UNIVERSITY OF NOVA GORICA

GRADUATE SCHOOL

DEFECTIVE EXON RECOGNITION IN DISEASE-ASSOCIATED GENES AND SPLICING CORRECTION STRATEGIES USING MODIFIED U1 snRNAS

DISSERTATION

Eugenio Fernandez Alanis

Mentor: Dr. Franco Pagani, M.D.

Nova Gorica, 2011

“L´amour est immense, mais il n´est pas infini; tandis que la science a des profondeurs sans limites où je ne saurais te voir aller seul. Je déteste tout ce qui peut se mettre entre nous. Si tu obtenais la gloire après laquelle tu cours, j´en serais malhereuse; ne te donnerait-elle pas des vives jouissances?”

Honoré de Balzac, *La Recherche de l´Absolu*

-¿Es usted un científico?- me preguntó él en tono descorazonado.
-El señor Farías es un hombre de letras- le aclaró Bermúdez.
-Menos mal –sentenció Frobenius. Y añadió, en un arranque de ira-: la Ciencia carece de todo “valor explicativo”. ¡No revela un corno!
La Enviada le tomó una mano y le dijo:
-¡Acuérdese, doctor!
-Urania –repuso él-, no estoy agitado. Pero una ciencia que arranca de la duda y se dirige a la duda por entre la duda, es un infierno, y no muy caro. Una verdad que no sea “indudable” no es una verdad. ¿Qué hacemos nosotros? ¿Me dirás que Ciencia? No, señor: nosotros hacemos algo así como el turismo de la duda.

Leopoldo Marechal, *El Banquete de Severo Arcángelo*

Abstract

A significant proportion of disease-causing mutations affect pre-mRNA splicing inducing skipping of the exon from the mature transcript. Using *F9* exon 5 and *CFTR* exon 12 models, I studied the effect of several natural mutations on splicing in haemophilia B and cystic fibrosis, respectively, and the role of U1 small nuclear RNA (snRNA) in the resulting splicing defects. Using hybrid minigene systems, I established the pathological effect of eighteen mutations. Sixteen were associated to exon skipping, one activated a cryptic 5' ss and one did not affect splicing. Unexpectedly, *in silico* analysis did not predict the effect on splicing of two 5' ss mutations (+3G and +4G in *CFTR* exon 12 and *F9* exon 5, respectively) indicating the importance of functional splicing assay for the diagnostic assessment of human genomic variations.

In minigene systems, loading of U1 snRNA by complementarity to the normal or mutant 5' ss corrected exon skipping caused by different types of mutations. To improve specificity and reduce potential off-target effects, I engineered U1 snRNA variants targeting nonconserved intronic sequences downstream of the 5' ss. For each gene system, I identified an exon-specific U1 snRNA (ExSpeU1) able to rescue splicing impaired by mutations at the polypyrimidine tract of the acceptor splice site, at the consensus 5' ss or at exonic regulatory elements. I characterised, by deletion analysis, two weak intronic splicing silencers (ISS) located in non conserved intronic sequences downstream of the 5' ss in *CFTR* exon 12 and *F9* exon 5, Interestingly these ISSs partially overlap with the binding sites of ExSpeU1s. By further site-directed mutagenesis and RNA binding experiments, I showed that, while *F9* exon 5 ISS binds to hnRNP A1/A2, *CFTR* exon 12 ISS is possibly forming a regulatory RNA secondary structure. Thus, part of the ExSpeU1-mediated splicing enhancement might be

due to interference with the corresponding ISS sequence.

Through splicing-competent cDNA constructs, I demonstrated that the ExSpeU1-mediated splicing correction of several *F9* exon 5 mutations at the polypyrimidine tract of the acceptor splice site and at the consensus 5' ss results in complete restoration of secreted functional factor IX levels. Thus, ExSpeU1s constitute a promising novel therapeutic strategy to correct splicing defects associated to defective exon definition in several human disorders.

Povzetek

Pomemben del mutacij, povezanih z nastankom bolezni, prizadene povezovanje (splicing) na nivoju pre-mRNA. To se najpogosteje zgodi z izključitvijo eksona iz zrelega transkripta. V doktorskem delu sem z modeloma *F9* eksona 5 and *CFTR* eksona 12 proučeval vpliv večih naravnih mutacij, ki prizadenejo povezovanje, pri hemofiliji B in cistični fibrozi. Poleg tega sem ugotavljal vlogo male jedrne RNA (snRNA) U1 v nastalih nepravilnostih. Z uporabo hibridnega minigenskega sistema sem ponazoril patološki vpliv osemnajstih mutacij. Šestnajst jih je bilo povezanih z izključitvijo eksona, ena s kriptičnim 5' ss, medtem ko ena ni prizadela povezovanja. Proti pričakovanju številni programi za napovedovanje povezovanja niso pravilno napovedali dveh mutaciji 5' ss (+3G v *CFTR* eksona 12 in +4G v *F9* eksona 5), kar poudarja pomembnost funkcijskega testiranja z ustreznimi testi pri diagnostični oceni različnih humanih genomskih variacij.

V minigenskem sistemu je dodajanje snRNA U1 skupaj z normalno ali mutirano 5' ss odpravilo izključevanje eksonov, ki nastane zaradi različnih vrst mutacij. Da bi izboljšali specifičnost in zmanjšali potencialni izventarčni vpliv, sem konstruiral serijo variant snRNA U1, ki so ciljale neohranjena intronska zaporedja navzdol od 5' ss. Za vsak genski sistem sem identificiral eksonsko-specifično snRNA U1 (ExSpeU1), ki je bila sposobna vzpostaviti pravilno povezovanje, prej onemogočeno zaradi mutacij v polipirimidinskem odseku akceptorskega mesta za povezovanje, na konsenzusnem mestu 5' ss ali na eksonskih regulatornih elementih. Z delecijsko analizo sem okarakteriziral dva šibka intronska utiševalca povezovanja (intronic splicing silencers – ISS), ki se nahajata na neohranjenih intronskih zaporedjih navzdol od 5' ss v *CFTR* eksona 12 in *F9* eksona 5. Zanimivo je, da se ti ISS delno prekrivajo z vezavnimi

mesti ExSpeU1s. Z usmerjeno mutagenezo in s testom vezave RNA sem pokazal, da se ISS *F9* eksona 5 veže s hnRNP A1/A2, medtem ko ISS *CFTR* eksona 12 zelo verjetno tvori regulatorno sekundarno strukturo RNA. Ojačeno povezovanja, ki ga povzroči ExSpeU1s, je lahko vsaj delno tudi posledica interference z odgovarjajočimi zaporedji ISS.

S pomočjo cDNA konstruktov, ki so kodirali komponente povezovanja, sem pokazal, da lahko z ExSpeU1 popravimo povezovalne nepravilnosti številnih mutacij *F9* eksona 5 na polipirimidinskem odseku akceptorskega mesta za povezovanje in na konsenzusnem mestu 5' ss. To se odraža v popolni povrnitvi ustrezne ravni izločanja funkcionalnega dejavnika strjevanja krvi IX. Zaradi tega lahko ExSpeU1s predstavlja obetavno novo terapevtsko orodje za odpravljanje nepravilnosti pri povezovanju, ki so povezane z neustrezno definicijo eksonov pri številnih humanih bolezenskih stanjih.

Contents

List of figures

List of Tables

Abbreviations

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

aa	Amino acid
bp	Base pairs
cDNA	Complementary DNA
CF	Cystic Fibrosis
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
ddH_2o	Double-distilled water
DMD	Duchenne muscular dystrophy
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphate (A, C, G and T)
dsRBM	double stranded RNA-binding motif
dsRBP	double stranded RNA-binding protein
dsRNA	double stranded RNA
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid
ESE	Exonic Splicing Enhancer
ESS	Exonic Splicing Silencer
FIX	coagulation factor IX protein
hnRNP	Heterogenous ribonuclear protein
IPTG	Isopropyl-β-d-thiogalactopyranoside
ISE	Intronic Splicing Enhancer
ISS	Intronic Splicing Silencer
kb	Kilobase
kDa	Kilodalton
N	Nucleotide (A or C or G or T)
NE	Nuclear Extract

NMD	Nonsense-mediated decay
nt	Nucleotides
PBS	Phosphate buffer saline
PBS	Phosphate buffer saline
PCR	Polymerase Chain Reaction
PPT	Polypyrimidine tract
R	Purine (G or A)
RNA	Ribonucleic acid
RNA pol II	RNA polymerase II
RNA pol III	RNA polymerase III
RRM	RNA Recognition Motif
RS	Arginine-Serine rich motif
RT	Room Temperature
SDS	N-lauroylsarcosine sodium salt
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
SMN	Survival of Motor Neuron
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein particles
SR	Arginine-serine rich protein
SRE	Splicing Regulatory Elements
ss	Splice site
TBE	Tris-borate-EDTA (buffer)
U2AF	U2 snRNP Auxiliary Factor
wt	wild-type
Y	Pyrimidine (T or C)
UTR	untranslated region
GV	genomic variant

1. Introduction

Section 1.1. ***Pre-mRNA processing and overview of splicing***

Nascent RNAs 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 et al. 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 A 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, in fact account for no more than 1% of the entire genome (Lander et al. 2001). Therefore the majority of the protein-coding genes are composed mostly by non-coding, intervening sequences (introns). For example, an average human protein-coding gene spans about 27 Kb, contains a mean of 8.8 exons with a mean size of 145 nt., 3365 nt-long introns and 5′ and 3′ untranslated regions (UTR) of approximately 770 and 300 nt., respectively (Lander et al. 2001). 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 (Wang et al. 2008; Pan 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 2004a; Nilsen et al. 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.

1.1.1. Exons and introns have characteristic sequence elements which distinguish them

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.

1.1.1.1. 5' ss or donor splice site. The 5' ss marks the exon/intron junction at the 5' end of the intron (Fig.1.1.1) 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

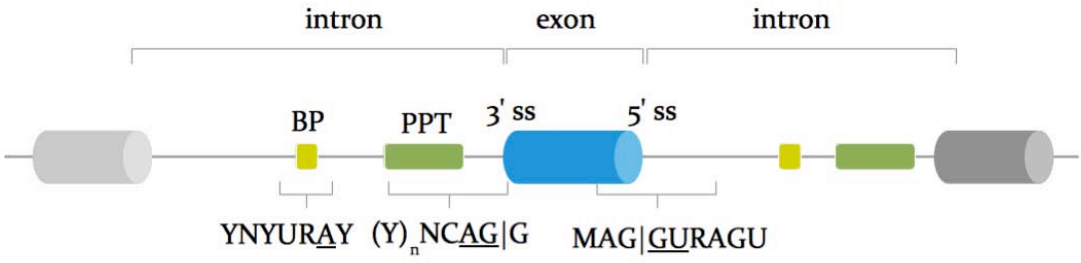

Figure 1.1.1. **Schematic representation of exon-intron boundaries**

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.

long ago to be MAG|GURAGU (M indicates A or C; R indicates purines and the | the exon/intron boundary) (Mount 1982; Shapiro et al. 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 et al. 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 (Roca et al. 2008; Carmel et al. 2004). 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 et al. 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 (Sahashi et al. 2007)

1.1.1.2. 3' ss, a composite signal. The intronic element that identifies the 3' ss usually appears several thousand 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) (Fig. 1.1.1). The BP is characterized by the presence of a conserved A surrounded by a highly degenerated motif YNYURAY (Y=pyrimidine and R= purine)(Reed et al. 1985)It is commonly found about 18-40 nucleotides upstream of the AG dinucleotide (Ruskin et al. 1984; Reed et al. 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 utilisation 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).

