vector so far successfully exploited for gene therapy purposes (Rogers

and Herzog 2015). Furthermore, the ExSpeU1 expression is driven by its strong ubiquitous endogenous promoter, and in the model of Spinal Muscular Atrophy a single ExSpeU1 copy can trigger an appreciable rescue (Dal Mas et al. 2015b), which could lead to the use of low doses of viral vector to achieve a therapeutic impact.

On the other hand, we are aware that ExSpeU1, like other personalized approaches based on the mutation type, must cope with the fact that many diseases, such as Hemophilia B, have highly heterogeneous mutational patterns (www.hgmd.cf.ac.uk). As a matter of fact, skipping of a given exon is a relatively frequent pathogenic mechanism that can be caused by several different mutations either at the 5′ss or 3′ss, or within the exon, which could weaken the applicability of the modified U1.

Here, to assess the ability of ExSpeU1 in restoring splicing in mice, we exploited the hydrodynamic injection of a single ExSpeU1 (U1fix9) toward two model exon-skipping mutations at the opposite 5′ss (c.519A>G) or 3′ss (c.392-8T>G) positions in the FIX exon 5 context to drive the transient expression of the splicing-defective hFIX cassette. Interestingly, even in the presence of the non-liver specific SV40 promoter, the hydrodynamic tail vein injection approach guaranteed an appreciable delivery of plasmid DNA with an appreciable expression of the hFIX transgene mainly in liver. It is worth nothing that optimization of species-specific nucleic acid amplification and immunologic assays enabled us to evaluate the hFIX in wild-type mice without the confounding effects of the endogenous murine FIX, as indicated by the undetectable hFIX transcripts and protein in untreated mice. Moreover, aPTT-based coagulation assays were optimized to assess the additive effect of hFIX to the endogenous murine one. Taken together these features provided us with an informative phenotype after treatment and thus a straightforward strategy to overcome the lack of mouse models for the HB mutations under investigation (Sabatino et al. 2012).

The use of the strong SV40 promoter instead of the physiological one might have further promoted in mice exon-skipping of the minigene, which contains a poorly defined exon. As a matter of fact, mRNA splicing and RNA polymerase II elongation rate/processivity are intricately intertwined, with an highly elongating pol II (SV40 promoter) favoring exon exclusion, and *vice versa* (Dujardin et al. 2013). Notwithstanding the unfavorable exon 5 context, the co-expression of the ExSpeU1 U1fix9, as clearly demonstrated by U1fix9-specific PCR, was able to remarkably recover exon inclusion in the FIXwt and, most importantly, in the presence of the defective 5′ss or 3′ss. Moreover, the strong rescue effect was comparable to that of the first generation U1fix directly targeting the 5'ss, which however holds the high risk of off-targets. Conversely, the U1wt was ineffective, thus highlighting the crucial role of the engineered U1snRNA 5′s tail in defining both target recognition and splicing correction.

A mandatory issue to evaluate the therapeutic potential of an approach is to assess the impact on protein levels and particularly on the activity in the affected pathway. Consistently, in mice treated with the pU1fix9, we measured a remarkable increase of circulating hFIX protein, and the appearance of the full-length protein isoform. Most importantly, this resulted in a robust shortening in the coagulation times (15-20 s) produced by the increase of hFIX protein endorsed of a normal specific activity. From this observation, it is tempting to speculate that the rescue extent would be beneficial if translated in the severe HB patients in whom the therapeutic threshold is rather low.

In addition, a computation search (Blast) in the human genome and transcripts database using the U1fix9 target sequence, which however does not consider RNA-RNA wobble base-pairing, identified only few alternative mRNA targets (CCDC64, MYNN, FSHR, FAM174A mRNAs), all of them within exons and not at exon-intron boundaries. Although the prediction has to be validated by future studies (i.e. RNA sequencing), this observation

points towards *F9* gene specific features of the ExSpeU1 U1fix9, which would decrease the risk of off-target effects and thus favor the translation into clinic.

Taken together these findings indicated that multiple HB splicing-defective variants causing exon-skipping can be efficiently rescued *in vivo* by a single ExSpeU1, able to promote appreciable synthesis of hFIX endorsed of a normal coagulant activity.

In conclusion, albeit our mouse models permit the assessment of the impact on coagulation phenotype and not on provoked bleeding, which will require the creation of splicing-specific HB mouse models, and despite the unfavorable FIX exon 5 definition, these data provide the first proof-of-concept about the *in vivo* properties of ExSpeU1s, which substantially extends their applicability.

MATERIALS AND METHODS

Expression vectors, delivery in mice and sampling

Expression vectors for the i) splicing-competent hFIX cDNA cassette, either wild-type (pFIXwt) or mutated (pFIX-2G$^{5'ss}$; pFIX-8G$^{3'ss}$), or ii) the U1snRNA, either wild-type (pU1wt) or exon-specific (pU1fix9) were previously reported.[23]

Eight-week-old BALB/C mice were hydrodynamically injected with plasmid DNA in 2.5 mL phosphate buffer saline-PBS through the tail vein. Blood samples were collected from the retro-orbital plexus into 3.8% sodium citrate 48 hours post-injection. Upon sacrifice, total RNA was isolated from random sections of mouse liver using TRIZOL reagent (LifeTechnologies, Carlsbad, CA, USA).

All procedures were approved and conducted under the guidelines established by the Italian Ministry of Health.

Evaluation of hFIX mRNA and U1 expression

Two micrograms of RNA from mouse tissues were retro-transcribed using SuperScript III Reverse Transcriptase (Life Technologies) with random primers and subsequently PCR amplified with hFIX-specific primers $^{5'}$ATTCCTATGAATGTTGGTGTCCCT$^{3'}$ (4F) and $^{5'}$GGGTGCTTTGAGTGATG TTATCCAA$^{3'}$ (6R). The PCR was run for 40 cycles of 95 °C for 30 s, 60 °C for 20 s and 72 °C for 30 s. Evaluation of U1snRNA expression was carried out using U1 specific primers $^{5'}$AGATCTGATACTTACCTG$^{3'}$ (U1wtF), $^{5'}$AGATCTCATTCTTATTCAG$^{3'}$ (U1f ix9F) and $^{5'}$GAACGCAGTCCCCCACTACCAC$^{3'}$ (U1R).

Assessment of hFIX antigen, protein isoforms and coagulant activity

Human FIX antigen levels and isoforms in plasma were evaluated by ELISA (Affinity Biologicals inc, Ontario, Canada) and Western blotting exploiting the polyclonal hFIX antibody 1 or 2 (GAFIX-AP or GAFIX-APHRP, respectively; Affinity Biologicals inc, Ontario, Canada).

FIX coagulant activity was assessed by coagulation assays based on the activated partial thromboplastin time (aPTT). Assays with serial dilution of mouse plasma spiked with recombinant hFIX were conducted to optimize the protocol and magnify the additive effect of hFIX to the endogenous murine one. In the final setting, mouse plasma was diluted 1:200 in Imidazole buffer and mixed 1:1 with hFIX-deficient plasma (George King, Bio-Medical Inc., Overland Park, KS, USA) before adding the activator SynthASil reagent (Instrumentation Laboratory, Bedford, USA) and calcium. Coagulation times were recorded by the ACLTOP 700 instrument (Instrumentation Laboratory, Bedford, USA).

Statistical analysis

Statistical differences among levels were evaluated by a Student′s t-test with a $P < 0.05$ considered significant.

REFERENCES

-Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M. An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice. J Thromb Haemost. 2014 Feb;12(2):177-185. doi: 10.1111/jth.12471.

-Balestra D, Barbon E, Scalet D, Cavallari N, Perrone D, Zanibellato S, Bernardi F, Pinotti M. Regulation of a strong F9 cryptic 5'ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides. Hum Mol Genet. 2015 Sep 1;24(17):4809-16. doi: 10.1093/hmg/ddv205. Epub 2015 Jun 10.

-Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, Ffrench-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 Mar;40(3):220-2.

-Baralle D, Lucassen A, Buratti E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. EMBO Rep. 2009 Aug;10(8):810-6. doi:10.1038/embor.2009.170. Review.

-Bolton-Maggs PH, Pasi KJ. Haemophilias A and B. Lancet. 2003 May 24;361(9371):1801-9. Review.

-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(13):4250-63. Epub 2007 Jun 18.

-Cazzella V, Martone J, Pinnarò C, Santini T, Twayana SS, Sthandier O, D'Amico A, Ricotti V, Bertini E, Muntoni F, Bozzoni I. Exon 45 skipping through U1-snRNA antisense molecules recovers the Dys-nNOS pathway and muscle differentiation in human DMD myoblasts. Mol Ther. 2012 Nov;20(11):2134-42. doi: 10.1038/mt.2012.178. Epub 2012 Sep 11.

-Dal Mas A, Fortugno P, Donadon I, Levati L, Castiglia D, Pagani F. Exon-Specific U1s Correct SPINK5 Exon 11 Skipping Caused by a Synonymous Substitution that Affects a Bifunctional Splicing Regulatory Element. Hum Mutat. 2015a May;36(5):504-12. doi: 10.1002/humu.22762. Epub 2015 Mar 19.

-Dal Mas A, Rogalska ME, Bussani E, Pagani F. Improvement of SMN2 pre-mRNA processing mediated by exon-specific U1 small nuclear RNA. Am J Hum Genet. 2015b Jan 8;96(1):93-103. doi: 10.1016/j.ajhg.2014.12.009. Epub 2014 Dec 31.

-Dujardin G, Lafaille C, Petrillo E, Buggiano V, Gómez Acuña LI, Fiszbein A, Godoy Herz MA, Nieto Moreno N, Muñoz MJ, Alló M, Schor IE, Kornblihtt AR. Transcriptional elongation and alternative splicing. Biochim Biophys Acta. 2013 Jan;1829(1):134-40. doi: 10.1016/j.bbagrm.2012.08.005. Epub 2012 Sep 7. Review.

-Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. Genes Dev. 2003 Feb 15;17(4):419-37. Review.

-Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. Hum Mol Genet. 2012 Jun 1;21(11):2389-98. doi:10.1093/hmg/dds045. Epub 2012 Feb 23.

-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 Jan 1;26(1):265-8.

-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 May;19(5):936-41. doi: 10.1038/mt.2011.7. Epub 2011 Feb 15.

-Gorman L, Mercatante DR, Kole R. Restoration of correct splicing of thalassemic beta-globin pre-mRNA by modified U1 snRNAs. J Biol Chem. 2000 Nov 17;275(46):35914-9.

-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 Oct 8;87(4):480-93. doi: 10.1016/j.ajhg.2010.08.016.

-Matos L, Canals I, Dridi L, Choi Y, Prata MJ, Jordan P, Desviat LR, Pérez B, Pshezhetsky AV, Grinberg D, Alves S, Vilageliu L. Therapeutic strategies based on modified U1 snRNAs and chaperones for Sanfilippo C splicing mutations. Orphanet J Rare Dis. 2014 Dec 10;9:180. doi: 10.1186/s13023-014-0180-y.

-Michienzi A, Prislei S, Bozzoni I. U1 small nuclear RNA chimeric ribozymes with substrate specificity for the Rev pre-mRNA of human immunodeficiency virus. Proc Natl Acad Sci U S A. 1996 Jul 9;93(14):7219-24.

-Nizzardo M, Simone C, Dametti S, Salani S, Ulzi G, Pagliarani S, Rizzo F, Frattini E, Pagani F, Bresolin N, Comi G, Corti S. Spinal muscular atrophy phenotype is ameliorated in human motor neurons by SMN increase via different novel RNA therapeutic approaches. Sci Rep. 2015 Jun 30;5:11746. doi:10.1038/srep11746.

-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 Mar 1;111(5):2681-4. Epub 2007 Dec 21.

-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 Jun 18;113(25):6461-4. doi: 10.1182/blood-2009-03-207613. Epub 2009 Apr 22.

-Pinotti M, Bernardi F, Dal Mas A, Pagani F. RNA-based therapeutic approaches for coagulation factor deficiencies. J Thromb Haemost. 2011 Nov;9(11):2143-52.doi: 10.1111/j.1538-7836.2011.04481.x. Review.

-Pollak, E and High, K.A. (2003) Genetic disorders of coagulation. In: Warrell, D., Cox, T., Firth, J. and Benz, E. (Eds) *Oxford Textbook of Medicine*, 4th ed, pp. 757–767, Oxford University Press, Oxford. Vol. 3

-Roca X, Krainer AR, Eperon IC. Pick one, but be quick: 5' splice sites and the problems of too many choices. Genes Dev. 2013 Jan 15;27(2):129-44. doi:10.1101/gad.209759.112. Review.

-Rogers GL, Herzog RW. Gene therapy for hemophilia. Front Biosci (Landmark Ed). 2015 Jan 1;20:556-603. Review.

-Sabatino DE, Nichols TC, Merricks E, Bellinger DA, Herzog RW, Monahan PE. Animal models of hemophilia. Prog Mol Biol Transl Sci. 2012;105:151-209. doi:10.1016/B978-0-12-394596-9.00006-8. Review.

-Sánchez-Alcudia R, Pérez B, Pérez-Cerdá C, Ugarte M, Desviat LR. Overexpression of adapted U1snRNA in patients' cells to correct a 5' splice site mutation in propionic acidemia. Mol Genet Metab. 2011 Feb;102(2):134-8. doi:10.1016/j.ymgme.2010.10.013. Epub 2010 Nov 5. PubMed PMID: 21094621.

-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 Jul;32(7):815-24. doi: 10.1002/humu.21509.

-Schmid F, Hiller T, Korner G, Glaus E, Berger W, Neidhardt J. A gene therapeutic approach to correct splice defects with modified U1 and U6 snRNPs. Hum Gene Ther. 2013 Jan;24(1):97-104. doi: 10.1089/hum.2012.110. Epub 2012 Nov 23.

-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 Sep;24(3):225-35.

-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 Feb;30(2):255-63. doi:10.1002/humu.20861. Erratum in: Hum Mutat. 2009 Feb;30(2):274.

-van der Woerd WL, Mulder J, Pagani F, Beuers U, Houwen RH, van de Graaf SF. Analysis of aberrant pre-messenger RNA splicing resulting from mutations in ATP8B1 and efficient in vitro rescue by adapted U1 small nuclear RNA. Hepatology. 2015 Apr;61(4):1382-91. doi: 10.1002/hep.27620. Epub 2015 Feb 23.

V

Regulation of a strong F9 cryptic 5'ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides

Dario Balestra[1,*], Elena Barbon[1,*], Daniela Scalet[1,*], Nicola Cavallari[1], Daniela Perrone[2], Silvia Zanibellato[1], Francesco Bernardi[1] and Mirko Pinotti[1]

[*]These authors contributed equally to this work

[1]Department of Life Science and Biotechnology, University of Ferrara and LTTA, Ferrara, Italy and [2]Department of Chemical and Pharmaceutical Sciences, University of Ferrara, Ferrara, Italy

Human Molecular Genetics 2015;24:4089-16

As mentioned earlier, the number of mutations affecting pre-mRNA splicing (Sterne-Weiler et al. 2011) is largely underestimated, particularly when considering nucleotide variations within exons (Teraoka et al. 1999; Ars et al. 2000), for which the impact on this finely orchestrated process (Chen and Manley 2009; Wahl et al. 2009; Nilsen and Graveley 2010) is hardly predictable. As a matter of fact, the splicing process output depends on the type of the splicing regulatory element affected by the mutation and on the network of interactions defining a given exon in the specific gene sequence context. Taking into account the frequency of each class of splicing regulatory elements (Zhang and Chasin 2004; Stadler et al. 2006; Kralovicova and Vorechovsky 2007), it is predicted that their interplay, particularly in the presence of inherited mutations, will produce an extremely variegate series of additive, negative or compensatory effects. The intriguing role of exonic splicing enhancer and silencer (ESE and ESS, respectively), sequences that substantially contribute to alternative splicing by regulating the choice of splice sites, is still largely unexplored, particularly when the proper exon inclusion is mandatory to encode a functional protein. The ESS, which are estimated to be significantly less frequent in "real" exons as compared to "pseudoexons" (Kralovicova and Vorechovsky 2007), appear to evolve with a strength that correlates with that of the 5'ss (Xiao et al. 2007). On the other hand, the influence of ESSs on cryptic 5'ss and their interplay with mutations could produce aberrant mRNA patterns (Teraoka et al. 1999; Cartegni and Krainer 2002; Pagani and Barelle 2004; Cavallari et al. 2012) through poorly explored combinations of mechanisms, the definition of which would in turn help implementing our knowledge of ESS physiological functions.

In this study we provide an intriguing model for positive and negative interactions among regulatory elements leading to severe Haemophilia B forms.

First, we demonstrate that several mutations at the exon 2 donor splice site (5'ss) of coagulation *F9* gene produce aberrant splicing by inducing the usage of a strong exonic cryptic 5'ss, which is regulated by adjacent exonic splicing regulatory elements, both ESE and ESS. Intriguingly, numerous HB-causing missense changes at the ESS (Green et al. 1989) increase the usage of the cryptic 5'ss. This provides the rationale for the use of antisense molecules (Bruno et al. 2004; Meyer et al. 2009; Nlend Nlend wt al. 2010; Spitali and Asrtsma-Rus 2012; Porensky et al. 2012; Havens et al. 2013) and variants of the small nuclear RNA U1 (U1snRNA) (Baralle et al. 2003; Susani et al. 2004; Pinotti et al. 2008; Pinotti et al. 2009; Tanner et al. 2009; Pinotti et al. 2011; Schmid et al. 2011; Glaus et al. 2011; Fernandez Alanis et al. 2012; Roca et al. 2013; Balestra et al. 2014) for a tailored

correction approach, exploiting the well-known advantages of RNA-based strategies (Pinotti et al. 2011). Antisense molecules, such as chemically modified oligonucleotides or engineered U7snRNA, or modified U1snRNA have been reported to counteract mutations and restore exon inclusion by masking either cryptic sites or improving exon definition in several models of human diseases (Baralle et al. 2003; Susani et al. 2004; Pinotti et al. 2008; Pinotti et al. 2009; Tanner et al. 2009; Pinotti et al. 2011; Fernandez Alanis et al. 2012; Schmid et al. 2011; Glaus et al. 2011; Balestra et al. 2014).

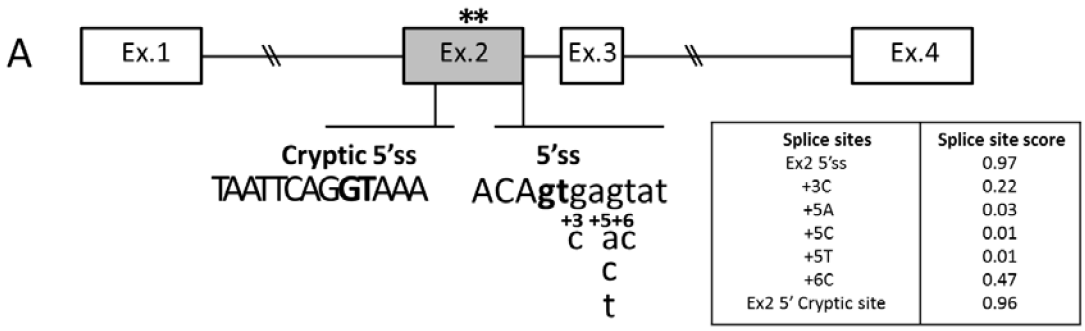
For the first time we demonstrate that splicing correction, in the presence of a strong cryptic 5'ss and of mutations at the authentic one, can be achieved only by the combined effect of modified U1snRNA and antisense oligonucleotides, which also contribute to better understand the mechanisms underlying proper exon definition through the involvement of positive and negative splicing regulatory elements.

RESULTS

We investigated a panel of mutations occurring at the positions +3 (c.252+3G>C), +5 (c.252+5G>A, c.252+5G>C, c.252+5G>T) and +6 (c.252+6T>C) of the 5'ss of exon 2 of *F9* gene (Fig. 1A) that have been reported in the Haemophilia B (HB) mutation database (http://www.factorix.org). All mutations cause severe HB forms characterized by factor IX (FIX) levels in plasma below 1%.

The *F9* exon 2 5'ss mutations induce usage of an exonic cryptic 5'ss regulated by *cis*-elements, including an ESS

To investigate the effect of *F9* Exon 2 5'ss mutations on pre-mRNA splicing we created a *F9* minigene including the genomic region spanning exon 1 through 4 (Fig. V.1).



Splice sites	Splice site score
Ex2 5'ss	0.97
+3C	0.22
+5A	0.03
+5C	0.01
+5T	0.01
+6C	0.47
Ex2 5' Cryptic site	0.96

Figure V.1
Schematic representation of the F9 genomic sequence cloned as minigene in the pCDNA3 vector. Exonic and intronic sequences are represented by boxes and lines, respectively. The sequences, with exonic and intronic nucleotides in upper and lower cases respectively, report i) the authentic 5'ss with the positions of the investigated changes detailed below, and ii) the cryptic 5'ss. The highly conserved dinucleotide GT of authentic or cryptic 5'ss is in bold. Asterisks represent the position of the two consecutive stop codons (TAA). The table reports the scores of the authentic and mutated 5'ss, and of the cryptic 5'ss.

Expression of the wild-type minigene in mammalian Baby Hamster Kidney (BHK) cells indicated that exon 2 was not completely included into the mature mRNA (Fig. V.2, lane 1). RT-PCR analysis and sequencing revealed a smaller amount (21±6%) of an alternative transcript that originates from the usage of a cryptic 5'ss located 104 bp upstream of the authentic 5'ss (Fig. V.2). This gives rise to an aberrant *F9* mRNA form (F9del) including only the 5' portion (60 bp) of exon 2 that accounts for a deleted and frame-shifted (F9del) mRNA harboring a premature nonsense triplet at position c.151. To rule out the possibility that the result depends on minigene artifacts, we performed RT-PCR on human liver mRNA and demonstrated a comparable amount (~20%) of the F9del transcript (Fig. V.2, lane 7). These findings are consistent with the computational analysis that predicts the cryptic 5'ss in exon 2 with a score (0.96) comparable to the authentic one (0.97) (Fig. V1). The mutations at the authentic *F9* exon 2 5'ss were then inserted into the validated F9wt minigene. Expression studies demonstrated that all mutations, reducing the 5'ss score lead to the virtually exclusive usage of the exonic cryptic 5'ss and synthesis of F9del transcripts (Fig. V.2, lanes 2-6).

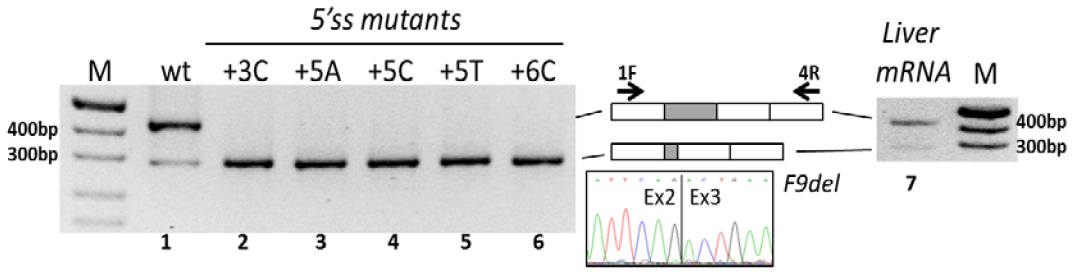


Figure V.2
Evaluation of F9 alternative splicing patterns in BHK cells transiently transfected with minigenes (left panel, lanes 1-6) or in human liver (right panel, lane 7). The schematic representation of the transcripts, and of primers used for the RT–PCR (arrows), is reported in the middle panel. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. The chromatogram reports the sequence of the shorter transcript (F9del), which demonstrates the usage of the cryptic 5' ss in exon 2.

However, the nonsense-mediated RNA decay (NMD) might lead to underestimate the levels of the F9del form, or prevent the evaluation of that arising from exon 2 skipping, both frame-shifted and introducing a premature nonsense triplet in exon 3. To rule out this possibility, we performed splicing assays in the presence of the NMD inhibitor cycloheximide (CHX). As shown in figure V.3, the splicing patterns in cells expressing the F9wt (lanes 1-2) or the +3C (lanes 3-4) constructs were virtually unaffected upon CHX treatment. This finding is consistent with the fact that the premature nonsense triplet in the F9del form occurs in a very short exon, a condition that would prevent NMD (Kervestin and Jacobson 2012). To further validate our observation, we also introduced two early nonsense triplets downstream of the cryptic site in exon 2 of the F9wt cassette (F9wt-xx), a situation known to promote NMD. As shown in figure V.3 (lane 5-6), we observed that the correctly spliced transcript, but not the F9del lacking the premature nonsense changes, undergoes major NMD. Altogether these data indicate that NMD has a negligible impact on splicing patterns of the *F9* constructs under investigation.

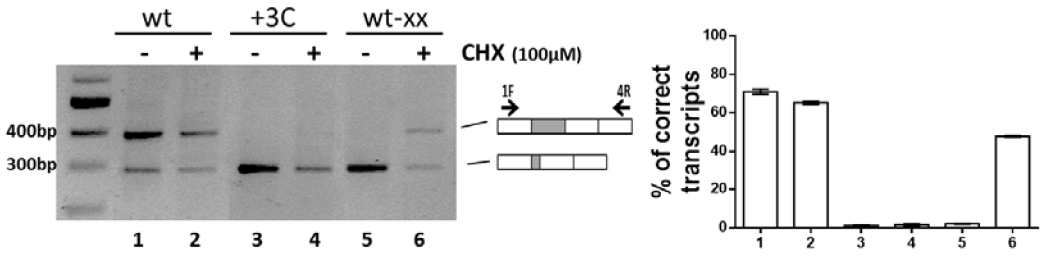


Figure V.3
Evaluation of F9 alternative splicing patterns in BHK cells transiently transfected with pF9wt (lanes 1-2), pF9+3C (lanes 3-4) and pF9wt-xx (lanes 5-6) alone (-) or upon treatment with 100µM cycloheximide (CHX). Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. The wt-xx represents the F9wt cassette harboring two early nonsense triplets downstream of the cryptic site in exon 2 inserted to investigate, as control, the occurrence of non-sense mediated decay. The schematic representation of the transcripts, and of primers used for the RT–PCR (arrows), is reported in the right panel. Histograms report the relative percentage of correct transcripts, which is expressed as mean ± standard deviation (SD).

The bioinformatics analysis of *F9* exon 2 predicted several Exonic Splicing Enhancers and also a Silencer element (AAAGAGGT) upstream of the cryptic 5'ss (Fig. V.5), and revealed that this region is highly conserved among mammals (Fig. V.4).

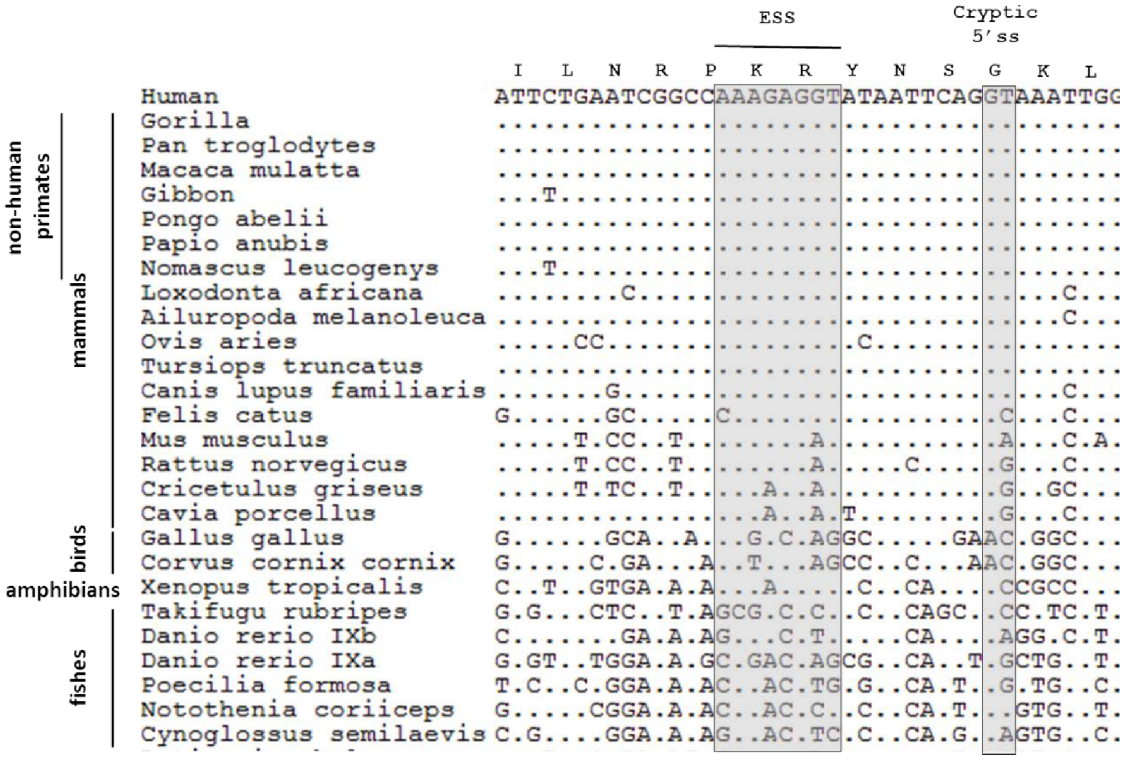


Figure V.4.
Multiple alignment of the FIX cDNAs from multiple species showing the sequence under investigation. The cryptic 5'ss and the ESS element are boxed in grey. Dots represents conserved nucleotides. The missense changes are reported above.

Intrigued by the hypothesis that a silencer favors the usage of the authentic 5'ss by down-regulating the cryptic one, we investigated its presence by exploiting natural models consisting of all missense mutations (c.135g/c, c.135g/t, c.137g/a, c.137g/c, c.138g/t, c.138g/c) (Fig. V.6) at the putative ESS sequence so far reported in moderate/severe HB patients (http://www.factorix.org). All of them, in addition to introducing amino acid substitutions (p.Lys45Asn, p.Lys45Asn, p.Arg46Lys, p.Arg46Lys, p.Arg46Ser, p.Arg46Ser), are predicted to slightly reduce the score of the putative ESS (from 0.92 to 0.75-0.80). Investigation of splicing patterns revealed that, with the exception of the c.135g/t, these mutations decreased the levels of normally spliced transcripts from 80% to ~40% (Fig. V.6, lanes 1-6).

The function of this region was further explored by well-established approaches such as i) multiple mutagenesis (three nucleotide changes, F9mut1) or ii) a deletion (6 nucleotides, F9ΔES1)(Fig. V.5A). The triple mutant F9mut1 resulted in the preferential usage of the

cryptic 5' ss and, conversely, in the remarkably decreased use of the authentic 5' ss (Fig. V.6B, lane 10). Coherently, albeit to a minor extent, the deletion in the F9ΔES1 minigene decreased the usage of the authentic 5' ss (Lane 8). RNA secondary structure might influence these splicing outputs (Qian and Liu 2014). Our prediction analysis (RNAfold) suggests that the accessibility of the cryptic 5' ss is maintained in the presence of nucleotide changes but appears to be reduced upon E1 deletion (Fig.V.5).

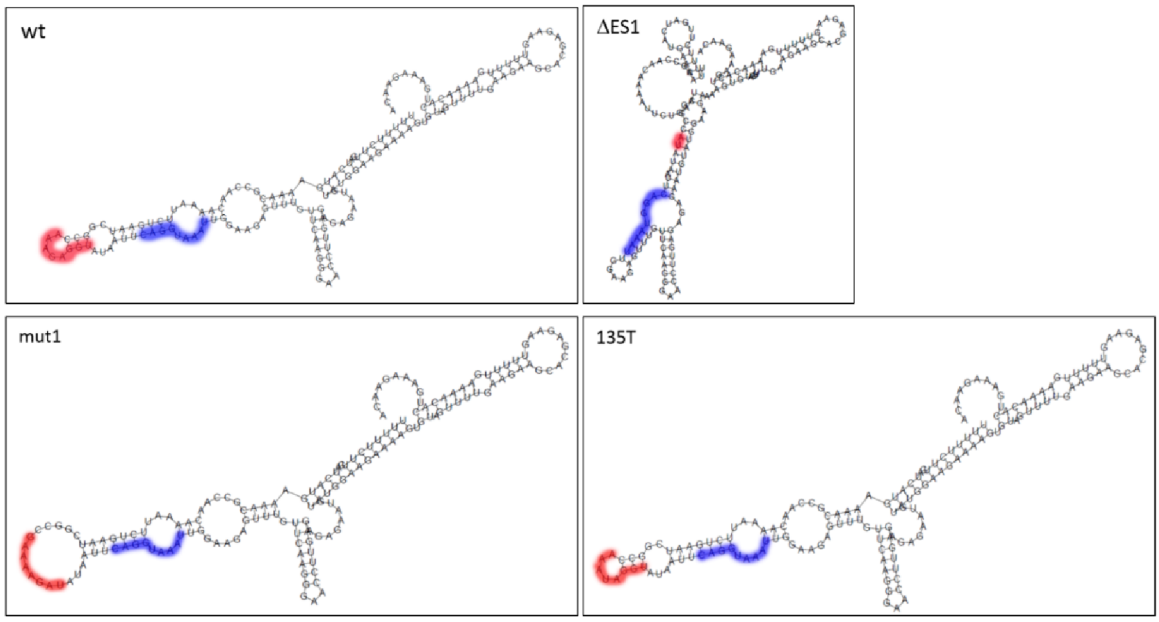


Figure V.5
Prediction by the RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) of the secondary RNA structure of the exon 2 region including the cryptic 5' ss (blue) and the regulatory element (red).