1.1.2. Splicing reaction: a tale of two transesterifications

In chemical terms, the RNA splicing reaction is conserved among all intron-containing organisms and consists of two consecutive transesterification steps (Lamond 1993; Moore et al. 1993) (Fig. 1.1.2.). First, the 2'-hydroxyl group of the

conserved A residue at the BP attacks the phosphate at the 5' ss. This leads to cleavage of the 5' exon from the intron and the concerted ligation of the intron 5' end to the BP 2'-hydroxyl (Ruskin et al. 1985; Konarska et al. 1985). Two reaction intermediates are then produced: a detached 5' exon and an intron/3' exon fragment in a lariat configuration containing a branched A nucleotide at the BP. The following and second transesterification step is the attack on the

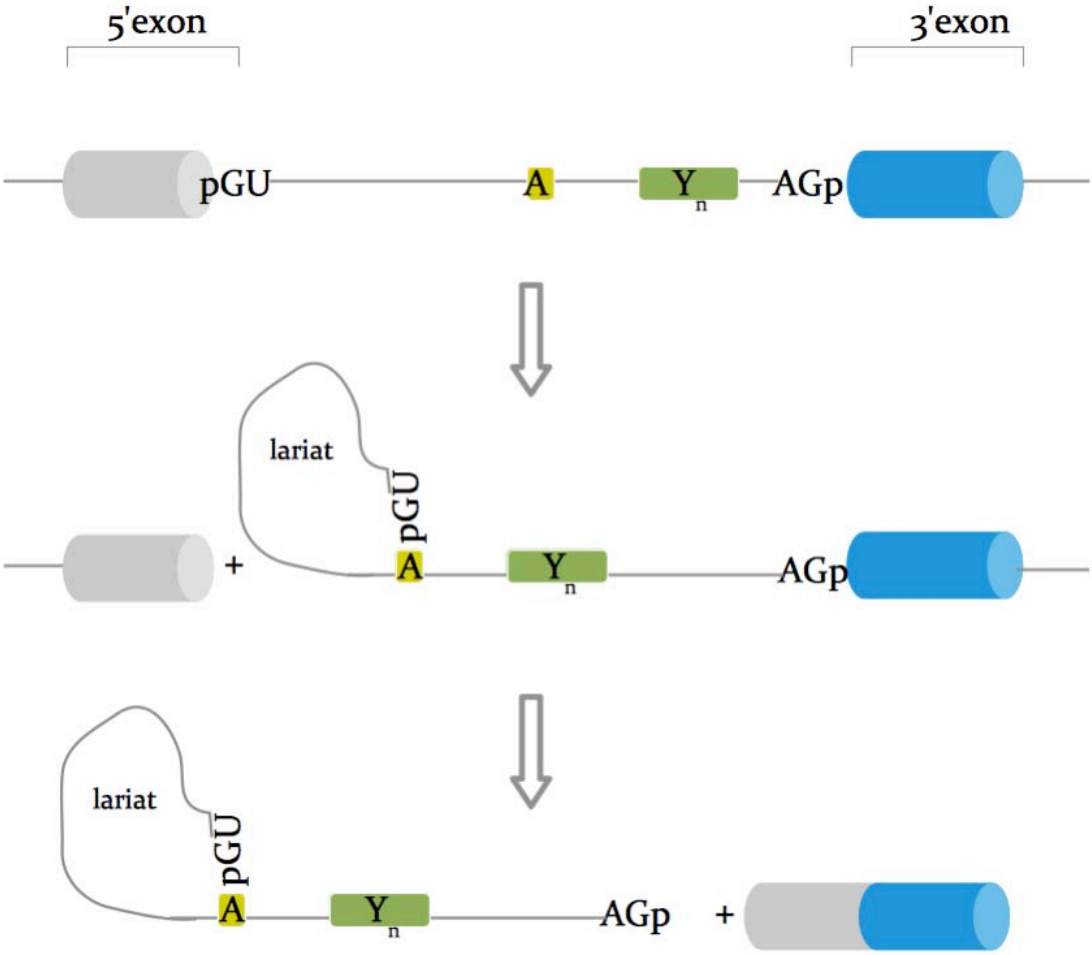

Figure 1.1.2. **Overview of a splicing reaction**

Essential splicing signals are GU/AG dinucleotides at the exon-intron and intron-exon junctions, respectively (5'- and 3'-ss), a PPT (Y_n) and the conserved A nucleotide at the BP. Splicing takes place in two transesterification steps. In the first step, the 2'-hydroxyl group of the A residue at the branch site attacks the phosphate at the GU 5'-splice site. This leads to cleavage of the 5'exon from the intron and the formation of a lariat intermediate. In the following step, a second transesterification reaction, which involves the phosphate (p) at the 3' end of the intron and the 3' OH group of the detached exon, ligates the two exons. This reaction releases the intron, still in the form of a lariat.

phosphate at the 3' end of the intron by the 3'hydroxyl of the detached exon. This ligates the two exons and releases the intron, still in the form of a lariat (Konarska et al. 1985; Moore et al. 1993). This reaction is similar to the one found in self-splicing group II introns (Fedorova et al. 2007). However, unlike them, splicing of eukaryotic organisms require the assistance of numerous assistant factors.

1.1.3. Splicing machinery: the spliceosome

The splicing reaction is orchestrated by the spliceosome, a large multicomponent modular ribonucleoprotein (RNP) complex that recognizes the exon-intron junctions, catalyzes the intron removal and joins the exons. Five uridine-rich (U-rich) small nuclear ribonucleoproteins (snRNPs) form the major building blocks of the spliceosome: U1, U2, U4, U5 and U6 snRNPs. Human spliceosomes also interact with a large number (from 150 to 300) of non-snRNP proteins, including heterogeneous nuclear RNPs (hnRNPs) and serine-arginine (SR) proteins (Wahl et al. 2009). A group of less abundant snRNPs, U11, U12, U4atac, U6atac, functionally analogous to the U1, U2, U4 and U6 snRNPs, are components of the so-called minor spliceosome and responsible for the excision of a rare class of introns, named U12-type. The U5 snRNP is shared between the two machineries (Patel et al. 2003).

Each snRNA have a U-rich RNA molecule and a variable number of associated proteins, which make up more than two-thirds of its mass. The snRNAs are transcribed by RNA pol II, with the exception of U6 and U6atac which are transcribed by RNA pol III. The U snRNPs undergo a complex maturation process before becoming functional in the spliceosome. Sm class-snRNAs are capped in the nucleus and exported to the cytoplasm, where the SMN complex assembles the Sm seven protein core (Raker et al. 1996). Sm proteins (B/B', D1, D2, D3, E, F and G) are common factors of U1, U2, U4 and U5, which bound to

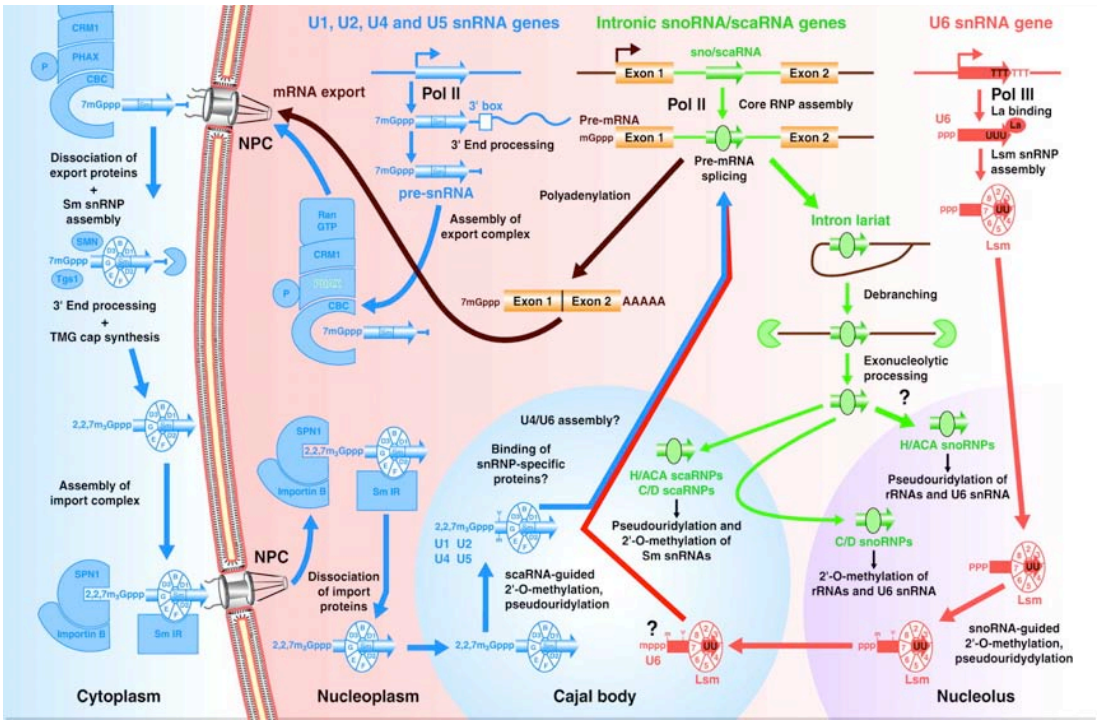

Figure 1.1.3. **Biogenesis of U snRNPs**

U1, U2, U4 and U5 are transcribed by RNA pol II and exported to the cytoplasm. There the Sm proteins are loaded onto the Sm site, 5' and 3' ends are modified and the complexes are imported to the nucleus. In the Cajal bodies scaRNA-guided modifications (2'-o-methylation and pseudouridynilation) take place, probably including the association of U snRNPs-specific proteins On the other hand, U6 snRNA is transcribed by RNA pol III and further maturation occurs initially in the nucleoplasm (Lsm proteins assembly) and later on in the nucleolus (sequence modifications carried out by snoRNAs). U6 snRNP assembly with U4 snRNP may take place in the Cajal bodies. Taken from Kiss (2004).

a conserved Sm motif on the each snRNA. The Sm proteins form a seven-member ring core structure that encircles the RNA. All the Sm proteins share a conserved Sm motif, consisting of two sets of conserved sequences (Sm1 and Sm2) separated by a large loop region, which appear to be the sites of protein-protein interactions necessary for the core structure to form. The Sm core is essential for the biogenesis, transport and function of the snRNP particles. Then the cap is hypermethylated to form the TMG cap structure exonuclease (Mouaikel et al. 2002; Huang et al. 1999). Both TMG and the SMN complex promotes the assembly of an import complex that facilitates the export back to the nucleus of the U snRNP particles and their accumulation in Cajal bodies. It is in these nuclear bodies where the small Cajal body-specific RNAs (scaRNAs) carry out some sRNA-specific modifications such as pseudouridinylation and

2'-O-methylation (Will et al. 2001).

Biosynthesis of U6 snRNA differs from the other spliceosomal U snRNAs in many aspects. First, it is transcribed by RNA pol III and the entire maturation process takes place in the nucleus, unlike the other snRNA which have a cytoplasmic processing phase. Second, U6 snRNA has a Sm-like (Lsm) motif which is bound by Lsm proteins, a protein family with high homology to Sm. Third, the formation of its 3' end is a transcription termination event rather than a RNA processing as the terminal poly (U) sequences acts as a transcription termination signal, a mechanism share by all RNA pol III genes (Paule et al. 2000). Fourth, as other RNA pol III transcripts it does not receive a m1G cap but rather contains a 5'-monomethylphosphate cap. Fourth, sequence modification in the forms of 2'-O-methylations and pseudoridinylations are carried out by small nucleolar RNAs (snoRNAs) rather than small Cajal body-specific RNAs (scaRNAs) (Ganot et al. 1999; Tycowski et al. 1998).

Besides Sm and Sm-like proteins, other particle-specific proteins associate with snRNAs. The number of specific associated proteins vary greatly among the U snRNAs, from only 3 for U1 snRNP to more than 20 for U2 snRNP (Stanek et al. 2006). The association of the U snRNP-specific proteins is believed to happen in the nucleus since it is known that import of some of them (U1A and U1C for U1 snRNP; U2A' and U2B''for U2 snRNP) occurs independently of snRNA nuclear import (Feeney et al. 1990; Jantsch et al. 1992; Kambach et al. 1994).