Therefore, it is temptative to speculate that the modest impact on splicing of F9ΔES1 minigene would arise from a compensatory effect in which the usage of the cryptic 5' ss is favored by the removal of the silencer element and concurrently disfavored by a reduced accessibility of the U1 snRNP.

Intriguingly, a 6-nucleotide deletion upstream of the putative ESS (F9ΔES2, Fig. V.6A), made as control, remarkably decreased use of the cryptic 5' ss (Fig. V.6B, lane 9), consistent with the bioinformatics prediction of ESEs in this region.

Taken together these multiple and complementary mutants point towards the presence of a network of positive and negative regulatory elements, with the natural missense changes mapping an ESS sequence down-regulating the adjacent cryptic 5' ss.

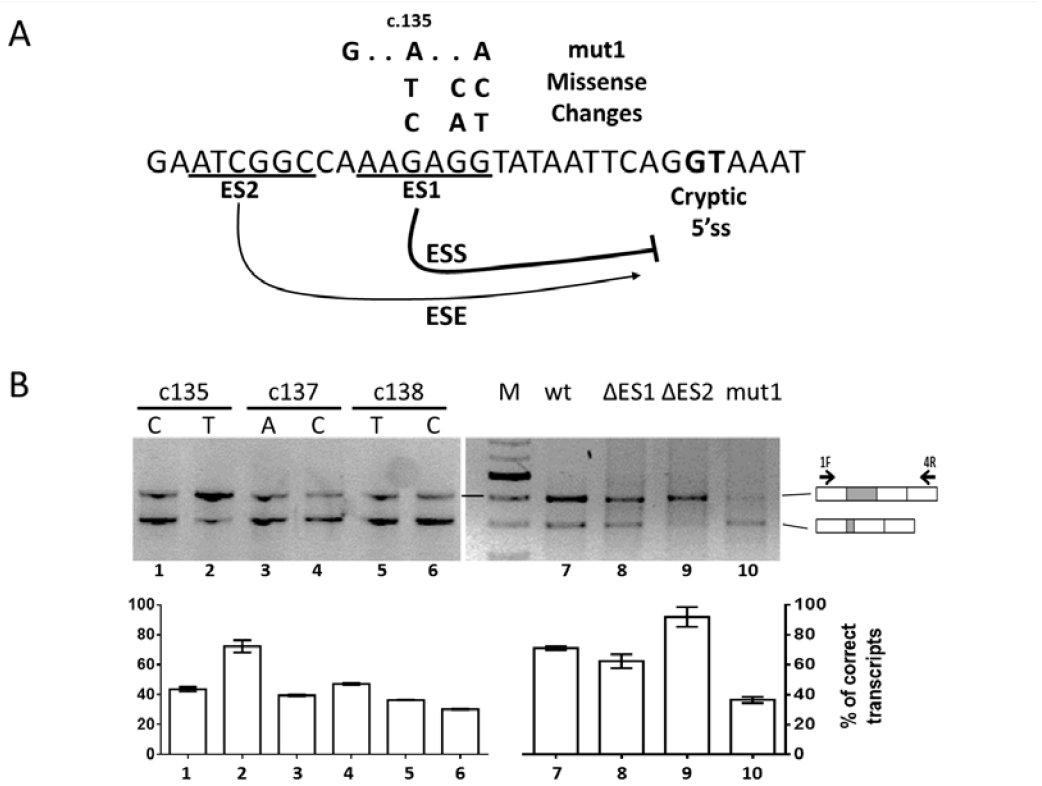


Figure V.6
The F9 exon 2 5'ss mutations induce usage of an exonic cryptic 5'ss regulated by cis-elements, including an ESS
A) Sequence of the F9 exon 2 region upstream of the cryptic 5'ss (bold). The underlined sequences indicate the examers (ES1, ES2) deleted to identify functional regulatory elements, which appear to act as silencer and enhancer (curved arrows) of the cryptic 5'ss. The nucleotide changes introduced in the ES1 element, either missense or artificial (mut1), are reported above together with a reference position in the F9 cDNA.
B) Evaluation of F9 alternative splicing patterns in BHK cells transiently expressing the F9wt (lane 7), missense variants (lane 1-6), mut1 (lane 10) or the deleted mutants ΔES1 (lane 8) and ΔES2. Histograms report the relative percentage of correct transcripts and results are expressed as mean ± standard deviation (SD) from three independent experiments. The schematic representation of the transcripts, and of primers used for the RT–PCR (arrows), is reported in the right panel.

Antisense oligonucleotides masking the cryptic 5'ss did not recover the selection of the defective *F9* exon 2 5'ss but rescued missense mutations within the ESS

We designed and synthesized two synthetic 2'-Omethylphosphothioate antisense oligoribonucleotides (AON360, AON353) to mask and inhibit the cryptic 5'ss (Fig. V.7).

Figure V.7
Sequence of the F9 gene region containing the ESS (underlined) and the cryptic (bold) and the authentic (italics) 5'ss with the schematic representation of the two antisense oligoribonucleotides AON353 and AON360.

To set up experimental conditions, we tested the AONs on the splicing pattern of the F9wt construct. At 10nM, the AON360 had its maximal effect and remarkably reduced the relative amount of the F9del form as compared to the correctly spliced one (Fig. V.8; Supplementary Material, Fig. S2A, lanes 1-2), thus indicating that it effectively inhibited the usage of the cryptic 5' ss.

On the other hand, the AON360, designed on the F9wt sequence and thus not perfectly matching the mutated ESS, was also able to partially recover correct splicing of the *F9* missense variants (Fig V.8; Supplementary Material, Fig. SV.1, lanes 3-14), sharing with the F9wt a functional authentic 5' ss.

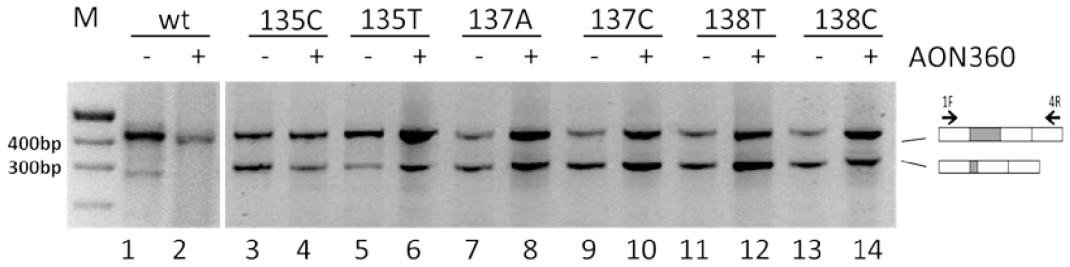


Figure SV.1 Effect of antisense oligonucleotides (AON) 360 on *F9* exon 2 mutants.
Evaluation of *F9* alternative splicing patterns in BHK transiently transfected with the F9wt (left, lanes 1-2) or minigenes harboring missense changes (right, lanes 3-14) in the presence (+) or absence (-) of 10nM AON360. The schematic representation of the transcripts, and of primers used for the RT–PCR (arrows), is reported in the right panel. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker.

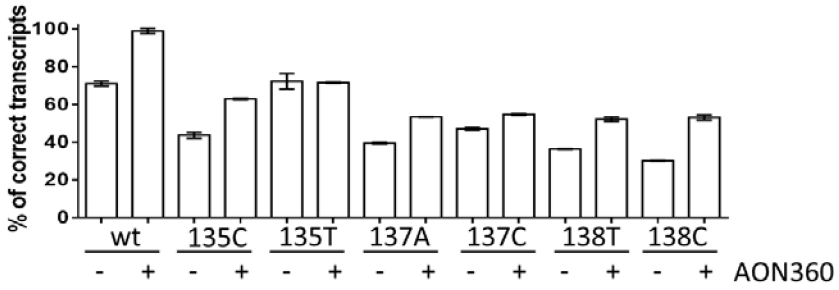


Figure V.8
Evaluation of F9 alternative splicing patterns in BHK cells transiently expressing the F9wt and the missense minigene variants alone (-) or with 10 nM AON360 (+). RT–PCR was conducted as in figure 1, and a representative gel is provided in Supplementary Fig. S1. Histograms report the relative percentage of correct transcripts and results are expressed as mean ± SD from three independent experiments.

Conversely, the antisense approach failed to correct the *F9* exon 2 5' ss mutations, and in addition induced complete skipping of exon 2 (Fig. V.9, lanes 2-11). Comparable results were obtained with the AON353 whereas a scrambled AON was ineffective

(Supplementary Material, Fig. SV.2, lanes 1-17).

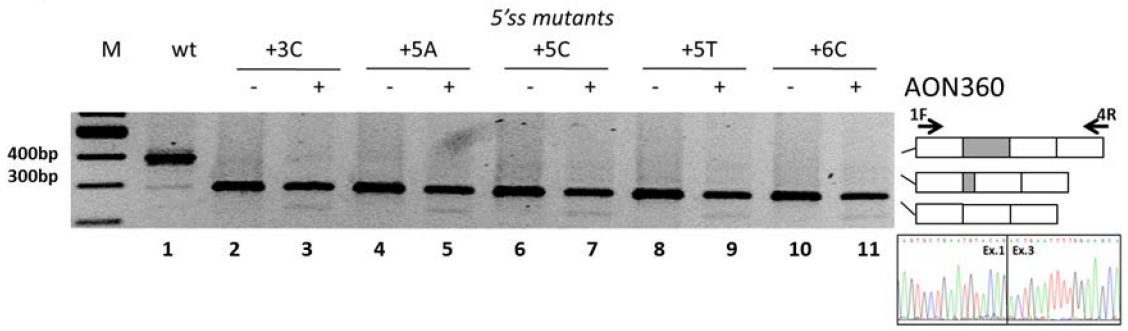


Figure V.9
Evaluation of F9 alternative splicing patterns in BHK cells expressing minigenes variants of the authentic 5' ss alone (-) or with 10nM AON360 (+). RT–PCR products, obtained with primers 1F and 4R (arrows in the right panel) were separated on 2% agarose gel. M, 100 bp molecular weight marker. The schematic representation of the transcripts is reported on the right. The chromatogram reports the sequence of the transcript lacking exon 2.

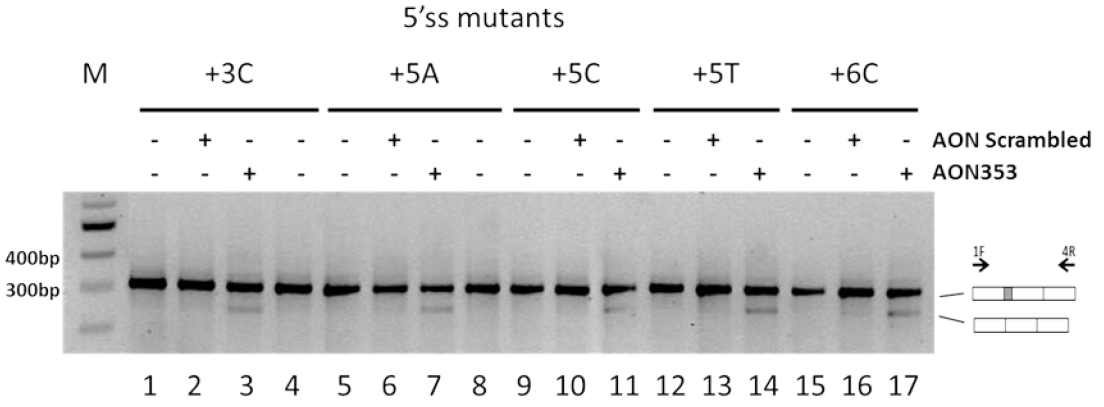


Figure SV.2 Effect of antisense oligonucleotides (AON) 353 on *F9* exon 2 mutants.
Evaluation of *F9* alternative splicing patterns in BHK expressing minigene variants of the authentic 5' ss alone (-) or with 10nM AON360 or scrambled AON (+). RT–PCR products, obtained with primers 1F and 4R (arrows) were separated on 2% agarose gel. M, 100 bp molecular weight marker. The schematic representation of the transcripts is reported on the right.

U1snRNAs with increased complementarity to the defective *F9* exon 2 5'ss have poor effects on the correct 5' ss selection

We designed a modified U1snRNA with perfect complementarity to the authentic *F9* exon 2 5' ss (U1F9wt) (Fig. V.10), which therefore has a single mismatch for each mutant.

Figure V.10
Sequence of the F9 gene region containing the ESS (underlined), the cryptic (bold) and the authentic (italics) 5'ss with the schematic representation of the sequence of the 5' tail of the compensatory U1snRNAs (right panel, where D=U,G or A).

Co-expression of the U1F9wt with each 5'ss mutant did not result in any appreciable rescue of splicing (Fig. V.11, lanes 2-11). On the other hand, the U1F9wt had a negligible effect on exon 2 definition even in the F9wt context, characterized by a strong authentic 5'ss (Supplementary Material, Fig. SV.3).

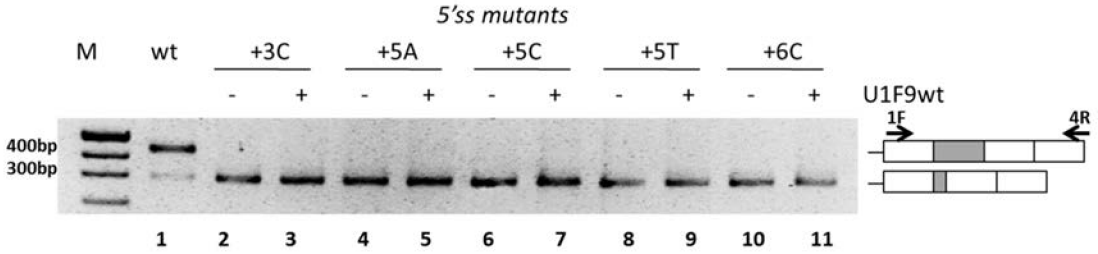


Figure V.11 U1snRNAs with improved (U1F9wt) or perfect (U1F9spec) complementarity to the mutated F9 exon 2 5'ss have poor correction effects.
Evaluation of F9 alternative splicing patterns in BHK cells transiently transfected with minigenes without (-) or with (+) the pU1F9wt (A) or the pU1F9spec (B).
The schematic representation of the transcripts and of primers used for the RT–PCR (arrows) are reported on the right. Amplified fragments were separated on 2% agarose gel. M, 100 bp molecular weight marker.

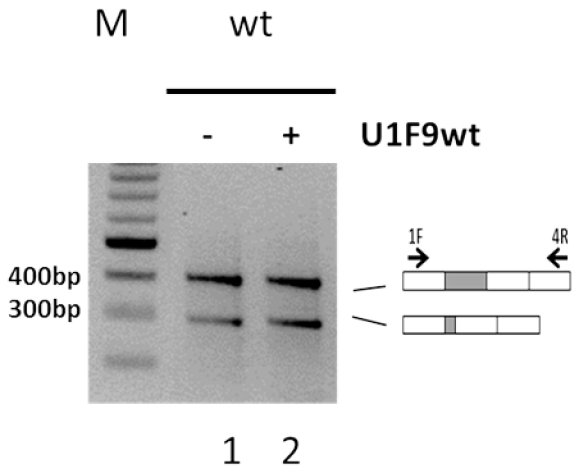


Figure SV.3 Control experiments for the activity of the U1F9wt and AON360.
Evaluation of *F9* alternative splicing patterns in BHK cells expressing the *F9*wt minigene alone (-) or with the U1F9wt (+). The schematic representation of the transcripts, and of primers used for the RT–PCR (arrows), is reported in the right panel. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker.

To further increase complementarity to mutant 5' ss, we created U1snRNAs specific for each mutation (U1F9spec; Fig. V.10). Complementation assays revealed that only the mutation +5C was partially corrected by the corresponding U1F9+5C variant, as indicated by the appreciable fraction of correct transcripts (26±3% of total FIX transcripts) (Fig. V.12, lane 6 vs 7). Conversely, the mutant-specific U1snRNA variants had no significant effects on aberrant splicing patterns generated by the expression of all the other 5' ss variants.