1.1.4. 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

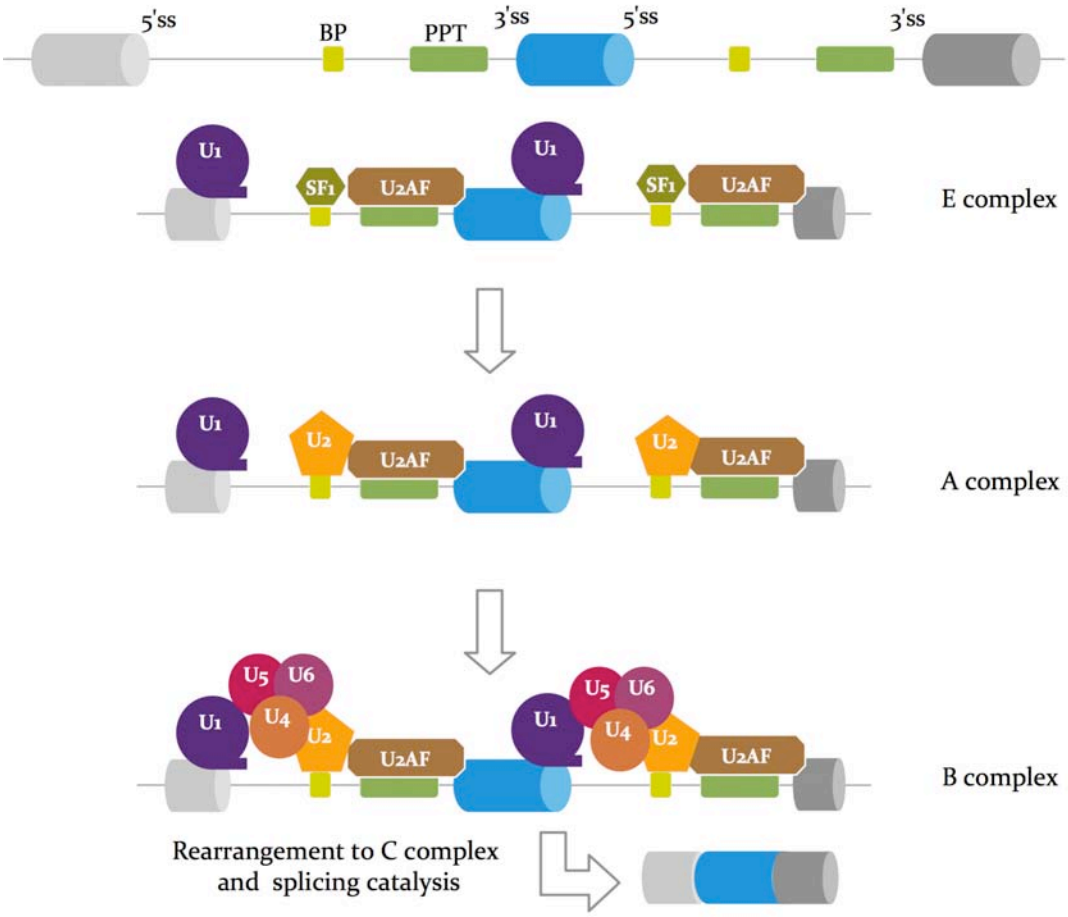

Figure 1.1.4. **The assembly of the spliceosome.**

The spliceosome assembly onto the pre-mRNA occurs in a stepwise manner. Exons are noted by cylinders, and introns by grey lines. Above the general splice signals on thre pre-mRNA are shown. The E complex contains U1 snRNP bound to the 5' ss, SF1/mBBP bound to the BP, and U2AF complex loaded onto the PPT ($U2AF^{65}$) and the acceptor site ($U2AF^{35}$). The A complex is formed after U2 snRNP takes over theBP, displacing SF1/mBBP. The tri-snRNP U4/U6/U5 then enters to form the B complex . Finally, spatial rearrangements lead to the C complex, in which U6 snRNP replaces U1 snRNP at the 5' ss and establishes contacts with U2 snRNP at the BP. Progression to catalytic steps of splicing gives rise to the mature, spliced pre-mRNA.

nature, such as RNA-RNA basepairing, RNA-protein and protein-protein binding and a lot of 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 (Hong et al 1997; Das et al. 1999; Kent et al. 2005; Frendewey et al. 1985; Jamison et al. 1992;

Tardiff et al. 2006; Bindereif et al. 1987; Bringmann et al. 1986).

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 (Will et al. 1996; Rossi 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 $U2AF^{65/35}$) 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 (see below). 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 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 (Xiao et al. 1997; Tripathi et al. 2010; Heyd et al. 2010; Stamm 2008; Shi et al. 2007; Tazi et al. 1993).

1.1.5. Alternative splicing

The use of different pairs of splice sites generates distinct mRNA isoforms from a single gene, a phenomena known as alternative splicing. Although initially alternative splicing was thought to be only a minor RNA processing pathway (Sharp 1994)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 (Black 2003; Tazi et al. 2009; Smith et al. 2000; Nilsen et al. 2010; Pajares et al. 2007). Therefore, alternative splicing is a key regulator of gene expression that also

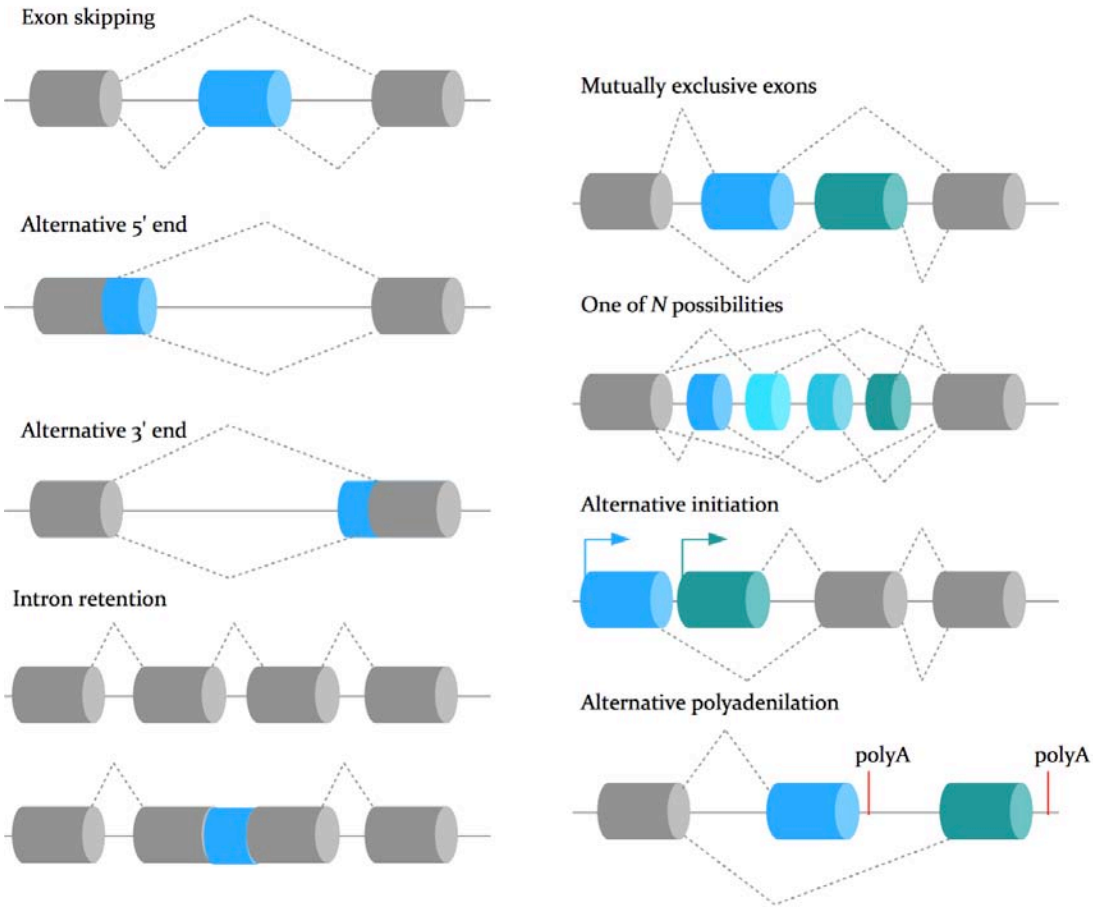

Figure 1.1.5. **Patterns of alternative splicing**

Grey boxes represent constitutive exons and coloured boxes indicate alternative exons.

contributes to proteome complexity. In fact, the enormous contribution of alternative splicing to the generation of proteomic diversity is believed to constitute a driving force in metazoan evolution (Nilsen et al. 2010; Keren et al. 2010; Xing et al. 2006; Calarco et al. 2007).

Every conceivable pattern of splicing can be found in nature (Fig. 1.1.5). A regulated exon can sometimes be included and sometimes excluded 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(Black 2003; Smith et al. 2000) 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 although it is clear that further studies will be required to fully understand these processes. Most alternatively spliced exons are thought to be controlled by multiple auxiliary *cis*-acting splicing sequences (see below),whose activity depends on their location relative to the canonical splice sites and the *trans*-acting factors that recognized them (Black 2003; Cartegni et al. 2002). Indeed 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; Tilgner et al. 2009; Batsché et al. 2006). 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 (Das et al. 2006, 2007; Cramer et al. 1999; 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 (Luco et al. 2010; Blaustein et al. 2005).
The emerging picture of alternative splicing events underscores its extreme richness and consolidates the view that its regulation is at the heart of gene expression and cell fate.

1.1.6. Auxiliary non-canonical cis*-acting regulatory sequences are needed for proper exon/intron discrimination*

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; Romano 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: this fact put in evidence that these consensus elements 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 et al. 2000; Bruce et al. 2001).
In addition to the canonical signals reported above, there are other *cis*-acting elements in pre-mRNA that regulate splicing process. Back in 1987 Mardon and colleagues reported the first evidence that internal exonic sequences far from 5' and 3'ss are fundamental for exon recognition (Mardon et al. 1987). Since then a wide array of exonic and intronic sequence elements have been

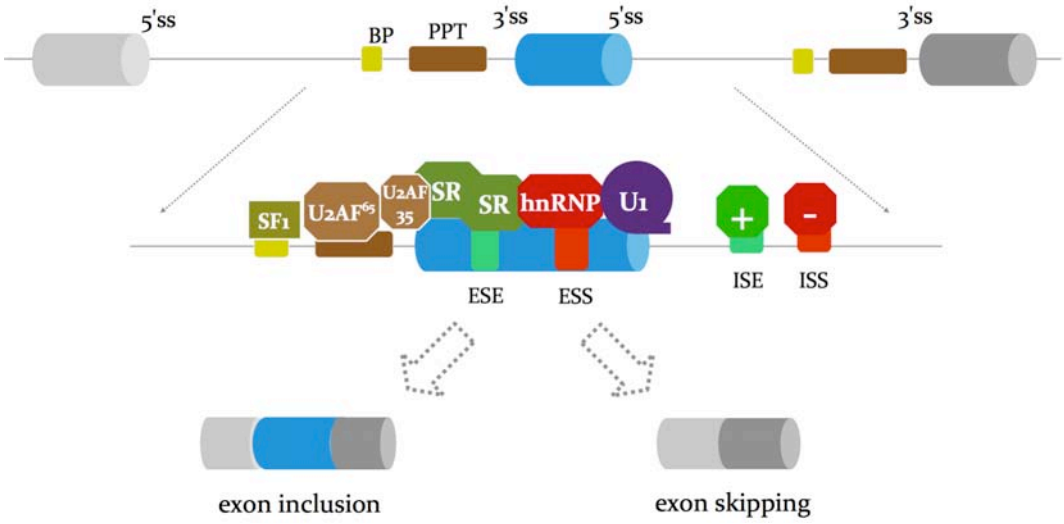

Figure 1.1.6. **Auxiliary *cis*-acting elements in pre-mRNA splicing**
Schematic representation of possible distribution of canonical and additional splicing *cis*-acting elements.
In addition to canonical splicing signals (5' and 3' ss, BP and PPT), enhancer and silencer elements in the exons (ESE; ESS) and/or introns (ISE; ISS) allow the correct splice sites to be distinguished from the many cryptic splice sites that have identical signal sequences. The multiple, weak interactions between the spliceosomal components bound at the canonical splice sites and the trans-acting splicing factors loaded onto the auxiliary cis-acting splicing elements define whether an exon will be included or skipped in the mature mRNA.

reported to regulate splicing in both constitutive and alternative exons (Stamm et al. 1994; Horabin et al. 1993; Guo et al. 1993; Schaal et al. 1999; Q. Wu et al. 1998; Tanaka et al. 1994; van Oers et al. 1994; Xiao et al. 2007; Lavigueur et al. 1993).

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) (Cartegni et al. 2002)and the distribution of these regulatory elements are higher in real exon and introns (Sironi et al. 2004; Zhang and Chasin 2004). In spite of this, these elements are not always well defined and their functions may overlap. In fact, in some systems it may be more appropriate to talk about composite exonic regulatory elements of splicing (CERES) as described for CFTR exon 9 and 12 (Pagani et al. 2003a, b).

1.1.6.1. 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). Indeed 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. The first ESE mapped was purine-rich, a alternate run of As and Gs, nonetheless a run of either Gs or As doesn't refer a ESEs (Tanaka et al. 1994; Marcucci et al. 2007). Through interactions with a subset of SR proteins, purine-rich ESEs recruit or strengthen the binding of basic splicing factors to suboptimal splice sites and stimulate spliceosome assembly (Graveley and Maniatis 1998; Lavigueur et al. 1993; Sun et al. 1993; Zuo et al. 1996; Roscigno et al. 1995). However, an exon sequence having one or more SR binding sites does not necessarily function as an ESE (Zheng et al. 1999, 1998) since SR proteins also bind splicing suppressors (ESSs) (Mayeda et al. 1999; Zheng et al. 1998) .

The other class is the non-purine-rich ESE. This class comprises the exonic AC-rich enhancer and exonic pyrimidine-rich enhancer. The AC-rich enhancers were first identified by *in vivo* selection experiments and were found to stimulate splicing both *in vivo* and *in vitro* (Coulter et al. 1997). AC-rich ESEs have been shown to be involved in the regulated splicing of both viral and cellular genes (Coulter et al. 1997; Gersappe et al. 1999; Zheng et al. 2000).

In addition to these classes, exonic pyrimidine-rich enhancers have been described in β-globin RNA (Schaal et al. 1999) and other mammalian RNAs (Staffa et al. 1997; Dirksen et al. 2003).
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.

1.1.6.2. 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).
1.1.6.3. 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 (G_3), that acts as ISE elements to enhance recognition of adjacent 5' or 3' ss (McCullough et al. 1997). For example, G runs in *THPO* genes shows that the intronic G motifs are involved with 3' ss definition by a combinatorial effect (Marcucci et al. 2007). In this model, precisely the G7 and G10 motifs present in the intron 2 collectively facilities the recognition of proper 3'ss by interacting with hnRNP H, whereas deletion of these two G motifs activates the cryptic 3' splice site.
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) .

1.1.6.4. 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 cell 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 renamed 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, b) and exon 10 of Luteinizing hormone receptor type 1 and 2 (Gromoll et al. 2007) 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).

1.1.7. 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 (Shepard et al. 2009; Huang et al. 2005; Graveley 2000). 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 o 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; 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 (Stamm 2008; Sanford et al. 2005; Blaustein et al. 2005; Misteli et al. 1997; Bourgeois et al. 2004).

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 (Hocine et al. 2010; Köhler et al. 2007) Some hnRNP are extremely abundant (about 100 milion copies per nucleus), while others are present in a lower amount (Kamma et al. 1995; Markovtsov et al. 2000).

The most prominent structural feature of hnRNPs is that they contain RNA-binding motifs as well as auxiliary domains that have been shown mediate protein-protein interactions (Dreyfuss et al. 1993; 2002).

The hnRNP proteins frequently mediate splicing repression, particularly through binding to exonic splicing silencer (ESS) elements or by sterical interference with other protein interaction (Cartegni et al. 2002; Zhu et al. 2001; 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 (Piñol-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. A classical example is hnRNP A1, which was first identified to counteract SF2 during 5' splice site selection of the E1A premRNA (Mayeda et al. 1992). It also modulates 3' ss choice as in the case of K-SAM exon in fibroblast growth factor receptor-2 mRNA (Del Gatto et al. 1997)). Although hnRNP A1 exhibits the highest binding affinitiy for the UAGGGY sequence element, it can also bind other RNA sequences with lower efficiency. It is known that hnRNP A1 can interact with itself, other hnRNPs proteins and spliceosomal U2 and U4 snRNPs (Cartegni et al. 1996; Honoré et al. 1995). Thus it is believed that hnRNP A1 negatively influences splicing through both direct binding to RNA sequences and protein-protein interactions, forming large complex which sterically interfere with spliceosome loading onto the pre-mRNA or reduces its affiity for the canonical splice sites.

Many other cases of hnRNPs antagonizing SR proteins and promoting exon skipping have been reported over the last years (Okunola et al. 2009; Rooke et al. 2003; Guil et al. 2003; Singh et al. 2004b; Jun Li et al. 2003; Dreumont et al. 2010). Alternative mechanisms by which hnRNP proteins promote exon skipping consist in inhibiting intron or exon definition crosstalks (Izquierdo et al. 2005; Sharma et al. 2008), direct competition with U2AF65 for the binding to the PPT (Lin et al. 1995; Matlin et al. 2007) or looping of exons through contacts between hnRNPs flanking these exons (Lamichhane et al. 2010; Oberstrass et al. 2005).

1.1.8. RNA secondary structure and alternative splicing

RNA has an intrinsic tendency to form highly stable and complex secondary and tertiary structures. However, *in vivo,* as soon as the pre-mRNA leaves the polymerase pore, it is bound by many sequence-specific and non-specific RNA binding proteins generating mRNP complexes. Consequently, pre-mRNA has a very short timespan to form intramolecular interactions. These RNA secondary structures might act as *cis*-elements and their influence on splicing can occur in multiple ways.

RNA duplexes can be recognized in a sequence-independent fashion by double-stranded (ds) RNA binding proteins, characterized by a dsRNA-binding motif (dsRBM) (Fierro-Monti et al. 2000). Proteins carrying dsRBMs are classified in nine families with varied functions. some of which are well-established and others are yet-unknown. One of these families is composed by the dsRNA adenosine deaminase family, which is involved in the covalent modification of dsRNA. These proteins promote the deamination of adenosine to inosines (I) by the RNA editing process, which is dependent on pre-mRNA secondary structures (Bass et al. 1988; U. Kim et al. 1994; Melcher et al. 1996). These enzymes bind to the RNA via dsRBM and can decrease the stability of

the structure by editing a A-U base pair into a less stable, non-canonical I-U base pair (Patterson et al. 1995; Bass et al. 1988). The destabilization of the helix can affect the recognition of splicing functional sequences by the mechanisms previously mentioned. Another possibility is that editing destabilize a double-stranded structure, thereby allowing access for proteins and spliceosomal components to splice sites or cis-acting elements. RNA editing could modulate splicing processing through base changes in *cis*-acting signals (Schoft et al. 2007).

RNA secondary structures may hinder the accessibility of splicing factors to functional sequences (either splice sites or other regulatory sequences) present in the pre-mRNA by sequestering them in stems or looping them out (Hiller et al. 2007) On the other hand, they can stimulate splicing processing when silencer elements are trapped within a RNA structure or vary the relative distance between regulatory elements, consequently altering their ability to modify splice site recognition (Buratti et al. 2004)..

1.1.9. U1 small nuclear RNA (U1 snRNA)

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 et al. 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 (Fig. 1.1.7). 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; Gridasova et al. 2005; Hirsch et al. 2000). 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). 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). The transcriptional stimulatory activity of these factors relies on multiple interactions with the general transcription machinery components associated to the SNAPc (Ford et al. 1998; Mittal et al. 1996).

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 (Kiss 2004; Huang et al. 1997). 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 (Egloff et al. 2008; Uguen et al. 2003).

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

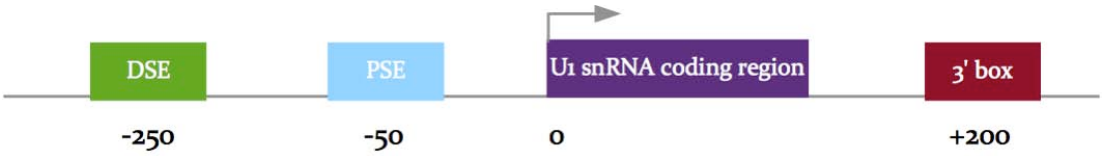

Figure 1.1.7. **The structure of human U1 snRNA gene transcribed by RNA pol II.**
The diagram shows the DSE and PSE *cis*-acting promoter elements and the 3'box *cis*-acting RNA-processing element of pol II-transcribed U1-snRNA gene boxed, with their position relative to the transcription start site noted below. The start site of transcription is marked with an arrow above the line.

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).

1.1.10. 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 (Fig. 1.1.8). 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 (Hamm et al. 1990; Surowy et al. 1989; Yuo and Weiner 1989). 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

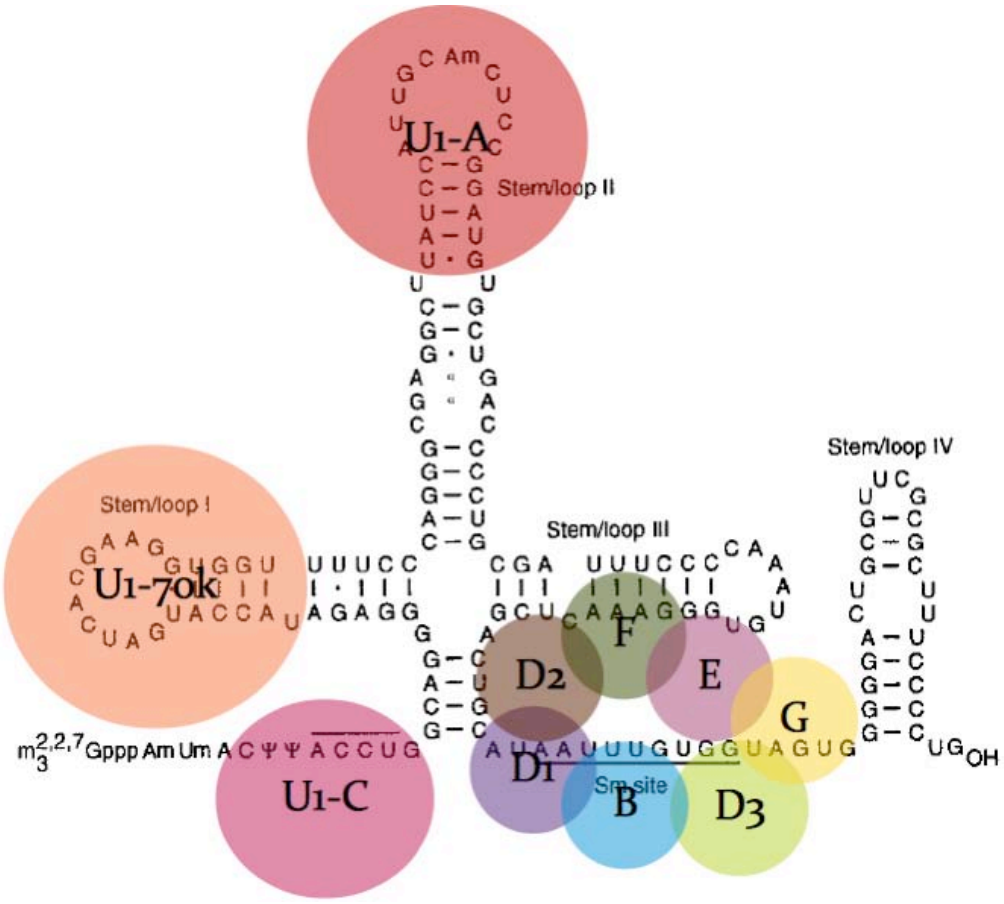

Figure 1.1.8. **Schematic structure of U1 snRNP**

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 (Pommeranz-Krummel et al 2009; Hernandez et al 2009).

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 et al. 1986). Although 5' ss recognition can occur in absence of a U1 snRNP 5'-tail (Du et al. 2002), the splicing efficiency is compromised. U1C protein contributes to 5' ss recognition and its stabilizes

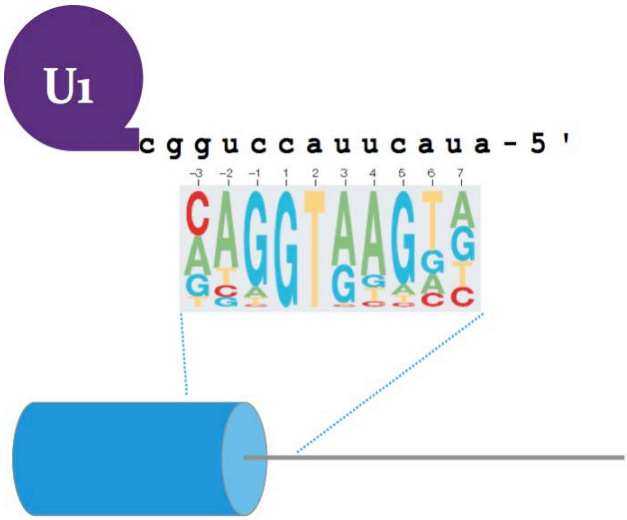

***Figure 1.1.9*. Schematic representation of U1 snRNA 5'-tail that binds to 5'ss consensus sequence.**

the RNA duplex between the 5' ss and the U1 snRNP 5'-tail (Du et al. 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 et al. 1996).

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 et al. 1993; Cao et al. 1998; Jamison et al. 1995; Wang et al. 1998; Adams et al. 2001) through its RS domain. These interactions are expected to facilitate U1 snRNP recruitment to the 5' ss and subsequent formation of the E complex (Cho et al. 2011). The importance of U1-70k behind U1 snRNP activity is further evidenced since it s the only U1 snRNP-associated protein which is regulated by phosphorylation (Tazi et al. 1993). Interestingly, U1-70k phosphorylation do not impede spliceosome formation but blocks splicing catalysis (Mermoud et al. 1994).

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 et al. 1986; Siliciano et al. 1988; Liao et al. 1990; Séraphin et al. 1988). Increased base pairing to a mutant 5' ss can nevertheless maintain usage of cryptic splice sites located nearby (Siliciano et al. 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 et al. 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 et al. 1993; Lesser et al. 1993; Tarn et al. 1994).

Section 1.2. ***Defective splicing and disease***

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

1.2.1. Pre-mRNA splicing and disease

Pre-mRNA splicing defects are likely to have an impact on medical practice because they seem to have a role in almost all diseases with a genetic aetiology (Tazi et al. 2009; Baralle et al. 2009).It is in fact now clear that substitutions which 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; Buratti et al. 2006). Although the frequency of splicing mutations varies considerably between individual genes, initial estimates considered that approximately 15% of pathogenic mutations cause disease through the defect they introduce in the splicing mechanism (Table 1.2.1).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; Ohno et al. 2001; Aznarez et al. 2003). Thus the primary mechanism of disease behind mos 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 (Wang et al. 2007; Baralle et al. 2005). As research has 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 (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 years ago (Vibe-pedersen et al. 1984) has allowed a relatively fast approach to identify splicing spoilers and to study their underlying functional mechanism (Baralle et al. 2005; Cooper 2005).