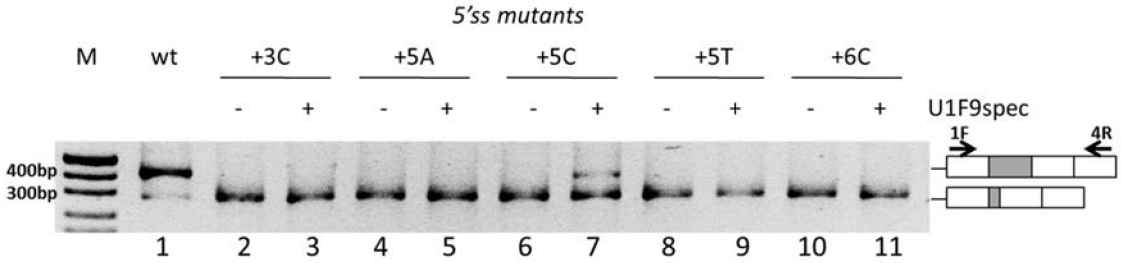


Figure V.12 U1snRNAs with improved (U1F9wt) or perfect (U1F9spec) complementarity to the mutated F9 exon 2 5'ss have poor correction effects.
Evaluation of F9 alternative splicing patterns in BHK cells transiently transfected with minigenes without (-) or with (+) the pU1F9wt (A) or the pU1F9spec (B).
The schematic representation of the transcripts and of primers used for the RT–PCR (arrows) are reported on the right. Amplified fragments were separated on 2% agarose gel. M, 100 bp molecular weight marker.

AONs and modified U1snRNAs have combinatorial correction effects

With the exception of the +5C mutant, neither the AON360, masking the cryptic 5' ss, nor the modified U1snRNAs, restoring complementarity to the 5' ss, were able to rescue exon 2 inclusion. We therefore explored the combination of these approaches to test their possible concerted effects. Strickingly, co-transfection of the AON360 with the mutant-specific U1snRNAs resulted in appreciable rescue of correctly spliced transcripts for the +3C (14±5 % of total transcripts), +5T (42±3%) and +6C (15±4%) mutations (Fig. V.13, lanes 3, 9, 11) but not the +5A mutation (lane 5). The combined approach did not significantly improved the correction of the +5C mutant (31±9%, lane 7) as compared to the U1F9spec alone (26±3%). To provide insights into the specificity, we also tested the U1F9spec together a scrambled AON on the +5T mutant, which failed to rescue splicing (Supplementary Material, Fig. SV.4). The U1-F9wt in combination with AON360 was ineffective for all mutants (data not shown).

Regarding the missense variants of the ESS, the additive contribution of the compensatory U1snRNA to the productive AON splicing rescue was reproducibly marginal (data not shown), a finding consistent with the presence of a normal authentic 5' ss.

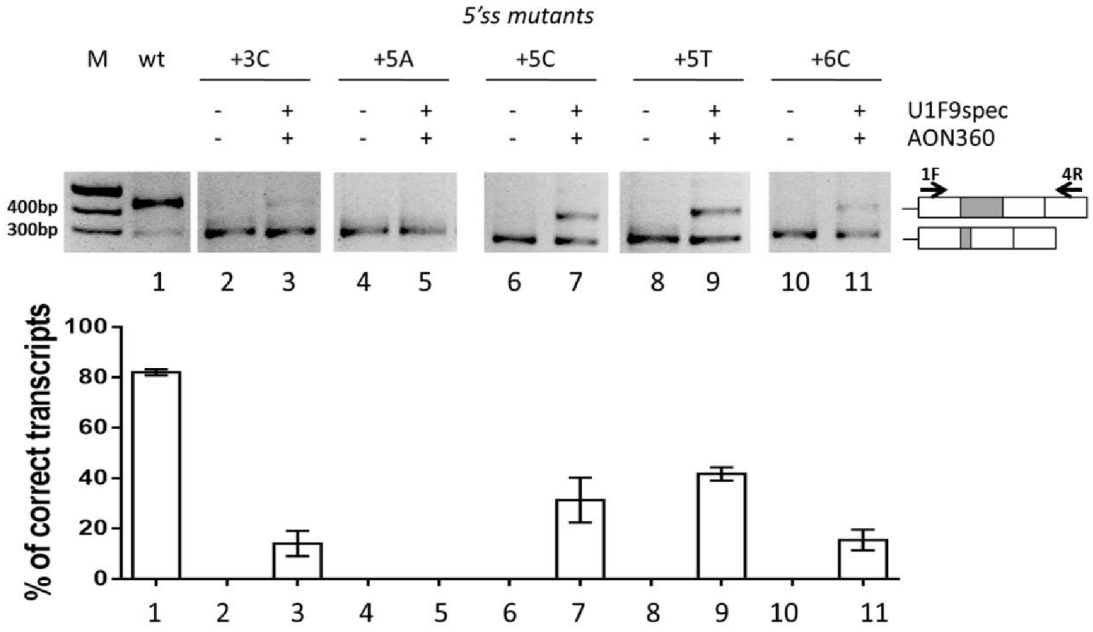


Figure V.13 Combinatorial correction effects of AON360 and mutation-specific U1snRNA variants.
Evaluation of F9 alternative splicing patterns in BHK cells transiently transfected with minigenes alone (-) or with the combination of AON360 (10nM) and the U1F9spec variants (+). RT–PCR and electrophoresis were conducted as in Fig. V.12. The schematic representation of the transcripts and of primers used (arrows) are reported on the right. The histograms report the percentage of correct transcripts expressed as means ± SD from three independent experiments.

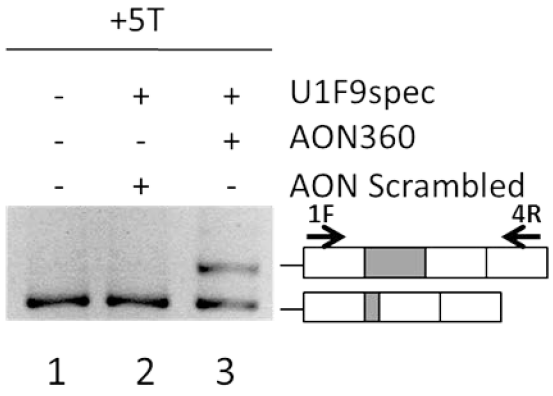


Figure SV.4 Control experiments for the activity of the U1F9wt and AON360.
Evaluation of *F9* alternative splicing patterns in BHK cells expressing the F9+5T minigenes variant alone (-) or with combinations of the U1F9spec with the scrambled AON (lane 2) or the AON360 (lane 3). RT–PCR products, obtained with primers 1F and 4R (arrows) were separated on 2% agarose gel. The schematic representation of the transcripts is reported on the right.

DISCUSSION

A crucial step in exon definition is represented by the selection of the proper 5'ss, which occurs in the presence of numerous sequences that resemble 5'ss but are not or inefficiently used in normal conditions (Roca et al. 2013). Among the signature of regulatory elements governing this choice and the definition of a given exon, the ESSs, which are less represented in constitutive exons (Královicová and Vorechovsky 2007), possess an intriguing and still poorly defined role.

Here we provide a paradigmatic example of the role of an ESS both in normal and pathological conditions, either in the presence of missense changes or of mutations at the authentic 5'ss. In particular, by exploiting the *F9* exon 2 context as study model, we demonstrated that both mutation types, through distinct effects, shift the balance toward the usage of a strong exonic cryptic 5'ss.

It has been suggested that the strength of ESS correlates with that of the 5'ss (Xiao et al. 2007). Intriguingly, the comparative analysis of the *F9* exon 2 across species (Fig. V.4) (Davidson et al. 2003) showed that the cryptic 5'ss is always paralleled by the conserved ESS. These observations support that the reported correlation between ESS and authentic 5'ss should be extended to, and is particularly relevant for cryptic 5'ss leading to unproductive splicing, as in the model of *F9* exon 2. Since the exon 2 encodes the γ carboxyl glutamic domain, which is absolutely required for the interaction of the serine protease with membranes, we infer that the ESS plays an essential role for the FIX protein function. By reducing the efficiency of unproductive splicing, the ESS would limit the impact of the cryptic 5'ss on the expression of functional FIX (Butenas et al. 2004).

On the other hand, natural mutations affecting the ESS would favor the cryptic 5'ss usage and reduce FIX expression. To support this mechanistic hypothesis we investigated all known missense mutations introducing changes within the ESS and potentially weakening it, and demonstrated that almost all of them favor the cryptic 5'ss recognition and significantly decrease the levels of correct FIX transcripts. As such, they have an additional, and probably significant, detrimental impact on FIX expression through combination of reduced amounts of correct FIX mRNA (~40%) with the alterations produced by amino acid substitutions. Interestingly, FIX plasma levels below 1%, which define a severe bleeding phenotype, were measured in the HB patients carrying these missense mutations.

Although we have not detailed the entire splicing regulatory network, these coherent

results indicate a functional role for the ESS, and offer an example of aberrant splicing caused by missense mutations that might represent a key determinant of the disease phenotypic variability.

The splicing patterns in the *F9* exon 2 context indicate a competition between 5'ss with comparable strength, and a modulatory role of the ESS that disfavors the selection of the cryptic 5'ss. The nucleotide changes at positions +3, +5 and +6 remarkably reduce the strength of the authentic 5'ss and its complementarity to the U1snRNA 5' tail, thus shifting the balance to the usage of the cryptic one.

In the attempt to rescue splicing we created antisense oligonucleotides designed to mask the cryptic 5'ss, a strategy that has been successfully used in several human disease models (Spitali and Aartsma-Rus 2012). This is a straightforward correction approach in the presence of the authentic 5'ss, as demonstrated by the remarkable inhibition of the cryptic 5'ss and rescue of correct splicing obtained with the AON360 in the context of the wild-type and of missense mutants. Consistently, in the presence of mutations impairing the authentic 5'ss and thus its recognition by the U1snRNP, this antisense strategy was ineffective and led to exon skipping.

We therefore focused on improving complementarity of the 5' tail of the U1snRNA to the defective 5'ss, which has repeatedly been shown to rescue splicing in many models of human disease characterized by different type of mutations (at 5'ss, at 3'ss, within exons) impairing exon-definition (Baralle et al. 2003; Susani et al. 2004; Pinotti et al. 2008; Pinotti et al. 2009; Tanner et al. 2009; Pinotti et al. 2011; Schmid et al. 2011; Glaus et al. 2011; Fernandez Alanis et al. 2012; Balestra et al. 2014). However, with the exception of the +5C mutant, the compensatory U1snRNA variants failed to elicit appreciable correction effects, a finding that might be attributable to the strong competition by the cryptic 5'ss.

We hypothesized that the concurrent increase in the efficiency of correct 5'ss selection by engineered U1snRNA and inhibition of the exonic cryptic 5'ss by antisense molecules would overcome the splicing defect. Our data demonstrated that the combination of mutant-specific U1snRNAs and AON360 was able to significantly increase the selection of the authentic 5'ss and rescue splicing (from 0 up to 40% of correct transcripts) in the presence of all, but one, mutations. Noticeably, the U1F9wt that differs for one mismatch only from the mutant-specific U1snRNAs (Fig. V.10) was ineffective on *F9* exon 2 5'ss mutants, thus strengthening the notion that the complementarity requirements between the U1snRNAs 5' tail and the 5'ss for a functional splicing outcome are far from being known (Roca et al. 2013).

In conclusion, by molecular characterization of a series of severe HB mutations affecting the *F9* exon 2 and its 5' ss, we propose, within a network of splicing regulatory elements, a function for ESS in the down-regulation of cryptic 5' ss in exons that are essential for protein function, which better defines ESS relevance in normal and pathological conditions.

Based on the evidence for the altered interplay among the ESS, cryptic and authentic 5' ss as a disease-causing mechanism, we produced novel experimental insights into the combinatorial activity of antisense oligonucleotides and compensatory U1snRNA in inducing splicing correction.

Material and Methods

Creation of expression vectors

To create the pF9wt vector, the genomic regions of human *F9* gene (NG_007994.1) spanning i) the first codon of exon 1 to nt. +173 in intron 1 (fragment 1), ii) nt.-264 of intron 1 to nt. +2254 of intron 3 (fragment 2) and iii) nt. -1435bp of intron 3 to the last codon of exon 4 (fragment 3) were amplified from genomic DNA of a normal subject with primers Ex1F-Ex1R, Ex2F-Ex3R and Ex3F-Ex4R using high-fidelity PfuI DNA-polymerase (Transgenomic, Glasgow, UK). The *F9* regions were sequentially cloned into the expression vector pcDNA3 (LifeTechnologies, Carlsbad, CA, USA) by exploiting the *KpnI-BamHI* (fragment 1), *BamHI-NotI* (fragment 2) and *NotI-XhoI* (fragment 3) restriction sites inserted within primers.

To create the pF9+3c, pF9+5a, pF9+5c, pF9+5t, pF9+6c, pF9c135c, pF9c135t, pF9c137a, pF9c137c, pF9c138t, pF9c138c, pF9wt-XX, pF9mut1, pF9ΔES1 and pF9ΔES2 vectors, the nucleotide changes/deletion were introduced into the pF9wt minigene by the QuickChange II Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA).

The pU1F9wt, pU1F9+3C, pU1F9+5C, pU1F9+5A, pU1F9+5T and pU1F9+6C expression vectors for the modified U1snRNAs were created by replacing the sequence between the sites BclI and BglII with oligonucleotides as previously reported (20).

Sequences of oligonucleotides are provided in Supplementary Material, Table S1. All vectors have been validated by sequencing.

Oligonucleotide		Sequence
Creation of the pF9wt		
1F	**forward**	5`-aaaggtaccatgcagcgcgtgaac-3`
1R	**reverse**	5`-aaaggatccgaagaaaacctagctaacaaagaacc-3`
2F	**forward**	5`-aaaggatccatccaaagtaaattcaaatatg-3`
3R	**reverse**	5`-aaagcggccgcggatgttttactgaaagggtctgg-3`
3F	**forward**	5`-aaagcggccgctgaataattcacatttgttggtc-3`
4R	**reverse**	5`-aaactcgagctaattcacagttctttccttcaaatcc-3`
Mutant *F9* minigenes at the authentic 5`ss		
F9 IVS2+3G>C	**forward**	5`-aaagaacagtcagtatttccacataa-3`
	reverse	5`-ttatgtggaaatactgactgttcttt-3`
F9 IVS2+5G>A/C/T	**forward**	5`-aaagaacagtgahtatttccacata-3`
	reverse	5`-tatgtggaaatadtcactgttcttt-3`
F9 IVS2+6T>C	**forward**	5`-gaacagtgagcatttccacataataccc-3`
	reverse	5`-gggtattatgtggaaatgctcactgttc-3`
Mutant *F9* minigenes at the ES1/ES2		
F9 c.135G>C/T	**forward**	5`-ctgaatcggccaaayaggtataattcagg-3`
	reverse	5`-cctgaattatacctrtttggccgattcag-3`
F9 c.137G>A/C	**forward**	5`-gaatcggccaaagamgtataattcagg-3`
	reverse	5`-cctgaattatacktctttggccgattc-3`
F9 c.138G>T/C	**forward**	5`-gaatcggccaaagagytataattcagg-3`
	reverse	5`-cctgaattatarctctttggccgattc-3`
F9 mut1	**forward**	5`-tgaatcggccgaaaagatataattca-3`
	reverse	5`-tgaattatatcttttcggccgattca-3`
pF9ΔES1	**forward**	5`-ccaacaaaattctgaatcggccatataattcaggtaaattggaaga-3`
	reverse	5`-tcttccaatttacctgaattatatggccgattcagaattttgttgg-3`
pF9ΔES2	**forward**	5`-tcttccaatttacctgaattatatggccgattcagaattttgttgg-3`
	reverse	5`-ttacctgaattatacctctttgtcagaattttgttggcgttttcatg-3`
Modified U1 snRNAs for *F9* exon 2 5'ss		
U1F9wt	**forward**	5'-gatctgatACTCACTGTgcaggggagataccat-3'
	reverse	5'-gatcatggtatctcccctgcACAGTGAGTatca-3'
U1F9+3C	**forward**	5'-gatctgat ACTGACTGTgcaggggagataccat-3'
	reverse	5'-gatcatggtatctcccctgcACAGTCAGTatca-3`
U1F9+5A	**forward**	5'-gatctgatATTCACTGTgcaggggagataccat-3'
	reverse	5'-gatcatggtatctcccctgcACAGTGAATatca-3'
U1F9+5C	**forward**	5'-gatctgatAGTCACTGTgcaggggagataccat-3'
	reverse	5'-gatcatggtatctcccctgcACAGTGACTatca-3'
U1F9+5T	**forward**	5'-gatctgatAATCACTGTgcaggggagataccat-3'
	reverse	5'-gatcatggtatctcccctgcACAGTGATTatca-3'
U1F9+6C	**forward**	5'-gatctgatGCTCACTGTgcaggggagataccat-3'
	reverse	5'-gatcatggtatctcccctgcACAGTGAGCatca-3'

Table S1: Sequences of oligonucleotides

Antisense oligonucleotides

AON353 and AON360 against the *F9* exon 2 cryptic 5'ss, and the scrambled AON, contain 2′-O-methyl modified ribonucloetides and full-length phosphorothioate backbone (Rimessi et al. 2010).