Gene	Disease/ Phenotype	Splicing mutations (%)
ATM	Ataxia telegectasia	18
BRCA1	Breast cancer predisposition	9
CADM	Medium chain acyl CoA dehydrogenase deficiency	10
CFTR	Cystic fibrosis	14
DMD	Duchenne muscular dystrophy	9
HBA1/2	Blood disorders (thalassemias, anaemia etc.)	3
HBB	Blood disorders (thalassemias, anaemia etc.)	10
HPRT	Hypoxantine-guanine phosphoribosyltr. I deficency	15
IKBKAP	Dysautonomia, familial	33
MAPT	Frontotemporal dementia and Parkinsonism	33
MLH1	Colorectal cancer	18
MSH2	Colorectal cancer	9
NF1	Neurofibromatosis type I	19
NF2	Neurofibromatosis type II	22
RHO	Retinitis pigmentosa	3
SMN1/2	Spinal muscular atrophy	4
WT1	Wilms tumor	11

Table 1.2.1. **Frequency of splicing defects in common human disease.**
*The data are calculated from the public Human Gene Mutation Database (24/11/2008). Adapted from (Baralle et al. 2009).

1.2.2. Mutations at canonical splice sites

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 (Krawczak et al. 2007)..

The degenerate nature of splice site consensus sequences makes assessment of pathogenicity of a new genomic variants a difficult task. Except for the mutations that destroy the invariant GT or AG dinucleotides on donor sites and acceptor sites respectively, the presence of a GV on other positions it is not always indicative of pathogenicity. For example, while single nucleotide polymorphisms (SNP) display an even distribution over all base positions in human 5' ss, pathological GVs tend to cluster at certain positions, particularly -1 and +3 to +6 for donor sites and -3 for acceptor sites (Krawczak et al. 2007). Pathological PPT mutations are less abundant than those affecting splice sites and show a homogeneous distribution ((Krawczak et al. 2007).

At the 5' ss, mutations affecting the GT residue at position +1 and +2 are the most common, 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-mRNA 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 RNA 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). For instance, hybrid minigene system have been extensively used for evaluation of splicing-associated mutations in many different gene models, such as CFTR, NF1, ATM

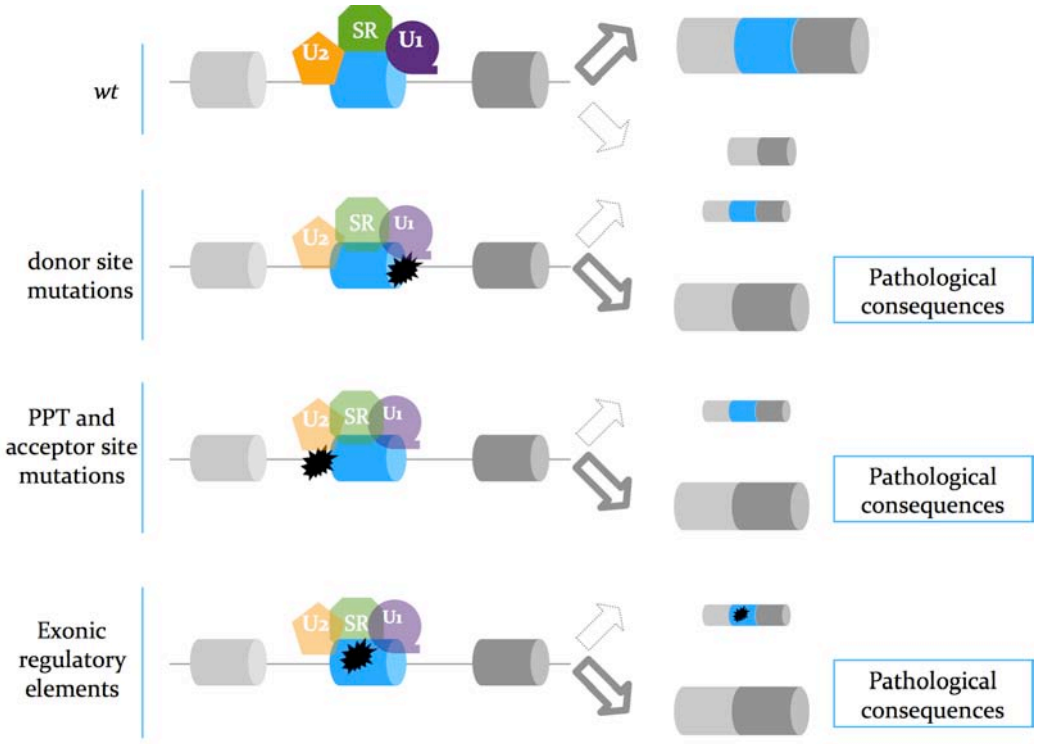

***Figure 1.2.1.*Mutations affecting canonical or exonic regulatory elements can trigger pathological aberrant splicing**

and others (Pagani et al. 2004). At the same time, this method allows to uncover novel regulatory elements, either intronic or exonic, which may participate in splice site selection (Cooper 2005).

1.2.3. Mutations involving non canonical splicing regulatory sequences

Nucleotide substitutions that occur within *cis*-acting elements can be classified as loss or gain 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 (Covaciu et al. 2011; Drögemüller et al. 2011; Whiley et al. 2010; Fukao et al. 2010; Vidal et al. 2009; Skoko et al. 2008; Petkovic et al. 2007; Pagani et al. 2003a; Hovhannisyan et al. 2005).

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 *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; 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 (Cartegni and Krainer 2002), and the other is that it creates an ESS to which the splicing suppressor hnRNP A1 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 65% of DMD mutations are genomic deletions, a large number of exonic and intronic point mutations cause disease through aberrant splicing (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).

1.2.4. Bioinformatic tools can predict pathogenicity of new GVs

Several bioinformatic tools have been developed over the last two decades, exhibiting a notorious growth in number and quality after the completion of the Human Genome Project.

As previously mentioned, an exonic GV can cause disease through an alteration of the coding sequence or mis-regulation of exon splicing. The latter likely occurs when the exonic GV affects ESE, which are oftenly undescribed for many cases. The use of systematic evolution of ligands by exponential enrichment (SELEX) has been utilized to identify sequences that can act as ESEs in response to specific SR proteins characterizing SR-protein-specific sequence motifs (Liu et al. 1998, 2000). The frequencies of individual nucleotides at each position were used to derive score matrices in order to predict the location of SR protein-specific putative ESEs and design an ESEfinder web-based program (http://rulai.cshl.edu/tools/ESE/) (Liu et al. 2000, 1998; Cartegni et al. 2003). In order to identify enhancer sequences additional *ab initio* computational approaches have also been developed. For instance, the RESCUE-ESE (Relative Enhancer and Silencer Classification by Unanimous Enrichment) identifies putative ESE motifs by selecting hexamers that are enriched in exons against introns and weak against strong splice site scores (Fairbrother et al. 2002, 2004). Similarly, the PESX approach (Putative Exonic Splicing Enhancers/Silencers) allows the identification of putative exonic splicing regulatory elements by comparing over-represented octamers in internal non-coding exons versus unspliced pseudo exons and the 5' untranslated regions of intronless genes. Oligonucleotides that are sufficiently overrepresented in noncoding exons are selected as PESE elements (Zhang and Chasin 2004; Zhang et al. 2005).

The first attempt to compute splicing sites strength has been reported more than 20 years ago by Shapiro & Senapathy (1987). They developed a position-

specific weight matrix (PSWM) for 3' and 5' ss that reflects the degree of sequence conservation of the 3755 5' and 3' known splice sites at that time. This approach assumes independence of each nucleotide position. For 5' ss they used appropriate weights or frequency percentages in order to score every 8bp-lenght subsequence in the input sequence, based on the PSWM. For 3'ss the method is somewhat more complicated, since scoring comprises the NCAGG consensus around the invariant AG dinucleotide and the PPT. The formula weights the PPT and the AG consensus equally. From these matrices, the S&S score has been derived within a range 0-100, being 100 the score of consensus 5' or 3' ss sequences. are then given the maximal score, that is 100, which decreases as input sequence varies from consensus. Therefore input splice sites with high similarity to consensus sequences are considered strong splicing sites and increasing dissimilarity is perceived as weakening of the splice site strength (Shapiro et al. 1987). Although this intuitive reasoning is commonly still considered valid, it is known that simple resemblance to consensus sequences is not a reliable variable to predict splice sites usage/strength. Weight matrix models (WMM) represent an extension to the S&S score, since they take into consideration the relative importance of each base at every position. Base position within the putative splice site sequence are no longer considered as an independent factor, as the matrix quantifies the relative likelihood of a given candidate splice site sequence with respect to the background nucleotide distribution from a training set of splice sites signals. Nevertheless they still fail to incorporate nucleotide interdependencies. Weight array methods slightly increase the discriminative power of WMM by using conditional probabilities for pairs of neighbouring nucleotides. Both WMM and WAM strategies are integrated into the maximum decomposition model (MDD) (Yeo et al. 2004) which is an iterative decision-tree approach that captures the strongest dependencies, including those between non-neighbouring positions. A least biased approach is constituted by the

maximum entropy model (MaxEnt), which does not make assumptions based on distribution but rather ponders consistency of the input sequence with a set of known real and decoy signal sequences. This approach incorporates local adjacent and non-adjacent positions into the calculation of the score, taking into account a series of constraints estimated from the available data. In the case of 3' ss, since the consensus motif is longer than the 5'ss MaxEnt initially breaks up the input sequence into overlapping pieces and the final likelihood is calculated from the appropriate ratio of individual segment distributions using second-order marginal constraints in each segment.

The computational tools described so far require previous manual selection of relevant information features. Machine learning techniques on the other hand deduce classification functions or rules from a pool of training data sets (known real and decoy splice signals). Neural network (NN) is a machine learning approach that recognizes sequence patterns once it is trained with sets of DNA sequences containing known real and decoy splice signals. Using a backpropagation feedforward neural network with one hidden layer it produces an output score between 0 and 1 for each candidate sequence (Reese et al. 1997).

For 5' ss, some strategies have been developed based on U1 snRNP recognition and duplex formation between donor site sequence and 5'-tail. Thermodynamic calculation of the 5' ss:U1 snRNA duplex is quantified by its free energy ΔG, and the output value reflects the contribution of hydrogen bonding, base stacking, mismatches and Watson-Crick and wobble base pairs (Carmel et al. 2004). Hydrogen bond formation is the principle of another 5' ss: U1 snRNA based approach to estimate splice site strength. Hbond score considers the formation of hydrogen bonds between 5'ss and the 11 nt of the U1 snRNA 5'-tail, taking into the calculation positions +7 and +8 of 5'ss (usually dismissed in other splice site scoring tools) and the nucleotide interdependence beyond neighbour bases (Freund et al. 2003).

1.2.5. Corrective therapies for splicing defects

1.2.5.1. *Antisense oligonucleotides and small molecules.* Because splicing is a prevalent cause of disease, significant efforts have been made to modulate splicing in a tissue-specific, transcript-specific manner. For that purpose, several strategies have been tested in order to reverse disease-causing splicing defects and alleviate its symptoms. Antisense oligonucleotides (ASOs) and small molecules are two of the most promising approaches (Cooper et al. 2009; Hammond et al. 2011). ASOs are designed to target and block pre-mRNA sequences that are functionally important for a disease-causing alternative splicing event. The use of ASO as therapeutic molecules is providing encouraging results for the treatment of SMA. ASOs 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).

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

1.2.5.2. 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 U1 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 (Müller 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 human *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 use of modified U1 snRNPs suppress the impact of disease-associated splice site mutations, complementing the loss of the normal U1 snRNP 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 (FVII) deficiency (Pinotti et al. 2008, 2009), 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 U1 snRNPs (Pinotti et al. 2009; Hartmann et al. 2010).

Section 1.3. ***Coagulation Factor IX deficiency***

1.3.1. F9 gene and protein.

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 (Fig. 1.3.1A).

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) (Fig. 1.3.1B).

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 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) (Fig. 1.3.1C). 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 (Fig. 1.3.1D) (Bowen 2002).

1.3.2. 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 (Fig. 1.3.2). Subsequent stages of the cascade then proceed, culminating in the deposition of fibrin, the structural polymer of the blood clot (Bowen 2002).

1.3.3. Molecular basis of the haemophilia B

The mutations causing haemophilia B have been localized and characterized in several hundreds of patients. What is clear is that the enormous number of mutations that have been elucidated is that the molecular basis of haemophilia is extremely diverse.

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) (Ketterling et al. 1999; Koeberl et al. 1990).