Expression in mammalian cells and mRNA studies

Baby Hamster Kidney (BHK) cells were cultured as previously described (13). Cells were seeded on twelve-well plates and transfected with Lipofectamine 2000 reagents (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's protocol.

One microgram of pF9 minigenes was transfected alone, with the AON (10nM) or a molar excess (1.5X) of the pU1 plasmids. Total RNA was isolated 24h post-transfection with Trizol (Life Technologies, Carlsbad, CA, USA) and reverse-transcribed and amplified using the SuperScript III One-Step RT-PCR System (Life Technologies, Carlsbad, CA, USA) with primers ex1F and ex4R. A similar approach was used to evaluated *F9* splicing patterns in human liver.

The correct and F9del fragments were also cloned into the pGEM vector and used as templates at known concentrations to verify the amplification efficiency, which appeared to be comparable (data not shown).

Densitometric analysis for the quantification of correct and aberrant transcripts was performed using the ImageJ software.

For denaturing capillary electrophoresis analysis, the RT-PCR amplified fragments were fluorescently labelled by using primers 1F and the 4R labelled with FAM and run on a ABI-3100 instrument.

Computational analysis

Computational prediction of splice sites and or splicing regulatory elements was conducted by using the http://www.fruitfly.org/seq_tools/splice.html, the Human Splicing Finder (http://www.umd.be/HSF/) and Rescue-ESE (http://genes.mit.edu/burgelab/rescue-ese/) online softwares. Computational prediction of RNA secondary structure was performed using the RNAfold software.

REFERENCES

-Ars E, Serra E, García J, Kruyer H, Gaona A, Lázaro C, Estivill X. Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. Hum Mol Genet. 2000 Jan 22;9(2):237-47. Erratum in: Hum Mol Genet 2000 Mar 1;9(4):659.

-Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M. An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice. J Thromb Haemost. 2014 Feb;12(2):177-185. doi: 10.1111/jth.12471.

-Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, Ffrench-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 Mar;40(3):220-2.

-Bruno IG, Jin W, Cote GJ. Correction of aberrant FGFR1 alternative RNA splicing through targeting of intronic regulatory elements. Hum Mol Genet. 2004 Oct 15;13(20):2409-20. Epub 2004 Aug 27.

-Butenas S, Orfeo T, Gissel MT, Brummel KE, Mann KG. The significance of circulating factor IXa in blood. J Biol Chem. 2004 May 28;279(22):22875-82. Epub 2004 Mar 23.

-Cartegni L, Krainer AR. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat Genet. 2002 Apr;30(4):377-84. Epub 2002 Mar 4.

-Cavallari N, Balestra D, Branchini A, Maestri I, Chuamsunrit A, Sasanakul W, Mariani G, Pagani F, Bernardi F, Pinotti M. Activation of a cryptic splice site in a potentially lethal coagulation defect accounts for a functional protein variant. Biochim Biophys Acta. 2012 Jul;1822(7):1109-13. doi:10.1016/j.bbadis.2012.03.001. Epub 2012 Mar 9.

-Chen M, Manley JL. Mechanisms of alternative splicing regulation: insightsfrom molecular and genomics approaches. Nat Rev Mol Cell Biol. 2009 Nov;10(11):741-54. doi: 10.1038/nrm2777. Epub 2009 Sep 23. Review.

-Davidson CJ, Tuddenham EG, McVey JH. 450 million years of hemostasis. J Thromb Haemost. 2003 Jul;1(7):1487-94. Review.

-Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. Hum Mol Genet. 2012 Jun 1;21(11):2389-98. doi:10.1093/hmg/dds045. Epub 2012 Feb 23.

-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 May;19(5):936-41. doi: 10.1038/mt.2011.7. Epub 2011 Feb 15.

-Green PM, Bentley DR, Mibashan RS, Nilsson IM, Giannelli F. Molecular pathology of haemophilia B. EMBO J. 1989 Apr;8(4):1067-72.

-Havens MA, Duelli DM, Hastings ML. Targeting RNA splicing for disease therapy. Wiley Interdiscip Rev RNA. 2013 May-Jun;4(3):247-66. doi: 10.1002/wrna.1158. Epub 2013 Mar 19. Review.

-Kervestin S, Jacobson A. NMD: a multifaceted response to premature translational termination. Nat Rev Mol Cell Biol. 2012 Nov;13(11):700-12. doi:10.1038/nrm3454. Epub 2012 Oct 17. Review.

-Královicová J, Vorechovsky I. Global control of aberrant splice-site activation by auxiliary splicing sequences: evidence for a gradient in exon and intron definition. Nucleic Acids Res. 2007;35(19):6399-413. Epub 2007 Sep 18.

-Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010 Jan 28;463(7280):457-63. doi: 10.1038/nature08909. Review.

-Nlend Nlend R, Meyer K, Schümperli D. Repair of pre-mRNA splicing: prospects for a therapy for spinal muscular atrophy. RNA Biol. 2010 Jul-Aug;7(4):430-40. Epub 2010 Jul 1. Review.

-Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet. 2004 May;5(5):389-96. Review.

-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 Mar 1;111(5):2681-4. Epub 2007 Dec 21.

-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 Jun 18;113(25):6461-4. doi: 10.1182/blood-2009-03-207613. Epub 2009 Apr 22.

-Pinotti M, Bernardi F, Dal Mas A, Pagani F. RNA-based therapeutic approaches for coagulation factor deficiencies. J Thromb Haemost. 2011 Nov;9(11):2143-52.doi: 10.1111/j.1538-7836.2011.04481.x. Review.

-Porensky PN, Mitrpant C, McGovern VL, Bevan AK, Foust KD, Kaspar BK, Wilton SD, Burghes AH. A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. Hum Mol Genet. 2012 Apr 1;21(7):1625-38. doi:10.1093/hmg/ddr600. Epub 2011 Dec 20.

-Qian W, Liu F. Regulation of alternative splicing of tau exon 10. Neurosci Bull. 2014 Apr;30(2):367-77. doi: 10.1007/s12264-013-1411-2. Epub 2014 Mar 14. Review.

-Rimessi P, Fabris M, Bovolenta M, Bassi E, Falzarano S, Gualandi F, Rapezzi C, Coccolo, Perrone D, Medici A, Ferlini A. Antisense modulation of both exonic and intronic splicing motifs induces skipping of a DMD pseudo-exon responsible for x-linked dilated cardiomyopathy. Hum Gene Ther. 2010 Sep;21(9):1137-46. doi:10.1089/hum.2010.010.

-Roca X, Krainer AR, Eperon IC. Pick one, but be quick: 5' splice sites and the problems of too many choices. Genes Dev. 2013 Jan 15;27(2):129-44. doi:10.1101/gad.209759.112. Review.

-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 Jul;32(7):815-24. doi: 10.1002/humu.21509.

-Spitali P, Aartsma-Rus A. Splice modulating therapies for human disease. Cell. 2012 Mar 16;148(6):1085-8. doi: 10.1016/j.cell.2012.02.014. Review.

-Stadler MB, Shomron N, Yeo GW, Schneider A, Xiao X, Burge CB. Inference of splicing regulatory activities by sequence neighborhood analysis. PLoS Genet.2006 Nov 24;2(11):e191. Epub 2006 Sep 28.

-Sterne-Weiler T, Howard J, Mort M, Cooper DN, Sanford JR. Loss of exon identity is a common mechanism of human inherited disease. Genome Res. 2011 Oct;21(10):1563-71. doi: 10.1101/gr.118638.110. Epub 2011 Jul 12.

-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 Sep;24(3):225-35.

-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 Feb;30(2):255-63. doi:10.1002/humu.20861. Erratum in: Hum Mutat. 2009 Feb;30(2):274.

-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.

-Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009 Feb 20;136(4):701-18. doi: 10.1016/j.cell.2009.02.009. Review.

-Xiao X, Wang Z, Jang M, Burge CB. Coevolutionary networks of splicing cis-regulatory elements. Proc Natl Acad Sci U S A. 2007 Nov 20;104(47):18583-8. Epub 2007 Nov 12.

-Zhang XH, Chasin LA. Computational definition of sequence motifs governing constitutive exon splicing. Genes Dev. 2004 Jun 1;18(11):1241-50. Epub 2004 May 14.

VI

Characterization and correction of the ATR c.a2101g synonymous change causing Seckel syndrome

Daniela Scalet[1], Dario Balestra[1], Sara Rohban[2], Matteo Bovolenta[1], Stefano Campaner[2] and Mirko Pinotti[1]

[1]Department of Life Science and Biotechnology, University of Ferrara and LTTA, Ferrara, Italy and [2] Center for Genomic Science, Fondazione Istituto Italiano di Tecnologia (IIT), Via Adamello 16, Milano 20139, Italy

Manuscript in preparation

As detailed in the introduction (Chapter II.2.2), Seckel syndrome 1 or ATR-Seckel is an extremely rare autosomal recessive neurodevelopmental disorder, which is caused by mutations within the gene encoding ataxia-telangiectasia and RAD3-related protein (ATR). ATR protein, likes Ataxia-Telengectasia Mutated (ATM) protein involved in ataxia-telengectasia disease, is a phosphatidylinositol 3-kinase like kinase (Zhou and Elledge 2000; Durocher and Jackson 2001; Shiloh 2001) playing an important role in DNA damage response. In 2003, O'Driscoll et all reported the first human ATR mutation causative of SS-1, the synonymous change c.A2101G. The mutation occurs in exon 9 and has been identified in homozygous condition in two unrelated Pakistani families. In patients' fibroblast, the mutation, through an unknown mechanism, leads to exon 9 skipping and synthesis of a shorter mRNA harboring a premature stop codon. However, trace levels of correct transcripts are present, which explain the survival of the affected SS-1 patients, although with a very severe clinical phenotype. While knockout mouse models of ATM have been reported and are valuable tools to dissect ATM function details, complete defective ATR mice are lethal at pre-gastrulation level. Therefore, despite the apparent overlapping function in response to DNA damage, ATR is essential for viability, while ATM is not. The sequences of ATR exon 8 – 10 are highly homologous between human and mouse, but the relative introns differ significantly, making a mouse model of the disease hardly feasible. In fact, the attempt to introduce the mutation into the equivalent residue in mouse ATR gene did not impair the splicing of the affected exon (Ragland et al. 2009). To recapitulate the splicing defect caused by the A2101G mutation, Murga et al. introduced the human genomic region including the affected exons, and relative introns, into the corresponding mouse locus. This approach created MEF cells that fully recapitulated the splicing defect and provided an animal model that almost reproduced all the clinical features of the disease. Despite these findings, the molecular mechanisms underlying the aberrant splicing caused by the synonymous c.A2101G mutation are unknown, and no therapeutic approaches are available yet.

In recent years, therapeutic strategies acting at RNA levels have gained promising results. Among these, the application of modified U1snRNAs, the RNA component of the ribonucleoprotein U1snRNP with an essential role in early steps of spliceosome assembly, shown the therapeutic potential in rescuing exon-skipping inducing mutations located at donor (5'ss) or acceptor (3'ss) splice sites, or within exons (Fernandez et al. 2012). The ability of modified U1 in rescuing splicing defect has been proven *in vitro* and recently *in vivo* for various genes associated to inherited diseases, such as FVII deficiency, SMA,

Cystic Fybrosis, and other (Baralle et al. 2003; Susani et al. 2004; Pinotti et al. 2008; Pinotti et al. 2009; Tanner et al. 2009; Glaus et al. 2011; Pinotti et al. 2011; Schmid et al. 2011; Fernandez Alanis et al. 2012; Balestra et al. 2014). Moreover, the tailored application of modified U1 with antisense oligonucleotides (ASO), used to mask splicing regulatory elements, proved to be successful in rescuing splicing mutations that failed to be corrected by modified U1snRNA alone (Balestra et al. 2015).

In the present study, we characterized at molecular level the effect of the c.A2101G synonymous change and provided a tailored therapeutic approach based on antisense oligonucleotide and modified U1snRNA. Our data demonstrate that the mutation c.A2101G, located within a weakly defined exon 9, creates an Exon Splicing Silencer (ESS) that impairs exon definition and leads to aberrant splicing. Moreover, we provided insights that a modified U1 with the ability to increase exon definition, alone or in combination with ASO masking the ESS, can efficiently rescue ATR expression at both RNA and protein levels. The correction was evident either *in vitro* or in $ATR^{S/S}$ MEF cells.

RESULTS

The mutation A2101G induces exon skipping of a poorly defined ATR exon 9.

To understand the molecular mechanism associated to the SS-1 caused by the synonymous change A2101G located in the exon 9 of ATR gene, the genomic region spanning intron 8 through intron 10 was cloned into the pTB expression vector (Fig. VI.1A), a plasmid useful to investigate the splicing process (Pagani et al. 2003). Expression of pATRwt minigene in HEK cells resulted in partial exon inclusion, with exon 9 skipping accounting for 70% of transcripts (Fig. VI.1B) and thus indicating that exon 9 is poorly recognized in normal condition. This finding was confirmed by bioinformatics analysis of splicing elements showing that the donor and acceptor splice sites (5'ss and 3'ss, respectively) of exon 9 are poorly defined (both 0.54/1.00).

In splicing assays, the introduction of the mutation (pATR2101G) induced complete exon skipping, with residual levels (4%) of correct transcripts.

In both conditions, RT-PCR analysis and sequencing also revealed the presence of two aberrant transcripts originated by the use of two cryptic 5'ss located upstream the mutation under investigation.

To investigate the relationship between exon definition and the A2101G mutation, we assessed the impact of the mutation in an artificial context in which exon 9 definition was

improved by increasing the 5'ss and 3'ss scores. Noticeably, the mutation exerted a negligible impact on exon 9 processing in the case of perfect 3'ss, with correct transcripts accounting for 60% and comparable to what observed in wt condition. In the case of perfect 5'ss, the mutation has no effect on splicing, with complete exon inclusion, a situation even better than that observed with pATRwt (Fig. VI.1B).

Taken together these findings show that the mutation impairs exon definition and that this occurs only in the naturally poorly defined exon 9 context.

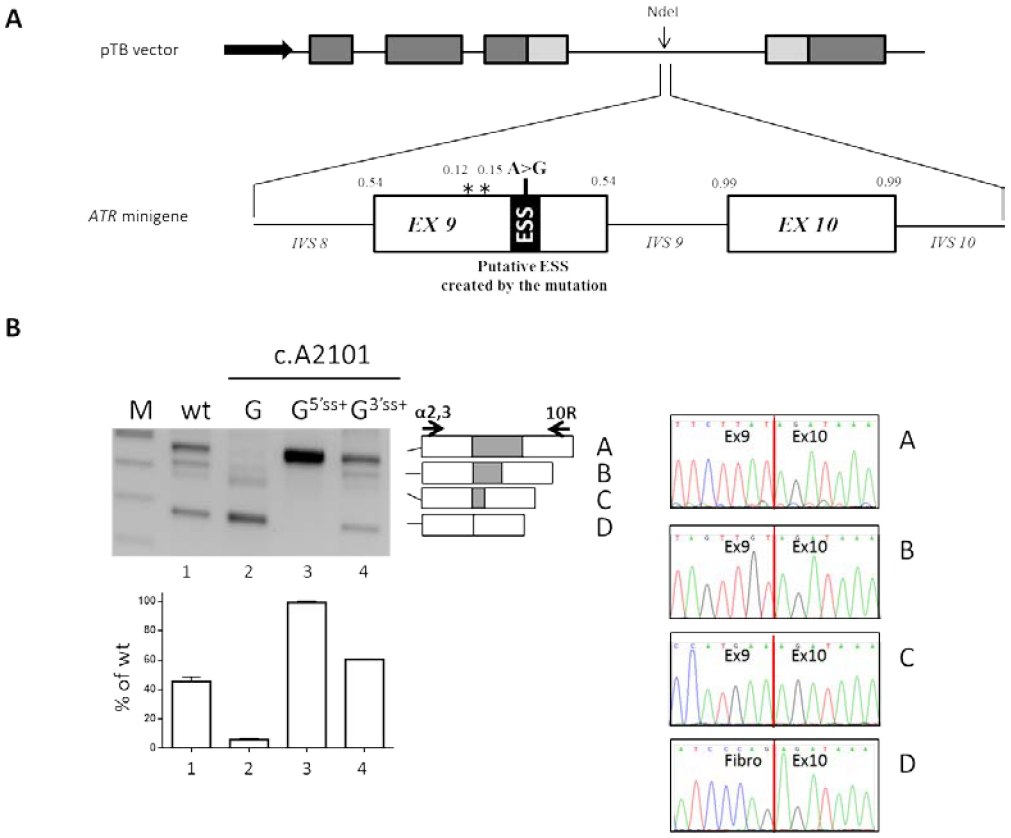


Figure VI.1 Splicing features of the human ATR exon 9 context and alternative splicing patterns
A) Schematic representation of the ATR genomic sequence cloned as minigene in the pTBNdeI vector. Exonic and intronic sequences are represented by boxes and lines, respectively. Exons of globin gene are indicated by grey boxes, while exon of fibronectin gene is indicated in light grey boxes. The scores of the splice sites (5'ss and 3'ss, authentic and cryptic), together with position of mutation A2101G under investigation, are indicated above the ATR minigene. Asterisks and the black boxe represent the cryptic donor splice sites and the putative exonic splicing silencer (ESS) created by the mutation, respectively.
B) Evaluation of ATR alternative splicing patterns in Hek293 cells transiently transfected with minigenes. The schematic representation of the normal and aberrant transcripts, and of primers used for the RT-PCR (arrows), is reported in the central panel. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. Chromatograms of transcripts are reported on the right. Histograms indicate the relative percentage of correctly spliced transcripts upon densitometric analysis of bands, which is expressed as mean ± standard deviation (SD).