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

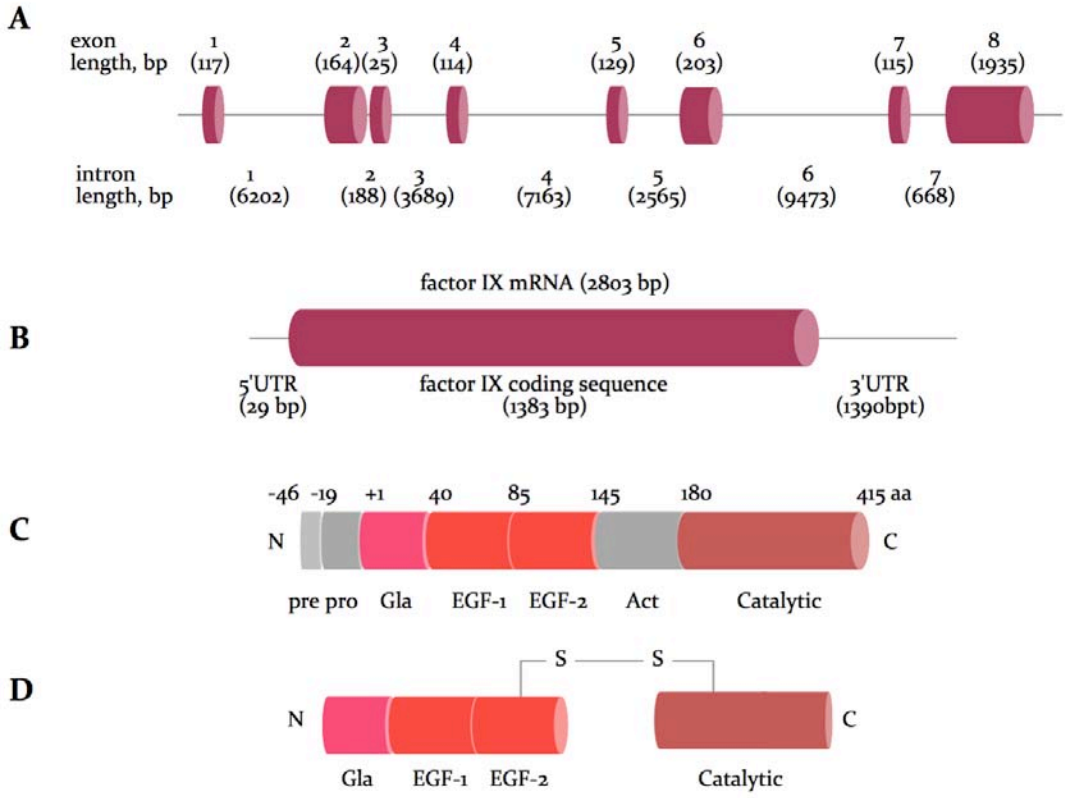

Figure 1.3.1. **Schematic representation of factor IX gene, mRNA and protein**

A, Genomic organization of F9 gene. Exon and intron number and size are reported above and below, respectively.

B, Factor IX mRNA showing the relative size and location of the ORF.

C, The newly synthesized factor IX protein molecule comprising a pre-pro signal sequence and a mature peptide of 415 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 residues 132 and 279. Gla, Gla domain; EGF, epidermal growth factor-like domain; Act, activation peptide released after proteolitic cleavage; catalytic, the serine protease domain.

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 point mutations arise from CG transitions while the remaining 70% do not occur a CG 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; Liu et al. 2001; Ketterling et al. 1999).

Deletions of factor IX 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 et al. 1994; Giannelli et al. 1996).

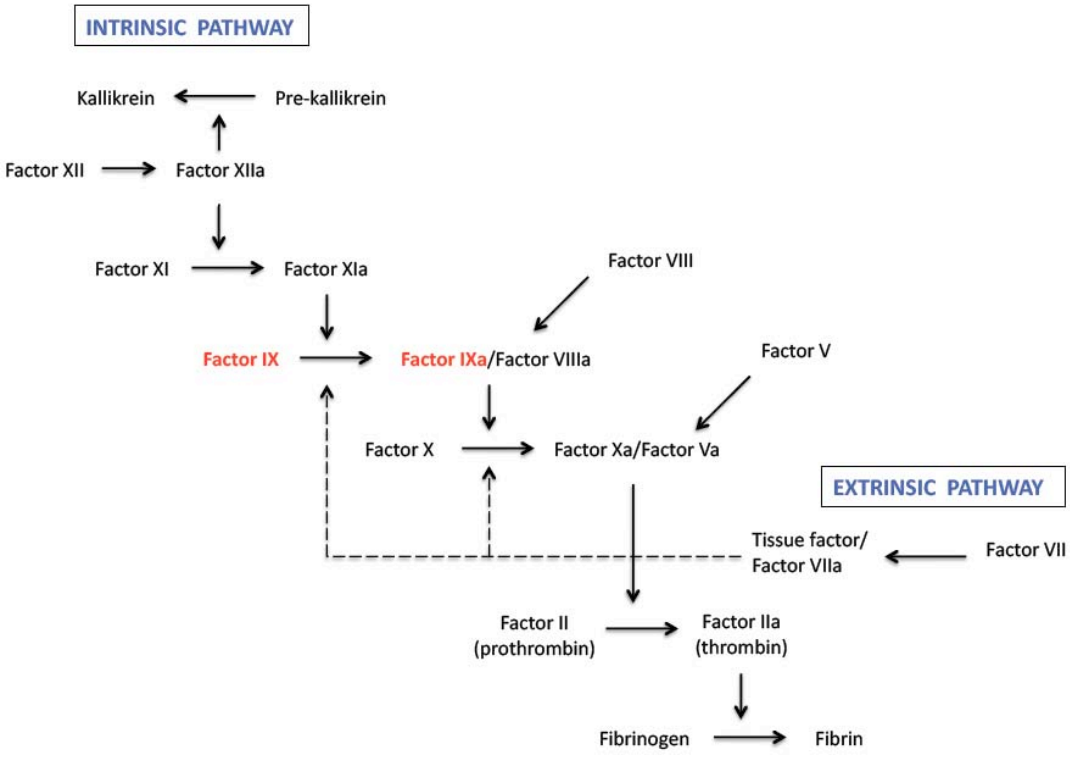

Figure 1.3.2. **Schematic showing the intrinsic and extrinsic pathways of coagulation cascade leading to fibrin formation.**
A deficiency of coagulation factor VIII or factor IX compromises the activation of coagulation factor X, the ensuing reactions are inefficient and haemophilia results.

Section 1.4. ***Cystic fibrosis***

1.4.1. General aspects of cystic fibrosis

Cystic fibrosis (CF), also known as mucoviscoidosis is an autosomal recessive genetic disease that mostly affects the entire body causing progressive disability and early death. The frequency of the disease differs among ethnic groups. However, it is more common in the Caucasian with a incidence of 1 in every 3300 whilst within Hispanics the occurrence is 1 in 9500 and it is an even rarer disorder in native Africans and native Asians, where it is estimated to occur in less than 1 in 50000 individuals (Goss et al. 2004)..

CF is an extremely heterogeneous disease both for the age of onset and for the clinical features. A number of individuals escape detection in the first decade or two of life, often because symptoms are mild. However, the typical patient with CF generally shows symptoms like thick production of mucus, which causes an obstructive chronic lungs disease, exocrine pancreatic insufficiency, intestinal obstruction of ileum in the newborns (meconium ileus) and sterility in 95% of men and in 10% of women(Kerem et al. 1996; Koch et al. 1993). Although several organs are affected in CF, the underlying principle mechanism of pathogenesis is considered to be the loss of epithelial plasmamembrane chloride conductance. Decrease in salt and fluid secretion is responsible for the blockage of exocrine outflow from the pancreas and the accumulation of heavy and dehydrated mucus in the airways. Later onset, due to the favourable condition or loss antimicrobial activity in the airway surface, causes respiratory infections by pathogens like *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Aspergillus fumigatus* (Goldman et al. 1997).

1.4.2. Molecular basis for CF

The gene CFTR (Cystic Fibrosis Transmembrane conductance Regulator) that

encodes the protein is located at the human chromosome 7 (region q31). This was the first gene identified by positional cloning (Riordan et al. 1989). The mRNA encoded from the gene is of 6kb and contains 27 exons. CFTR is a membrane associated glycoprotein of 169 kDa in size and consists of 1480 amino acid (Gregory et al. 1990). Structurally the protein has two nucleotide binding domains (NBD1 and NBD2), one regulatory domain (R) and two hydrophobic transmembrane domain (each consist of six membrane spanning segments) (Riordan et al. 1989) The NBD1 subdivision of the protein is encoded by exon 9 through exon 12. In particular, exon 12 encodes amino acid spanning 560 to 588 towards the C terminal of the NBD1 (Lewis et al. 1995; Pagani et al. 2003a; Slomski et al. 1992). In brief, the protein functions to regulate the channel through phosphorylation of the R and NBDs. The proposed mechanism is that partial phosphorylation of the R domain cause the dimerization with NBD1, which leads to the opening of the gate at the transmembrane domain. Whereas, complete phosphorylation of the R domain cause the closing of the gate by interacting with the NBD2 (Bompadre et al.

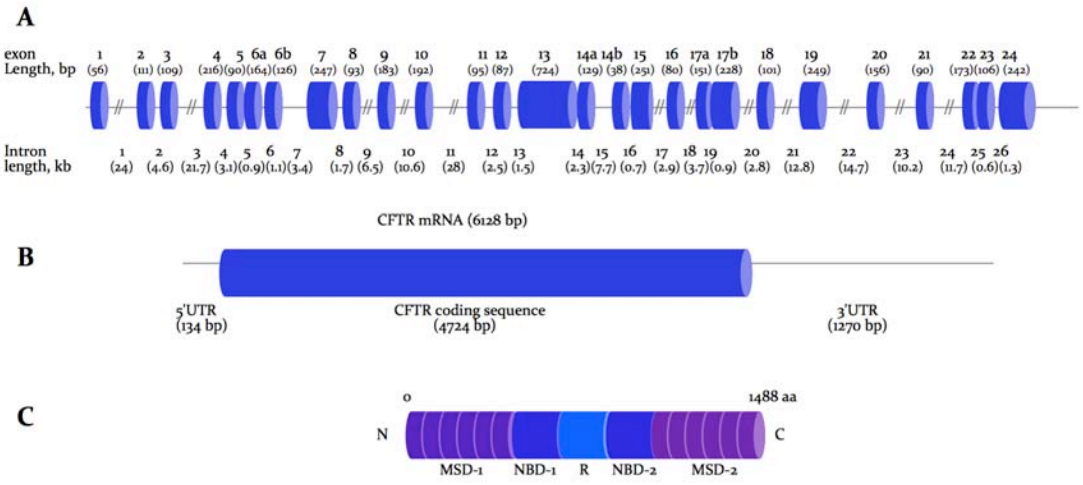

Figure 1.4.1. **Schematic representation of CFTR gene, mRNA and protein**

A, Genomic organization of *CFTR* gene. Exon and intron number and size are reported above and below, respectively.

B, *CFTR* mRNA showing the relative size and location of the ORF.

C, CFTR protein. MSD-1 and MSD-2, membrane-spanning domain 1 and 2; NBD-1 and NBD-2, nucleotide binding domain 1 and 2; R, regulatory domain.

2005a, b). Mutation to any of these functional domains produces defective proteins and leads to disease.

The symptoms among CF patients differ a lot and it is hard correlate the phenotypic symptoms with the genotype. However, phenotypic symptoms can be categories in three groups. In the first one, we can include the symptoms that are common to most CF patients, regardless of the type of mutations. An example is the abnormal electrolyte composition of sweat that is common to virtually all patients with classical CF. In the second category there are features which show a good correlation between genotype and disease phenotype. This category is best represented by the pancreatic function of the patients. The third category of phenotypic features includes symptoms that do not show significant correlation with genotype, such as the pulmonary status of CF patients, in which the severity of the disease is strongly affected by

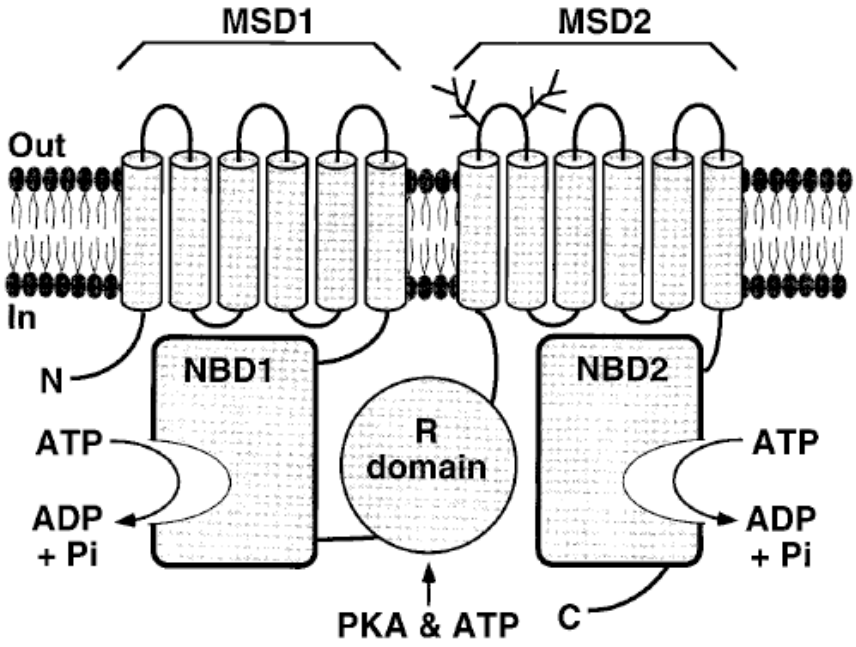

Figure 1.4.2. **Model of CFTR protein structure**

Model showing proposed domain structure of CFTR protein. MSD-1 and MSD-2, membrane-spanning domain 1 and 2; NBD-1 and NBD-2, nucleotide binding domain 1 and 2; R, regulatory domain; PKA, cAMP-dependent protein kinase. Taken from Sheppard and Welsh, 1999.

environmental and secondary genetic factors (CF modifiers) (Zielenski 2000; Zielenski and Tsui 1995).