The A2101G mutation induces exon 9 skipping by creating an ESS and can be efficiently rescued by an ASO and/or modified U1snRNA-based approach.

Bioinformatic analysis of exon 9 seeking for exonic regulatory sequences shown that the mutation creates a novel exonic splicing silencer (ESS), an element that is not predicted in wt condition.

To corroborate this finding, we designed and synthesized a synthetic 2'-Omethylphosphothioate antisense oligoribonucleotide (ASO) masking the ESS originated by the mutation (Fig. VI.2A). Co-transfection of HEK cells with pATR2101G and increasing amount (50 to 200nM) of ASO resulted in complete synthesis of normally spliced transcripts (Fig. VI.2B). Conversely, the ASO had no effect on the splicing pattern of the pATRwt, confirming the absence of the ESS in the wt context.

Data from the ATR minigene harboring the ameliorated 5'ss prompted us to explore a correction strategy based on a compensatory U1 with perfect complementarity with the authentic 5'ss ($pU1^{ATR}$) (Fig. VI.3A). Co-expression of the $pU1^{ATR}$ in HEK cells expressing the pATR2101G minigene rescued exon skipping. In fact, in the presence of a molar excess (1.5X) of $pU1^{ATR}$, the levels of correct transcript accounted for 62%. This approach also improved exon definition in the wt context, thus further demonstrating that exon 9 is poorly defined. On the other hand, co-expression of the pU1wt had negligible effects on splicing patterns in both conditions, thus suggesting that the correction is U1 specific.

In the attempt to restore complete exon definition, we combined the ASO and U1-mediated approaches. Co-delivery of ASO and $pU1^{ATR}$ in cells expressing the pATR2101G minigene resulted in a robust correction, with normally spliced transcripts that accounted for 77%. Comparable results were obtained in cells expressing the wt minigene, indicating a strong exon definition mediated by the U1ATR.

Altogether, these findings are consistent with the presence of an ESS originated by the presence of the mutation and this element induces the skipping of weakly defined exon 9. Moreover, the antisense and the U1-mediated strategies might provide a potential therapeutic approach to rescue ATR expression in the Seckel disease.

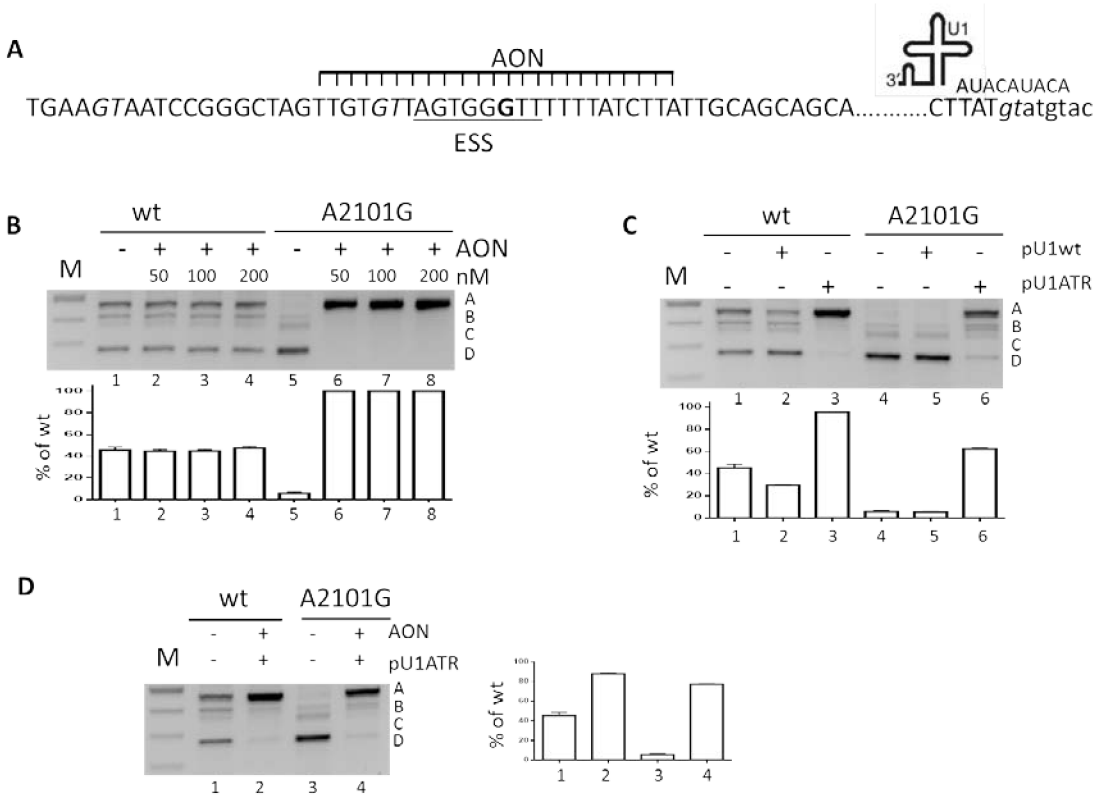


Figure VI.2. Regulation of ATR splicing by ESS and modulation by ASO and/or modified U1snRNA.
A) Sequences of ATR exon 9 surrounding the mutation c.A2101G and of the modified U1snRNA (U1ATR) with the 5' tail sequence placed above the corresponding ATR 5'ss target. The highly conserved dinucleotides GT belonging to the cryptic donor splices and the mutation under investigation are indicated in italics and bold, respectively. Exonic and intronic sequences are respectively showed in upper and lower case. The underlined sequence represents the eptamer predicted by the bioinformatics analysis and acting as silencer. B) Evaluation of ATR alternative splicing patterns in Hek293 cells transfected with the wt (pwt) or mutated (pA2101G) minigenes alone, or in combination with different concentration of ASO targeting the ESS. The normal and aberrant transcripts are represented as letters (A to C). Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. Histograms indicate the relative percentage of correctly spliced transcripts upon densitometric analysis of bands, which is expressed as mean ± standard deviation (SD). C) Evaluation of ATR alternative splicing patterns in Hek293 cells transfected with the wt (pwt) or mutated (pA2101G) ATR minigenes alone, or in combination with plasmid expressing the wt (pU1wt) or modified U1snRNA (pU1ATR) targeting the exon 9 5'ss. The normal and aberrant transcripts are represented as letters (A to C). Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. Histograms indicate the relative percentage of correctly spliced transcripts upon densitometric analysis of bands, which is expressed as mean ± standard deviation (SD). D) Evaluation of ATR alternative splicing patterns in Hek293 cells transfected with the wt (pwt) or mutated (pA2101G) ATR minigenes alone, or in combination with ASO targeting the ESS and modified U1snRNA (pU1ATR)-expressing plasmid. The normal and aberrant transcripts are represented as letters (A to C). Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. Histograms indicate the relative percentage of correctly spliced transcripts upon densitometric analysis of bands, which is expressed as mean ± standard deviation (SD).

Rescued in humanized MEF cells of ATR expression at RNA and protein levels by modified U1^{ATR} delivered by a lentivirus vector.

To evaluate the impact of the U1-mediated rescue at protein level, we switched from the *in vitro* model to humanized MEF cells (MEF$^{S/S}$), exploited to create the mouse model of Seckel disease. To stably introduce the foreign gene into fibroblast DNA, we introduced the modified U1 into a lentivirus (LV) that co-expressed the GFP as tracking gene. To determine whether expression of modified U1 was able to efficiently rescue ATR expression in MEF cells, we initially transduced MEF$^{S/S}$ cells with LV-U1ATR. To assess the overall LV transduction efficiency, we assessed the amount of GFP-positive MEF$^{S/S}$ cells. Transduction of fibroblast cells with LV-U1ATR resulted in 85% of cells expressing the GFP tracking gene, indicating an overall good transduction efficiency (Fig VI.3B). Conversely, assessment of U1ATR expression in MEF$^{S/S}$ shows an unexpected reduced expression level of U1snRNA, although the endogenous strong promoter drives its expression. While ATR expression in MEF$^{S/S}$ cells recapitulated the aberrant splicing profile observed *in vitro*, with traces of correctly spliced transcript, expression of U1ATR partially restore exon definition, as observed by increase of the correctly spliced transcript with the concurrent decrease of the exon skipping isoform. On the other hand, transduction of MEF cells with LV-U1ATR did not change the splicing profile of mouse ATR, indicating the U1 specificity toward the human 5' ss (Fig VI.3A).

The U1-mediated rescue of ATR expression was assessed even at protein level. While detection of ATR protein in MEF$^{S/S}$ cells did not show any traces of protein in cell extract, transduction with LV-U1ATR resulted in traces of ATR protein upon high exposure of the gel. Conversely, no changes upon LV-U1ATR infection and reduced detection of ATR protein in MEF cell were detected, indicating respectively the U1-specific correction and the human specificity of the antibody used in the assay (Fig. VI.3C).

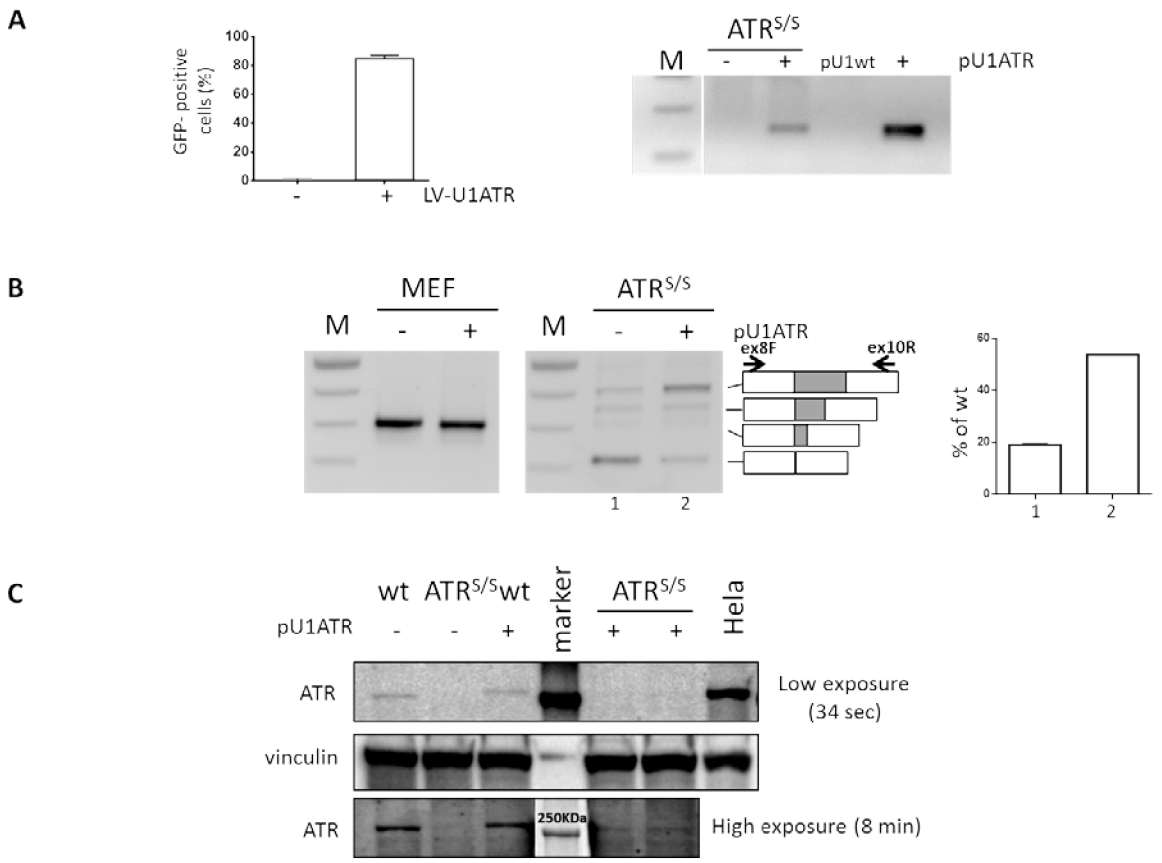


Figure VI. 3 Rescue of ATR expression in MEF$^{S/S}$ cells transduced with modified U1snRNA-expressing lentivirus.

A) Detection of modified U1snRNA expression in MEF$^{S/S}$ cells alone (-) or transduced (+) with a lentivirus expressing the U1ATRsnRNA targeting the exon 9 5'ss. Positive and negative control are represented by plasmids expressing the wt U1snRNA (pU1wt) or the modified U1snRNA (U1ATR) targeting the exon 9 5'ss used to create the lentivirus exploited to transduce MEF$^{S/S}$ cells. Amplified products were separated on 3% agarose gel. M, 100 bp molecular weight marker.

Histogram represents the GFP-positive MEF$^{S/S}$ cells upon infection with lentivirus expressing the modified U1snRNA and the GFP as tracking gene.

B) Evaluation of ATR alternative splicing patterns in MEF$^{S/S}$ cells (-) or transduced (+) with a lentivirus expressing the modified U1snRNA (U1ATR) targeting the exon 9 5'ss. The schematic representation of the normal and aberrant transcripts, and of primers used for the RT-PCR (arrows), is reported. Amplified products were separated on 2% agarose gel. M, 100 bp molecular weight marker. Histograms indicate the relative percentage of correctly spliced transcripts upon densitometric analysis of bands, which is expressed as mean ± standard deviation (SD).

C) Evaluation of ATR expression by western blotting analysis in MEF WT and MEF$^{S/S}$ cells alone (-) or transduced (+) with a lentivirus expressing the U1ATRsnRNA targeting the exon 9 5'ss, under low (upper panel) and high exposure time (lower panel). Vinculin expression and HeLa cells were used as control.

DISCUSSION

Pre-mRNA splicing is an essential process necessary for both proper gene expression and generation of transcript diversity. Intron removal and exon junction are directed by sequence signals within the pre-mRNA and are catalyzed by the spliceosome, a large ribonuclear protein complex with more than 100 players. The interaction of splicing regulatory sequences and their binding proteins (hnRNPs, SR proteins, and snRNPs) determines where and how splicing takes place. In this view, the spliceosome has to recognize exons by using sequences that are co-evolved with the amino acid sequence code and the splicing process selective pressure (Mueller et al. 2015).

While mutations within intronic sequences involved in spliceosome assembly are likely to alter splicing and are often associated to human diseases, missense mutations in exonic sequences can impair protein functionality, alter RNA processing, or both. On the other hand, synonymous mutations occurring within exons, not altering aminoacid sequence, can only impair protein expression through creation of regulatory defects. They have been shown to alter protein-folding abilities, RNA stability, and to induce exon skipping (Chamary et al. 2006, Sauna et al. 2007, Kimchi-Sarfaty et al. 2007). The synonymous change A2101G in the ATR gene, found in two unrelated families with Seckel syndrome, has been associated to altered exon 9 splicing. In particular, the mutation led mainly to exon skipping and the usage of two cryptic donor splice sites located upstream the mutation, both introducing a premature stop codon in the next exon. However, the mutation is associated to residual levels of correctly spliced transcript, as indicated by survival of patients. Albeit the aberrant splicing observed *in vitro* and *in vivo*, the molecular mechanism underlying the disease remained unknown. While some discrepancies, perhaps likely attributed to the cellular environment exploited, can be observed between our data and O'Driscool's findings using a similar exon trapping approach, it is clear that exon 9 is naturally poorly defined. In this context, we demonstrated that the A2101G mutation creates an Exonic Splicig Silencer (ESS), as indicated by the appearance of correctly processed transcript upon masking of the putative element using a specific ASO. The use of ASOs targeting the ESS might provide a potential therapeutic approach to rescue ATR expression in the Seckel disease. It is worth nothing that ASOs are generally used to induce exon skipping by masking regulatory elements favoring exon definition, while here we are forcing exon definition by masking a negative exon inclusion element.