The number of mutations in the CFTR gene causing these symptons is growing every day. The complete list of mutation identified up to now can be found at the world web access http://www.genet.sickkids.on.ca. In most of the cases, the pathological effect of single base substitution in CFTR gene were considered to be based on the change in coding sequence (or were mis-classified as benign polymorphism databases). Nowadays, several examples have been shown that many of these changes affect splicing pattern of the gene instead changing the specific amino acids (Pagani et al. 2000, 2003a; Zielenski et al. 1995; Vankeerberghen et al. 1998)

Section 1.5. ***Aim of the thesis***

Adequate splicing processing relies on a multitude of weak interactions between spliceosomal and *trans*-acting splicing factors, bound to canonical and less-defined splicing regulatory elements. The initial recognition of the donor splice sites by the U1 snRNP is a key event defining and committing the exon to splicing catalysis. The degenerate nature of the donor splice site sequences allow for some local mismatches to the U1 snRNP base pairing. Thus the pathological impact of new genomic variants is not a straightforward process and should be assessed through splicing functional assays.

In this thesis work, I have explored the impact on exon processing of several disease-associated mutations in C*FTR* exon 12 and *F9* exon 5 and the involvement of U1 snRNP on the defective splicing defect.

Using different minigene systems and modified U1 snRNPs the following aspects have been investigated: :

-the impact on pre-mRNA processing of several mutations affecting different splice signals (PPT, exonic regulatory elements and donor splice sites);

-the role of U1 snRNP complementarity in aberrant splicing pattern;

-the potential therapeutic effect of modified U1 snRNPs, targeted to intronic regions at a distance from the 5' ss, on mutation suppression and recovery of functional protein biosynthesis;

-the characterization of new intronic splicing silencer elements in *CFTR* exon 12 and *F9* exon 5 and their potential mechanism of action

2. Results

Section 2.1. **In silico *analysis of splice site strength of* CFTR *exon 12 and* F9 *exon 5 disease-associated variants.***

In order to evaluate the effect on splicing of genomic variants (GVs) I studied coagulation factor IX gene (*F9*) exon 5 and *c*ystic *f*ibrosis *t*ransmembrane *r*eceptor gene (*CFTR*) exon 12. I have evaluated ten different natural mutations associated with factor IX deficiency and eight mutations associated with cystic fibrosis as reported in the corresponding databases (Table 2.1.1).

In the case of *CFTR* exon 12, I focused on two exonic mutations, c.1696G>A (p.A566T) and c.1731C>T (p.Y577Y) and six mutations occurring within exon 12 donor splice site sequence: c.1766G>A, (p.S589N), c.1766G>T (p.S589I), c.1766+2T>C, c.1766+3A>C, c.1766+3>G and c.1766+5G>A. For simplicity's sake, I will refer to the exonic mutations according to their relative location to the CFTR exon 12 3' splice site: 17A and 52T, respectively. Similar criteria will be applied to the donor site mutations, which will be named according to their relative location to the 5' splice site (5' ss): -1A, -1T, +2T, +3C, +3G and +5A, respectively. Exonic mutations 17A and 52T were chosen as examples of mutations that affect exonic splicing regulatory elements. In fact, they are located in two previously described regulatory elements in *CFTR* exon 12, the Composite Exonic Regulatory Element of Splicing (CERES) 1 and 2, respectively (Pagani et al. 2003a, 2005). 52T is a synonymous mutation found in a patient with a severe cystic fibrosis phenotype, while 17A is a missense mutation found in an individual with mild CF (Table 2.1.1; *CFTR* Mutation Database). Alternatively, five mutations located at the *CFTR* exon 12 donor splice site have been associated to severe CF phenotype and one (+3G) to mild form of the disease.

Concerning *F9* exon 5, I have focused on mutations within polypyrimidine

(PPT) and donor splice site sequences. Specifically, within the PPT region I studied the c.17660T>G and c.17661T>G mutations, located upstream to the acceptor site of *F9* exon 5 at positions -9 and -8 respectively. These mutations will be further referred to as -9G and -8G. Both of them are found in patients with mild haemophilia B (Montejo et al. 1999), Haemophilia B mutation database). Within the donor splice site sequence, the mutations investigated consist of three synonymous variants at position c.17896 (A>G, A>C and A>T), in addition to the variants c.17897G>T, c.17897+1G>A, c.17897+2T>C, c.17897+4A>G and c.17897+13A>G (Table 1). These donor site mutations, according to their relative position to the 5' ss, will be further referred to as -2G, -2C, -2T, -1T, +1A, +2T, +4G and +13G. Concerning the pathological relevance of these variants, individuals carrying mutations in positions -2, -1T, +1A and +2C are affected with severe haemophilia B, mutation +4G has been found in patients with moderate phenotype and mutation +13G on individuals suffering from mild to moderate phenotype (Haemophilia B Mutation Database). To our knowledge these mutations were not investigated for their impact on *F9* exon 5 aberrant splicing.

To understand the effect of these mutations on exon processing, I first analysed the strength of normal and mutated splice sites in terms of consensus values by using several *in silico* methods and the results are shown in Table 2.1.2. Analysis with *in silico* programs shows that the scores of *CFTR* exon 12 wild-type (*wt*) donor site are not very different from the optimal consensus (CV=1=100%). In contrast, the scores of *F9* exon 5 *wt* were significantly reduced (Table 2.1.2). For example, considering the NN tool the *CFTR* exon 12 *wt* has a score of 0.99 whereas the *F9* exon 5 *wt* has only 0.21. Sequence inspection showed that both *wt* splice sites deviate from the consensus and have some reduced complementarity to the U1 snRNA. The *CFTR* exon 12 *wt* donor site shows two mismatches at position -3 (an A) and at position +4 (a T). The *F9* exon 5 *wt* deviates at three positions: it has a C at position 3, a T at position 5

and an A at position 6 (Fig. 2.3.1A and 3.4A). The scores for both *CFTR* exon 12 and *F9* exon 5 donor site mutations exhibit a great variability and based on this analysis it is difficult to clearly predict the effect of some mutations on splicing. This is particularly evident for *F9* exon 5, as the *wt* donor site itself has a very low score, and accordingly should not be used. Considering the mutations that abolish the nearly invariant GU dinucleotide (conserved in almost ~99.24% of human donor sites) (Burset et al. 2000), it is expected for the prediction tools not to recognize the donor splice sites at all (some of them are built assuming only U2 dependent GU donor sites; see Materials & Methods) or to assign them the lowest scores (CV=0). In accordance with these expectations, 10 out of 11 prediction tools used report null score for the +2C mutations in *CFTR* exon 12, (see Table 2.1.2). Conversely, calculation of free energy (ΔG) confers it a score similar to non-GU mutations as this tool do not give any preference to nucleotide position. Similarly, 7 out of the 11 prediction tools report the lowest scores or were unable to recognize any donor site for the +1A and +2C mutations in *F9* exon 5.

The interpretation of score predictions for mutations +3G in *CFTR* exon 12 and +4G in *F9* exon 5 donor sites can be puzzling. In fact, concerning the +3G mutant in *CFTR* exon 12, 7 out of 11 bioinformatic tools exhibit score values similar to the corresponding *wt* donor site. The seven prediction tools that suggest no significant effect on splicing include the most used programs like Neural Networks (NN), Shapiro and Senapathy (S&S), maximum entropy (MaxEnt) and weight matrix models (WMM). On the other hand, the mutation +4G on *F9* exon 5 donor site significantly reduces the score for 8 out of 10 bioinformatic tools, including NN, MaxEnt, S&S or WMM.

Mutation +13G of *F9* exon 5 occurs downstream the donor site but, according to sequence change and all prediction tools used, it creates a cryptic donor site with higher (S&S, MaxEnt, MDD, MM, WMM, NN) or equal (AST, ΔG, SplicePort, Hbond, HSF) strength of the natural one.

Mutation	Sequence	AA, change	Reported clinical features	Reference
CFTR exon 12				
Exonic mutations				
wt	AAAGATG			
17A	AAAGATA	566, Ala to Thr	Mild CF disease, pancreatic sufficiency	*CFTR* Mutation Database
wt	GGATAC	Healthy individuals		
52T	GGATAT	577, Tyr to Tyr	Severe CF	*CFTR* Mutation Database
Donor site mutations				
wt	AAGgtatgt			
-1A	AAAgtatgt	589, Ser to Asn	Severe CF	(Stanziale et al. 2005; Mérelle et al. 2006)
-1T	AATgtatgt	589, Ser to Ile	Severe CF	*CFTR* Mutation Database
+2C	AAGgcatgt	NA	Severe CF, steatorrhea; pancreatic insufficiency	(Choe et al. 2010)
+3C	AAGgtctgt	NA	Severe CF	(Kinnunen et al. 2005)
+3G	AAGgtgtgt	NA	Mild CF, pancreatic sufficiency	(Cremonesi et al. 1992)
+5A	AAGgtatat	NA	Severe CF	*CFTR* Mutation Database
F9 exon 5				
Polypyrimidine tract mutations				
wt	tgcttcttttagATG			
-8G	tgctgcttttagATG	NA	Mild severity	Haemophilia B Mutation Database
-9G	tgcgtcttttagATG	NA	Mild severity	(Montejo et al. 1999)
Donor site mutations				
wt	CAGgtcata			
-2T	CTGgtcata	127, Ala to Ala	Severe	Haemophilia B Mutation Database
-2C	CCGgtcata	127, Ala to Ala	Severe	
-2G	CGGgtcata	127, Ala to Ala	Severe	
-1T	CATgtcata	128, Val to Leu	Severe	
+1A	CAGatcata	NA	Severe	
+2C	CAGgccata	NA	Severe	
+4G	CAGgtcgta	NA	Moderate	
wt	tgaataaga			
+13G	tgagtaaga	NA	Mild-moderate	(Koeberl et al. 1990; Ketterling et al. 1999)

Table 2.1.1. **Genotype-phenotype correlation of genomic variants of *CFTR* exon 12 and *F9* exon 5 analyzed in this thesis.**
Severity of cystic fibrosis (CF) in patients is determined for each patient considering presence of specific phenotypic clinical components and their time course (Zielenski 2000) For coagulation factor IX deficiency, the clinical severity is determined by the plasma level of factor IX: severe, <0.01 IU/ml; moderate, 0.01-0.5 IU/ml; mild, >0.05 - <0.40 IU/ml (Haemophilia B database: http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html). Polypyrimidine tract and donor site mutations are named after their position with respect to 3' and 5'ss, respectively; cDNA names are described in Section 2.1. (-) or (+) denotes position from exon/intron junction. Uppercase letters represent exonic sequences and lowercase letters indicate the intronic sequences. Natural mutations are coloured. NA, not applicable.

Finally, the PPT transversions -8G and -9G shows a mild reduction of the *F9* exon 5 3' splice site (3' ss) score with all the prediction tools.