The observation that the ESS effect was marginal or null in presence of ameliorated 3' ss or 5' ss, respectively, suggests that exon definition is more crucial for exon 9 inclusion than the presence of the ESS element. Exon definition is regulated by the recognition of the 5' and 3' ss from ribonucleoprotein U1snRNP and U2AF35, respectively. The donor splice site includes an almost invariant sequence GU at the 5' end of the intron, within a larger, less conserved region of 9 nucleotides (the consensus sequence is MAG|GTRAGT where M is A or C and R is A or G; | denotes exon-intron boundary) recognized by the 5' tail of the U1snRNA in the early stages of spliceosome assembly. Generally, donor splice sites differing from the consensus sequence are poorly recognized by the U1snRNA, thus requiring accessory regulatory elements within exon to ensure proper exon inclusion.

The ability of modified U1snRNA, in which the 5' tail is engineered to recover or improve base pair with the target donor splice site, in restoring exon definition is a well-proven strategy to rescue splicing impaired by splicing mutations located within the donor splice site (Fernandez et al. 2012). With the aim of improving exon 9 definition, we created a modified U1snRNA (U1ATR) with perfect complementarity with the exon 9 5' ss. The co-expression of the U1ATR, despite the poorly defined exon 9 context, was able to remarkably recover exon inclusion in the ATRwt and, most importantly, in the presence of the mutation, which clearly demonstrates its efficacy in forcing exon definition. On the other hand, co-expression of the U1wt was ineffective, thus highlighting the crucial role of the modified U1snRNA 5' s tail in defining target recognition and splicing correction.

In the attempt to restore splicing, we combined the U1snRNA and ASO-mediated approaches. As demonstrated by a similar approach for coagulation factor IX (Balestra et al. 2015), the tailored approach has an additive effect in rescuing splicing when compared to the U1ATR-based strategy, indicating the combination of improved exon definition and ESS masking can efficiently restore exon 9 inclusion.

A mandatory concern to evaluate the therapeutic potential of a correction approach is to assess the impact on protein levels. It is worth nothing that traces of ATR protein were detectable in MEF^{ATR} cells, a finding supporting the survival of patients. Notably, in humanized MEF cell infected with a lentivirus expressing the modified U1, we observed an increase of correct transcripts. The divergence between the correction observed in minigene system and MEF cell relies on the low expression level of the modified U1 gene delivered by lentivirus vector rather than on the low transduction efficiency, as demonstrated by the high GFP-positive cells observed upon infection. In fact, the presence of CpG islands in the transgene is the main limiting factor in the level and duration of

transgene expression delivered by lentivirus vectors (Yew et al. 2002). Accordingly, the presence of CpG islands spanning all U1snRNA-expressing gene was confirmed by bioinformatics analysis (data not shown). On the other hand, it has been recently demonstrated (Dal Mas et al. 2015) that a single modified U1 is able to efficiently rescue splicing due to its endogenous high expression level. Despite the modest correction observed at RNA level in MEF cells, we were able to detect an increase of the full-length ATR protein in cells treated with U1-expressing lentivirus.

Taken together these findings provide the molecular mechanism underlying the Seckel syndrome caused by the synonymous change A2101G and the proof that a therapeutic approach based on ASO, modified U1snRNA or their combination is feasible.

Materials and methods

Creation of vectors:

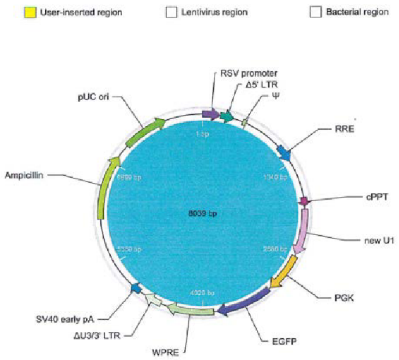
A 1561-bp genomic fragment spanning *ATR* intron 8 (-434bp) trough intron 10 (+416bp) was amplified from the genomic DNA of a normal subject using high-fidelity PfuI DNA-Polymerase (Transgenomic, Glasgow, UK) and cloned in the pTB expression vector using the unique *NdeI* restriction site (pATRwt). Mutants (pATR2101G, **pATR**2101**G**$^{5'ss+}$, **pATR**2101**G**$^{3'ss+}$) of this minigene were generated by site-directed mutagenesis (QuickChange II XL Site directed Mutagenesis Kit; Agilent Technologies). All ATR minigenes were checked by direct sequencing. For the construction of the 5'ss fully complementary U1^{ATR} expression plasmid, the BglII-BclI fragment of the pU1wt was replaced using the following oligonucleotides (5'-gatctgatACATACATAgcaggggagataccat-3' and 5'-gatcatggtatctcccctgcTATGTATGTatca-3') as previously described (Pinotti et al. 2008).

Antisense oligonucleotides (ASO)

ASO base pairing with nucleotides 32 of exon 9 was synthetized by Dr. Daniela Perrone (Department of Chemical and pharmaceutical Sciences, University of Ferrara, Ferrara, Italy) and contains 2'-0-metyl modified ribonucleotides and full-length phosphorotioate backbone (5'-TAAGATAAAAAACCCACTAACACAA-3').

Creation of Lentivirus and transduction of immortalized MEF cells.

For the creation of modified U1-expressing lentivirus vector, the sequence of U1snRNA was cloned in a lentiviral plasmid (pLV- PGK-EGFP, Cyagen Biosciences) using the *BamHI* site.



***Scheme** of LV-U1ATR*

The U1 lentiviral plasmid together with psPAX2 packaging plasmid and pMD2.G envelope plasmid were combined at a ratio of 3:2:1 respectively, and then used to transfect 293T packaging cells using calcium phosphate method. Twenty-four hours post transfection the medium was replaced with a new medium. The virus supernatants were collected at 48 and 72 hours post transfection and filtered through a 0.45 µm filter.

Immortalized wild type and seckel mouse embryonic fibroblast (MEF) cells were incubated with the filtered virus supernatant supplemented with polybrene (8 µg/ml) at least for 3 hours at 37°C in a humidified incubator. Afterwards the medium was replaced by fresh medium.

Expression in mammalian cells and mRNA studies

Hek293 cells were cultured in a humidified incubator at 37°C and 5% CO_2 in Dulbecco modified Eagle medium (DMEM, Euroclone) supplemented with 2mM L-glutamine, 10% fetal bovine serum, and Antibiotics (Sigma) according to the manufacturer's instructions. For each transfection experiment, cell were seeded in a 12-well plate at 80% confluency. After 24 hours, cells were switched to serum-free OptiMEM (Life technologies) and transiently transfected with Lipofectamine 2000 (Life technologies) according to the manufacturer's protocol. 1.5 ug of minigenes were transfected either alone, with ASO or with 1.5X molar excess of U1snRNA-encoding plasmids. Total RNA extraction was performed after 24h of incubation using TRIreagent (Invitrogen) and reverse transcripted using the M-MLV (Invitrogen) following the manufacturer's protocol. α-2,3 (5'-XX-3')

and 10R (5'-XX-3') primers were used to assess splicing process. The conditions used for the PCRs were 95°C for 2 minutes for the initial denaturation, 95°C for 30 s, 56 °C for 30 s, 72°C for 40 s for 40 cycles and 72°C for 10 min for the final extension.
Densitometric analysis was performed using ImageJ software.

Western blot analysis

Cells were washed twice with ice-cold PBS and then directly lysed in an appropriate volume of 2X laemmli buffer (120 mM Tris-HCl pH 6.8, 4% SDS and 20% Glycerol) supplemented with protease and phosphatase inhibitors. After sonication, the lysate was cleared by centrifugation and the supernatant was recovered. The protein concentration was quantified using optical absorbance at 280 nm. 25 µg of protein was boiled with β-mercaptoethanol (5% v/v) at 95°C for 10 minutes. Samples were then separated on 4-15% gradient precast TGXTM polyacrylamide gel (Bio-Rad Laboratories) and then were blotted onto nitrocellulose membrane using Trans-Blot® TurboTM transfer system (25 V, 1 A, 30 minutes). Following 1 hour blocking in TBST containing 5% BSA, the blot was incubated with primary antibody against ATR (N-19, Santa Cruz, 1:1000 dilution) for overnight at 4°C. Secondary peroxidase-coupled antibody (rabbit anti-goat, Sigma, 1:5000 dilution) was incubated with the blot at room temperature for 1 hour. ECL-based chemiluminescence (Bio-Rad Laboratories) was detected on BioRad ChemiDoc system and the image was processed using Image Lab 4.0 (Bio-Rad Laboratories).

Computational analysis

Bioinformatic prediction of splice sites and/or splicing regulatory elements was conducted by using the http://fruitfly.org/seq_tools/splice.html and the Human Splicing Finder (http://www.umd.be/HSF3/) online softwares.

REFERENCES

-Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M. An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice. J Thromb Haemost. 2014 Feb;12(2):177-185. doi: 10.1111/jth.12471.

-Balestra D, Barbon E, Scalet D, Cavallari N, Perrone D, Zanibellato S, Bernardi F, Pinotti M. Regulation of a strong F9 cryptic 5'ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides. Hum Mol Genet. 2015 Sep 1;24(17):4809-16. doi: 10.1093/hmg/ddv205. Epub 2015 Jun 10

-Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, Ffrench-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 Mar;40(3):220-2.

-Chamary, J. V., Parmley, J. L., and Hurst, L. D. (2006) Hearing silence:non-neutral evolution at synonymous sites in mammals. Nat. Rev. Genet.7, 98–108

-Dal Mas A, Fortugno P, Donadon I, Levati L, Castiglia D, Pagani F. Exon-Specific U1s Correct SPINK5 Exon 11 Skipping Caused by a Synonymous Substitution that Affects a Bifunctional Splicing Regulatory Element. Hum Mutat. 2015 May;36(5):504-12. doi: 10.1002/humu.22762. Epub 2015 Mar 19.

-Durocher D, Jackson SP. DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? Curr Opin Cell Biol. 2001 Apr;13(2):225-31. Review.

-Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. Hum Mol Genet. 2012 Jun 1;21(11):2389-98. doi: 10.1093/hmg/dds045. Epub 2012 Feb 23.

-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 May;19(5):936-41. doi: 10.1038/mt.2011.7. Epub 2011 Feb 15.

-Kimchi-Sarfaty, C., Oh, J. M., Kim, I. W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V., and Gottesman, M. M. (2007) A “silent” polymorphism in the MDR1 gene changes substrate specificity. Science 315, 525–528

-Mueller WF, Larsen LS, Garibaldi A, Hatfield GW, Hertel KJ. The Silent Sway of Splicing by Synonymous Substitutions. J Biol Chem. 2015 Nov 13;290(46):27700-11. doi: 10.1074/jbc.M115.684035. Epub 2015 Sep 30.

-Murga M, Bunting S, Montaña MF, Soria R, Mulero F, Cañamero M, Lee Y, McKinnon PJ, Nussenzweig A, Fernandez-Capetillo O. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. Nat Genet. 2009 Aug;41(8):891-8. doi: 10.1038/ng.420. Epub 2009 Jul 20.

-O'Driscoll M1, Ruiz-Perez VL, Woods CG, Jeggo PA, Goodship JA. A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet. 2003 Apr;33(4):497-501. Epub 2003 Mar 17.

-Pagani F, Stuani C, Tzetis M, Kanavakis E, Efthymiadou A, Doudounakis S, Casals T, Baralle FE. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum Mol Genet. 2003 May 15;12(10):1111-20.

-Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet. 2004 May;5(5):389-96. Review.

-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 Jun 18;113(25):6461-4. doi: 10.1182/blood-2009-03-207613. Epub 2009 Apr 22.

-Pinotti M, Bernardi F, Dal Mas A, Pagani F. RNA-based therapeutic approaches for coagulation factor deficiencies. J Thromb Haemost. 2011 Nov;9(11):2143-52.doi: 10.1111/j.1538-7836.2011.04481.x. Review.

-Ragland RL, Arlt MF, Hughes ED, Saunders TL, Glover TW. Mice hypomorphic for Atr have increased DNA damage and abnormal checkpoint response. Mamm Genome. 2009 Jun;20(6):375-85. doi: 10.1007/s00335-009-9195-4. Epub 2009 Jun 7.

-Sauna, Z. E., Kimchi-Sarfaty, C., Ambudkar, S. V., and Gottesman, M. M.(2007) The sounds of silence: synonymous mutations affect function.Pharmacogenomics 8, 527–532.

-Shiloh Y. ATM and ATR: networking cellular responses to DNA damage. Curr Opin Genet Dev. 2001 Feb;11(1):71-7.

-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 Sep;24(3):225-35.

-Yew NS, Zhao H, Przybylska M, Wu IH, Tousignant JD, Scheule RK, Cheng SH. CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo. Mol Ther. 2002 Jun;5(6):731-8.

-Zhou BB, Elledge SJ. The DNA damage response: putting checkpoints in perspective. Nature. 2000 Nov 23;408(6811):433-9. Review.

VII

Conclusions and Perspectives

The central dogma of Molecular Biology emerged originally as a collinear view of gene expression, in which the information flows from DNA to protein through messenger RNA molecules. The last decades of research have considerably expanded this paradigm by showing the multiplicity of transcripts that can be generated from a single DNA locus through alternative promoters, termination sites, and throught alternative splicing of exons and/or introns, with the latter that need to be removed to generate translatable mRNAs.

Removal of introns form pre-mRNA precursors (pre-mRNA splicing) is a necessary step for the expression of most genes in multicellular organisms, and about 90% of human genes undergo to alternative splicing, thus expanding the number of proteins produced by the relatively low number of genes in human genome. Mutations or natural variations in pre-mRNA sequences, as well as in spliceosomal components and regulatory factors, have been implicated in the etiology and progression of numerous pathologies. These ranges from monogenic to multifactorial genetic diseases, including metabolic syndromes, muscular dystrophies, neurodegenerative and cardiovascular diseases, and cancer. Understanding the molecular mechanisms associated with splicing-related pathologies can provide key insights into the physiological function of the complex splicing machinery and thus establish sound basis for novel therapeutic approaches.

The complexity of the splicing code offers many chances for mutations to disrupt normal pre-mRNA splicing. Mutations can destroy functional donor and acceptor splice sites, activate cryptic splice sites in exons or introns and weaken/strengthen auxiliary splicing signals (exonic and intronic splicing enhancers and silencers), thereby leading to aberrant splicing events such as exon skipping or intron retention in the mature mRNA. Moreover, since U1snRNA is involved even in poly A sites selection (Berg et al. 2012), mutations can be associated to aberrant splicing due to the ability of the mutant allele to recruit proteins that can interfere with the U1-mediated inhibition of cryptic poly A signals within introns. In fact, in Huntington disease caused by duplication of CAG triplets in exon 1, a proposed

mechanism of aberrant splicing relies on the ability of the SRSF6, a splicing factor that binds the expanded CAG repeats, to interfere with the ability of U1snRNA to inhibit premature poly A signals located within intron 1 of HTT gene (Gipson TA et al. 2013). In addition to mutations occurring within sequences involved in the spliceosome assembly, the real contribution of missense and synonymous changes within exons in splicing alteration remains largely underestimated. Overall, it has been estimated that mutations altering splicing process may account for up to 50% of all pathogenic DNA variants.

Nowadays, the most exploited therapeutic strategy for rare inherited diseases is the replacement gene therapy. Despite the great potential of this approach, and the results obtained in some diseases, (Cicalese et al. 2015; Gaudet et al. 2015; Kolb and Kissel 2015; George and Fogarty 2016) the size of the gene and the gene regulation limit the applicability of such approach. In fact, gene capacity of the most used viral vectors (AAV and lentivirus) is generally restricted to ~5 kbp, thus preventing its translability for diseases caused by mutations in large genes, such as dystrophin and *F8* genes, associated respectively to Duchenne Muscular Dystrophy and Hemophilia A. Moreover, metabolic diseases for which the gene expression has to be fine regulated, simple replacement gene therapy approaches cannot be easily applied.

These limitations boosted research toward alternative strategies based on intervention at RNA level. This approach is emerging as a promising alternative therapeutic approach for genetic diseases (Bonetta L. 2009; Hammond SM et al. 2011) as it permits restoration of gene expression while maintaining gene regulation in physiological tissues, and overcomes limitations related to vector-mediated delivery of large genes.

Increasing attention has been given to the U1 small nuclear RNA (U1snRNA) that, in the earliest splicing step, mediates the recognition of the donor splice site (5'ss) by the ribonucleoprotein U1snRNP (Horowitz DS, Krainer AR. 1994). Studies in various cellular models of human disease indicated the potential therapeutic effect of engineered

U1snRNAs to rescue aberrant splicing caused by mutations at 5'ss (Baralle et al. 2003; Susani et al. 2004; Tanner et al. 2009; Hartmann et al. 2010; Schmid 2011; Glaus et al. 2011; Dal Mas et al. 2015a).

In the last years, our group has successfully exploited U1snRNAs to correct splicing mutations causing severe coagulation factor deficiencies. We demonstrated for coagulation factor VII (FVII) deficiency caused by a 5'ss mutation that an U1snRNA, designed to restore complementarity to the mutated 5'ss, rescues FVII mRNA processing and secretion of pro-coagulant molecules *in vitro* and *in vivo* (Pinotti et al. 2008; Pinotti et al. 2009; Balestra et al. 2014). The approach has been also extended to several splicing mutations inducing skipping of exon 5 in F9 gene. In cellular models, we demonstrated that a unique U1snRNAs targeting intronic sequences downstream of the 5'ss (Exon-Specific U1snRNA, ExSpeU1) restores in vitro FIX function impaired by different mutations at the 5'ss but also at the acceptor (3'ss) splice site (Fernandez et al. 2012). Moreover, the ExSpeU1s have been successfully exploited to rescue SMN2 splicing and SMN expression in mouse models of Spinal Muscular Atrophy (Dal Mas et al. 2015b). In spite of this promising potential, the ExSpeU1 ability to rescue multiple mutations has not been explored *in vivo*.

Therefore, one of the main aims of the thesis was (chapter IV) to explore in mice the ExSpeU1-mediated rescue of two model Hemophilia B-causing mutations at the donor (c.519A>G) or acceptor (c.392-8T>G) splice sites of *F9* exon 5 and to assess the correction at RNA level and the pro-coagulant activity of the rescued protein. To overcome the absence of HB mouse models for specific splicing mutations we exploited the hydrodynamic injection of splicing defective human FIX cassettes that enable us to express them mainly in liver of wt BALB/C mice. Treatment with the plasmids expressing the wt and mutant (hFIX-2G$^{5'ss}$ and hFIX-8G$^{3'ss}$) constructs resulted in liver in hFIX transcripts lacking exon 5, and in low plasma levels of the inactive FIX deleted isoform, which

mimicked the coagulation phenotype of HB patients. Interestingly, co-injection of the selected ExSpeU1 U1fix9, but not of U1wt, restored exon inclusion of both variants and in the intrinsically weak FIXwt context. This resulted in appreciable circulating levels of pro-coagulant full-length hFIX (mean ± standard deviation, 1.0±0.5 µg/ml; 1.2±0.3 µg/ml; 1.9±0.6 µg/ml,), leading to a striking shortening (15-20 s) of FIX-dependent coagulation times. It is worth noting that the extent of rescue, if translated into HB patients, would be well beyond the therapeutic threshold that is around 2-5% (Pollak and High 2003).

Altogether, these data provide the first proof-of-concept *in vivo* that a unique ExSpeU1 can efficiently rescue gene expression impaired by distinct exon-skipping variants, which extends the applicability of ExSpeU1s to panels of mutations and thus cohort of patients, a major achievement to propose this strategy for diseases with highly heterogeneous mutational patterns.

On the other hand, mutations at 5'ss might lead to the activation of cryptic splice sites rather than inducing exon-skipping, a mechanism that might vanish the strategy based on engineered U1snRNAs. For this reason, we characterized a meaningful splicing model in which numerous HB-causing mutations, either missense or at the donor splice site (5'ss) of coagulation F9 exon 2, promote aberrant splicing by inducing the usage of a strong exonic cryptic 5'ss. Interestingly, splicing assays with natural and artificial *F9* variants indicated that the cryptic 5'ss is regulated, among a network of regulatory elements, by an exonic splicing silencer (ESS). This finding, and the comparative analysis of the *F9* sequence across species showing that the cryptic 5'ss is always paralleled by the conserved ESS, supporting a compensatory mechanism aimed at minimizing unproductive splicing.

In this context with a strong cryptic 5'ss and a defective authentic 5'ss, we observed a very poor correction effect by U1snRNA variants with increased or perfect complementarity to the defective 5'ss. This led us to test another extensively exploited RNA-based technology, namely antisense oligoribonucleotides (AONs), exploited to mask the cryptic 5'ss.

However, the AONs were effective on exonic changes but only promoted exon 2 skipping in the presence of mutations at the authentic 5'ss. For this reason, we explored the combination of the mutant-specific U1snRNAs with AONs, which resulted in appreciable amounts of correctly spliced transcripts (from 0% to 20-40%) for several mutants of the exon 2 5'ss.

Taken together these findings, which revealed a fine interplay among ESS, cryptic and the authentic 5'ss as a disease-causing mechanism, provided novel experimental insights into the combinatorial correction activity of antisense molecules and compensatory U1snRNAs.

Another poorly addressed issue regards the pathogenic splicing caused by synonymous and non-synonymous exonic changes at splicing regulatory elements (Cartegni et al. 2002, Cartegni et al. 2003), which represent additional therapeutic targets for splicing- switching molecules such as modified U1snRNAs or AONs. A paradigmatic example is provided by the synonymous mutation in SMN2 exon 7 that produces exon skipping, and which is a therapeutic target for SMA. As a matter of fact, SMN2 splicing has been successfully rescued both *in vitro* and *in vivo* in SMA mouse models by either ExSpeU1s (Dal Mas et al. 2015b) or AONs targeting an ISS (Hua et al. 2007; Hua et al. 2008; Passini et al. 2011).

In this thesis, we addressed the issue on the model of the *ATR* c.2101A>G mutation associated with Seckel syndrome-1. Expression studies demonstrated that the variant creates an exonic splicing silencer (ESS), leading to exon 9 skipping. Different strategies were explored to rescue *ATR* splicing: i) AON masking mutation and ESS ii) modified U1snRNA with increased complementarity to the 5'ss iii) combination of AON and U1snRNA. Both approaches produced appreciable amounts of correctly spliced transcripts (from 0% to 42 and 49%, respectively) in cellular models. Noticeably, combination of modified U1snRNAs with AONs led to a significant improvement of *ATR* rescue (~69%) at RNA level. Infection of SS-1 MEF cells with lentiviral vector harboring the modified

U1snRNA, although to a less extent, rescued *ATR* expression either at RNA and protein levels.

Altogether, these data demonstrated that in the presence of poorly defined exons and of exonic mutations creating splicing regulatory elements, the approach with compensatory U1snRNAs and AONs produces additive effects.

In conclusion, by investigations in *F9* and *ATR* genes, we demonstrated that the knowledge of the aberrant splicing mechanisms leading to diseases offers the intriguing opportunity to design and develop tailored correction approaches based on modified U1snRNA, improving exon definition, or antisense oligonucleotides, masking cryptic sites or regulatory elements, able to rescue splicing and gene expression. Depending on the exon units and on the mechanism, these two strategies can be combined with synergistic effects.

Regarding the future translatability, the two strategy are complementary.

The small size of the gene coding the U1snRNA makes it easily packageable into AdenoAssociated virus vectors (AAV), which actually represent the ideal vectors for gene delivery purposes. They are not pathogenic, not-integrative and, due to the many serotypes available, can be exploited to target different tissues of interest. In this view, the AAV-U1snRNA injection would permit a life-long therapeutic approach upon a single injection.

On the other hand, the AON are easy to produce and could be administered several times, and could be used in those patients not eligible for AAV injections (pre-existing anti-AAV antibodies). Not surprisingly, both strategies are currently under assessment in clinical trials for SMA.

Altogether our data, in combination with results obtained in other human genetic disorders, propose RNA-based approaches for further investigations in mouse models of disease aimed at assessing the therapeutic effect and the safety profile (including the assessment of off-targets by RNA sequencing), which might lead to the development of further

innovative therapies for the numerous human inherited diseases associated with aberrant splicing.

REFERENCES

-Balestra D, Faella A, Margaritis P, Cavallari N, Pagani F, Bernardi F, Arruda VR, Pinotti M. An engineered U1 small nuclear RNA rescues splicing-defective coagulation F7 gene expression in mice. J Thromb Haemost. 2014 Feb;12(2):177-185. doi: 10.1111/jth.12471.

-Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, Ffrench-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 Mar;40(3):220-2.

-Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, Dreyfuss G. U1 snRNP determines mRNA length and regulates isoform expression. Cell. 2012 Jul 6;150(1):53-64. doi:10.1016/j.cell.2012.05.029.

-Cartegni L, Krainer AR. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat Genet. 2002 Apr;30(4):377-84. Epub 2002 Mar 4.

-Cartegni L, Krainer AR. Correction of disease-associated exon skipping by synthetic exon-specific activators. Nat Struct Biol. 2003 Feb;10(2):120-5.

-Cicalese MP, Aiuti A. Clinical applications of gene therapy for primary immunodeficiencies. Hum Gene Ther. 2015 Apr;26(4):210-9. doi:10.1089/hum.2015.047. Review. PubMed PMID: 25860576; PubMed Central PMCID:PMC4410188.

-Dal Mas A, Fortugno P, Donadon I, Levati L, Castiglia D, Pagani F. Exon-Specific U1s Correct SPINK5 Exon 11 Skipping Caused by a Synonymous Substitution that Affects a Bifunctional Splicing Regulatory Element. Hum Mutat. 2015a May;36(5):504-12. doi: 10.1002/humu.22762. Epub 2015 Mar 19.

-Dal Mas A, Rogalska ME, Bussani E, Pagani F. Improvement of SMN2 pre-mRNA processing mediated by exon-specific U1 small nuclear RNA. Am J Hum Genet. 2015b Jan 8;96(1):93-103. doi: 10.1016/j.ajhg.2014.12.009. Epub 2014 Dec 31.

-Fernandez Alanis E, Pinotti M, Dal Mas A, Balestra D, Cavallari N, Rogalska ME, Bernardi F, Pagani F. An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. Hum Mol Genet. 2012 Jun 1;21(11):2389-98. doi:10.1093/hmg/dds045. Epub 2012 Feb 23.

-Gaudet D, Brisson D. Gene-based therapies in lipidology: current status and future challenges. Curr Opin Lipidol. 2015 Dec;26(6):553-65. doi:10.1097/MOL.0000000000000240.

-George LA, Fogarty PF. Gene therapy for hemophilia: past, present and future. Semin Hematol. 2016 Jan;53(1):46-54. doi: 10.1053/j.seminhematol.2015.10.002. Epub 2015 Oct 28. Review.

-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 May;19(5):936-41. doi: 10.1038/mt.2011.7. Epub 2011 Feb 15.

-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 Oct 8;87(4):480-93. doi: 10.1016/j.ajhg.2010.08.016.

-Hua Y, Vickers TA, Baker BF, Bennett CF, Krainer AR. Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon. PLoS Biol. 2007 Apr;5(4):e73.

-Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. Am J Hum Genet. 2008 Apr;82(4):834-48. doi: 10.1016/j.ajhg.2008.01.014. Epub 2008 Mar 27.

-Kolb SJ, Kissel JT. Spinal Muscular Atrophy. Neurol Clin. 2015 Nov;33(4):831-46. doi: 10.1016/j.ncl.2015.07.004. Review. PubMed PMID: 26515624.

-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 Mar 2;3(72):72ra18. doi:10.1126/scitranslmed.3001777.

-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 Mar 1;111(5):2681-4. Epub 2007 Dec 21.

-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 Jun 18;113(25):6461-4. doi: 10.1182/blood-2009-03-207613. Epub 2009 Apr 22.

-Pollak, E and High, K.A. (2003) Genetic disorders of coagulation. In: Warrell, D., Cox, T., Firth, J. and Benz, E. (Eds) Oxford Textbook of Medicine, 4th ed, pp. 757–767, Oxford University Press, Oxford. Vol. 3

-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 Sep;24(3):225-35.

-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 Feb;30(2):255-63. doi:10.1002/humu.20861. Erratum in: Hum Mutat. 2009 Feb;30(2):274.

Curriculum vitae

Daniela Scalet was born on 13th September 1988 in Feltre (Italy). After completing her secondary education at the Scientific High School "Vittorino da Feltre" in Feltre, she enlisted at Ferrara University where she obtained a Bachelor's degree in Biotechnology (2010) and a Master's degree in Cellular and Molecular Biology (2012) *cum laud.* Her Master thesis, entitled " Splicing mutations in F9 coagulation gene: characterizations and corrections with *small nuclear RNAs*" was based on a 9 months practical internship at the Department of Life Science and Biotechnologies at the University of Ferrara under the supervision of Prof. Mirko Pinotti. In January 2013 she started a PhD project on regulation of exon definition by intrinsec elements and by combination of tailored U1snRNA with antisense oligonucleotides in *ATR*, *FAH* and *F9* coagulation factor genes at the Department of Life Science and Biotechnologies at the University of Ferrara under the supervision of Prof. Mirko Pinotti and Prof. Francesco Bernardi. In the frame of these projects she also moved abroad like "visiting researcher" for two months to the laboratory of Dr. Stan Van de Graaf at Tytgat Institute for Liver and Intestinal Research, Department of Gastroenterology & Hepatology at the Academic Medical Center (Amsterdam) to study an RNA-based therapeutic strategy for FAH deficiency. As a PhD Student she attended several courses on thrombosis and haemostasis and rare diseases. Morover she presented her work at national and international congress, receiving a "Young Investigator Award" at the XXV Congress of the International Society on Thrombosis and Haemostasis (Toronto 2015) and "Ulla Hedner Award", Access to Insight Core Faculty and Novo Nordisk (Toronto 2015).

List of publications

1. Balestra D, Barbon E, **Scalet D***, Cavallari N, Perrone D, Zanibellato S, Bernardi F, Pinotti M. Regulation of a strong F9 cryptic 5'ss by intrinsic elements and by combination of tailored U1snRNAs with antisense oligonucleotides. Hum Mol Genet.2015 Jun10. ***First author**

2. Balestra D, **Scalet D**, Pagani F, Mari R, Bernardi F, Pinotti M. A single Exon-Specific U1snRNA induces in mice a robust FIX activity impaired by mutations at donor or acceptor splice sites. Mol Ther Nucleic Acids 2015.*Under Revision*.

3. **Scalet D**, Balestra D, Rohban S, Bovolenta M, Campaner S and Pinotti M. Characterization and correction of the ATR A2101G synonymous change causing Seckel syndrome. *In preparation*

Congress abstracts

1. **Scalet D**, Balestra D, Barbon E, Cavallari N, Perrone D, Bernardi F, Pinotti M. (ISTH congress 2015). Correction of aberrant splicing causing Haemophilia B through the combination of compensatory U1snRNAs and antisense oligonucleotides.

2. Balestra D, **Scalet D**, Pagani F, Bernardi F, Pinotti M. . (ISTH congress 2015). A unique Exon Specific U1snRNA rescues different Haemophilia B –causing splicing-defective factor IX variants in mice.

3. Balestra D, Dal Mas A, Rogalska Malgorzata E, Barbon E, **Scalet D**, Donadon I, Ferrarese M, Bussani E, Pianigiani G, Ferraresi P, Branchini A, Baroni M, Bovolenta M, Mattioli C, Pagani F, Pinotti M. (XVIII Convention Telethon 2015). Development of a RNA-based therapeutic approaches for hemophilia B caused by exon skipping mutations.

4. Balestra D, Cavallari N, Barbon E, **Scalet D**, Fernandez Alanis E, Dal Mas A, Rogalska Malgorzata E, Pagani F, Bernardi F and Pinotti M. 57th National Meeting of the Italian Society of Biochemistry and Molecular Biology (2013). Aberrant mRNA splicing in coagulation factor deficiencies: from molecular mechanisms to RNA-based therapeutic approaches.

Awards

Novo Nordisk

Ulla Hedner Award, Access to Insight Core Faculty and Novo Nordisk. “Correction of aberrant splicing causing haemophilia B through the combination of compensatory U1snRNAs and antisense oligonucleotides”, Toronto 2015

Young Investigator Award

XXV Congress of the International Society on Thrombosis and Haemostasis (ISTH), Toronto 2015