	Sequence	S&S	MaxEnt	MDD	MM	WMM	NN	AST	ΔG	SplicePort	Hbond	HSF	Base Pairs
consensus	CAGgtragt	100	10.86	15.08	12.75	13.07	1	100	-9.6	----------	23.8	100	9
CFTR exon 12													
Donor site mutations													
wt	AAGgtatgt	88.7	9.79	13.48	10.03	9.76	0.99	89.78	-5.4	1.17	15.8	90.24	8
-1A	AAAgtatgt	76.3	6.93	9.28	6.35	6.59	0.48	77.86	-1.7	-0.02	11.1	79.66	7
-1T	AATgtatgt	75.9	5.87	8.78	5.32	5.64	0.23	77.26	-0.6	-0.4	11.1	79.37	7
+2C	AAGgcatgt	70.4	2.03	5.72	2.28	2.01	ND	72.66	-1.3	ND	ND	ND	7
+3C	AAGgtctgt	78.6	6.43	10.58	5.57	5.24	0.2	79.63	-3.7	0.12	11.5	85.22	7
+3G	AAGgtgtgt	85.4	6.64	13.28	6.72	8.38	0.84	84.81	-5.1	0.47	12.4	89.08	8
+5A	AAGgtatat	74.3	7.84	11.08	6.68	6.31	0.46	77.93	-4.1	0.27	12.3	78.08	7
F9 exon 5													
Polypyrimidine tract mutations													
consensus	ttttctttccagG	100	12.05	------	13.58	17.73	1	100	---	---------	---------	100	------
wt	tgcttctttagA	91.4	9.23	------	11.07	12.29	0.94	86.42	---	ND	---------	82.09	------
-8G	tgctgcttttagA	88.6	8.15	------	10.1	10.28	0.85	83.86	---	ND	---------	80.97	------
-9G	tgcgtcttttagA	88.3	9.04	------	10.04	10.19	0.86	83.39	----	ND	---------	79.8	------
Donor site mutations													
wt	CAGgtcata	69.9	3	7.98	3.54	3.31	0.21	72.54	-3.3	-0.54	11.9	80.32	6
-2T	CTGgtcata	62	-2.79	3.98	1.03	1.27	0	64	-1	-1.33	8.3	75.77	6
-2C	CCGgtcata	61.7	-1.46	3.68	2.3	0.96	0	63.33	-1.3	-1.13	8.3	75.57	5
-2G	CGGgtcata	62	-1.86	4.68	1.79	1.03	0.01	63.38	-3.1	-1.29	8.5	75.47	6
-1T	CATgtcata	57.1	-11.52	-3.92	-5.04	-0.81	0	60.02	1.5	-1.54	4	69.46	5
+1A	CAGatcata	51.6	-5.18	-0.2	-4.64	-4.88	0	55.42	1.1	ND	ND	ND	5
+2C	CAGgccata	51.6	-4.75	0.22	-4.21	-4.45	0	55.42	-1.7	ND	ND	ND	5
+4G	CAGgtcgta	59.1	-0.49	5.68	0.92	0.93	0.01	62.4	-4.4	-0.64	11.8	71.98	5
+13G	tgagtaaga	74.1	5.11	8.28	5.18	4.25	0.41	69.59	-2.8	-0.53	12	80.25	5

Table 2.1.2. **Bioinformatic predictions of score values for *CFTR* exon 12 donor splice sites and *F9* exon 5 acceptor and donor splice sites.**
Predictions were always made using default settings. The last column denotes the number of putative base pairs between U1 snRNP and donor site sequence should be formed according to 5' ss sequence. Polypyrimidine tract (PPT) and donor site mutations are named after their position with respect to 3' and 5' ss, respectively; cDNA names are described in Section 2.1.(-) or (+) denotes position from exon/intron junctios. Uppercase letters represent exonic sequences and lowercase letters indicate the intronic sequences. Natural mutations are coloured. ND, not detected. S&S, consensus value based on sequence composition of splice sites (Shapiro and Senapathy, 1987); MaxEnt, based on maximum entropy principle; MDD, maximal dependance decomposition; MM, inhomogenous first-order Markov models; WMM, weight matrix models (Yeo & Burge, 2004); NN, neural networks (Reese et al 1997); AST, Analyzer Splice Tool; ΔG, free energy calculation of U1 snRNA and donor site sequence base pairing (Carmel et al 2004); SplicePort (Dogan et al 2007); Hbond, estimation of H-bonds formed between U1 snRNA and donor site sequence (Freund et al 2003); HSF, Human Splice Finder (Desmet et al 2009). Haemophilia B database is available at http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html. *CFTR* Mutation Database can be consulted at http://www.genet.sickkids.on.ca/app.

Section 2.2. ***Effect of* CFTR *exon 12 and* F9 *exon 5 variants on pre-mRNA splicing in hybrid minigenes.***

In order to assess whether the mutations listed in Table 2.1.1 impair pre-mRNA processing, I performed a splicing functional assays. For this purpose I used a modified version of the previously described pTB hybrid minigene that has been reported to reproduce the *in vivo* splicing pattern of several gene systems (Pagani et al. 2000, 2002, 2003a; Amaral et al. 2004; Baralle et al. 2003). The minigene I engineered is a derivative of pTB, which is a modified version of the α-globin-fibronectin-EDB (Extra Domain B) minigene. It contains the exons 1, 2 and 3 of the α-globin and part of the fibronectin gene. Its transcription is under the control of a minimal α-globin promoter and the SV40 enhancer (Pagani et al. 2003b). Both *CFTR* exon 12 and *F9* exon 5 along with part of their intronic flanking region (IVS 11 and IVS 12, and IVS4 and IVS5, respectively) were cloned using a unique *Nde*I restriction site, which is located in the large fibronectin intron (see Materials & Methods). HeLa cells were transfected with the *wt* and mutated forms of *CFTR* (Fig. 2.2.1A) or *F9* (Fig.2.3A) minigenes and the spliced mRNAs were analysed through RT-PCR using alpha2,3 and bra2 oligonucleotides specific for the minigene sequence therefore excluding the possibility of amplifying endogenous mRNA. The PCR products detected upon transfection of the *wt* minigenes showed the presence of two mRNA bands: the upper one corresponds to exon inclusion (*CFTR* exon 12: 333 bp; *F9* exon 5: 375 bp) and the lower one to exon exclusion (247 bp). The identity of the bands was verified by direct sequencing after elution of the bands from the gel. Quantitative analysis of band intensity showed that *CFTR* exon 12 *wt* is mostly included in mature mRNA, with a mean value of ~80% of exon inclusion (Fig. 2.2.1B and C, lane 1). The low level of *CFTR* exon 12 skipping in this minigene system is similar to the exon skipping observed *in vivo* in normal bronchial epithelial tissue as previously reported (Pagani et al. 2003a, 2005). Exonic

mutations 17A and 52T reduced exon 12 inclusion to ~11% and ~8.6%, respectively (Fig. 2.2.1B and C, lanes 2 and 3), in agreement with previous reports (Pagani et al. 2003a, 2005). Interestingly, all the seven donor splice site mutations analysed induced nearly complete exon skipping (Fig. 2.2.1B and C, lanes 4-9), indicating their potential pathological effect and involvement in the disease pathogenesis.

In the case of *F9* exon 5, I first compared the pFIX ex5 *wt* minigene splicing pattern with the splicing pattern in liver tissue. As shown in Fig. 2.2.2, in both the minigene and the liver exon 5 is not completely included, showing some degree of exon skipping in the mature mRNAs. These results indicated that *F9* exon 5 minigene successfully mimics the *in vivo* splicing pattern. *F9* exon 5 PPT mutations -8G and -9G and donor splice site mutations at positions -2, -1, +1 and +2 triggered exon skipping (Fig. 2.2.3B, C, lanes 2-7). The -2 mutants showed very low amount of exon inclusion (less than 3%) whereas the other mutants exhibited complete exon skipping. Surprisingly, and in contrast to the majority of the prediction programs, the mutation at position +4 displayed a normal splicing pattern, with a slight increase of the percentage of exon 5 inclusion if compared to the *wt* minigene (Fig 2.3B and C; compare lane 1 with lane 8). A possible explanation could be that this variation might create a novel cryptic GU donor site 3 bp downstream of the normal site. Therefore, to rule out this possibility, I performed direct sequencing analysis and detected the correct usage of the natural donor site. Thus, this mutation did not induce aberrant splicing, neither exon skipping nor formation of a cryptic 5' ss. Transfection of the +13G minigene apparently showed complete exon inclusion (Fig. 2.2.3B, lane 9), however direct sequencing of the band and further analysis with FAM oligonucleotide on capillary electrophoresis indicated that the mutation activates the usage of a cryptic 5' ss 13 bp downstream the normal 5' ss (see below and Section 2.6, Fig. 2.6.2). All *in silico* tools have assigned to the cryptic 5' ss a score value similar to that of the natural *F9* exon 5 donor

splice site or higher (Table 2.1.2).

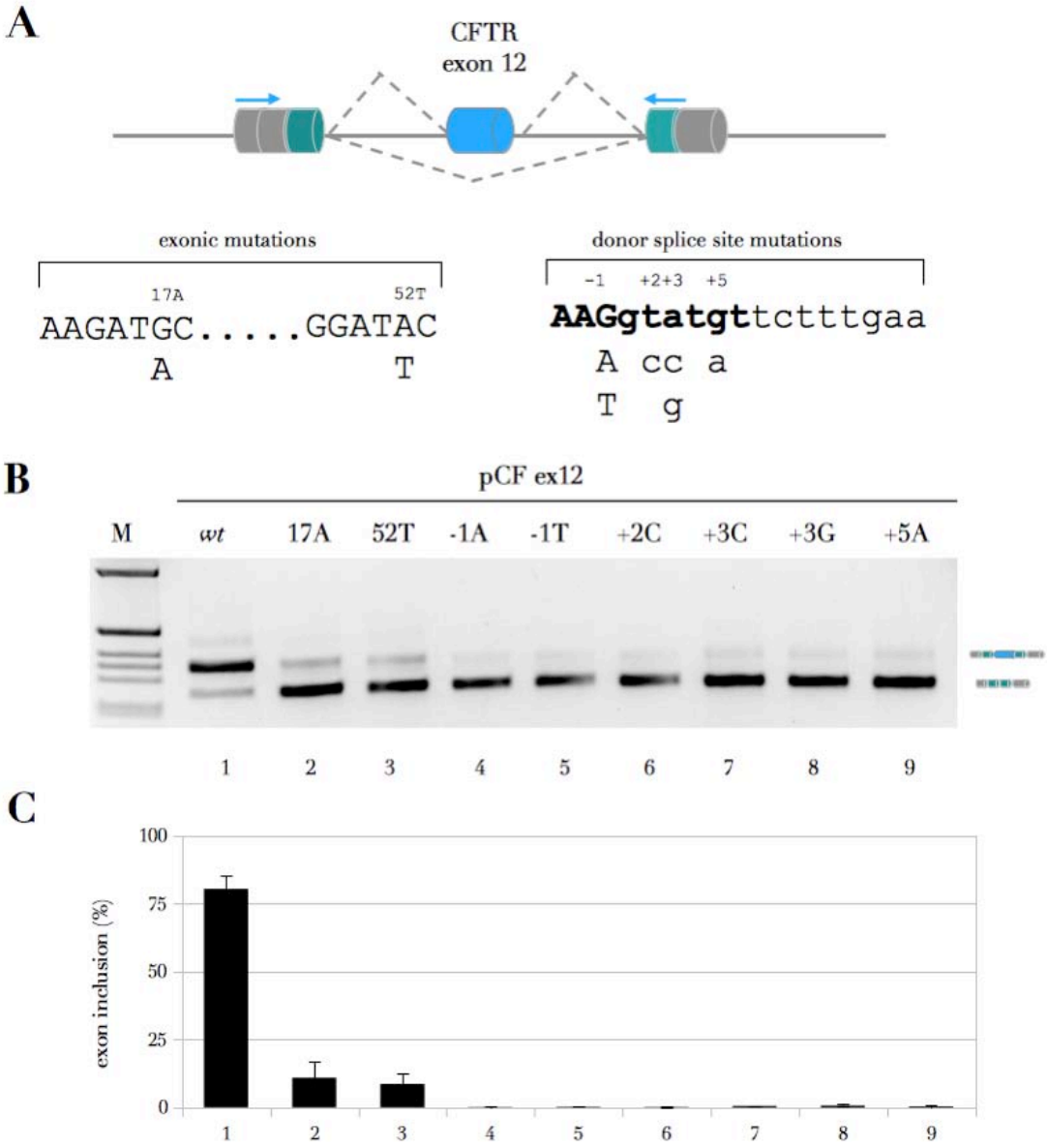

Figure 2.2.1. **Effect of exonic and donor site mutations on splicing of *CFTR* exon 12.**
A, Schematic representation of pCF ex12 minigene. Dashed lines (- - -) illustrate possible splicing outcome. Arrows represent primers for PCR amplification. For exonic mutations, sequence variation (below wild-type sequence) and nucleotide position (with respect to 3'ss; above) are shown; for donor splice site mutations, sequence variation (below wild-type sequence) and position (above) are shown. (-) or (+) denotes position from exon/intron junction. Uppercase letters, exonic sequences; lowercase, intronic sequences. The donor splice site sequence is bolded.
B, Analysis of spliced transcripts. HeLa cells were transfected with 0.5 µg of pCF ex12 *wt* or mutant minigenes and splicing pattern evaluated by RT-PCR with alpha2,3 and bra2 primers. Amplified products were resolved on a 2% agarose gel. The identity of the bands is indicated on the right side of the panel. M, molecular weight marker (1kB, Invitrogen).
C, Quantification of exon 12 inclusion of exonic and 5' splice site mutations. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

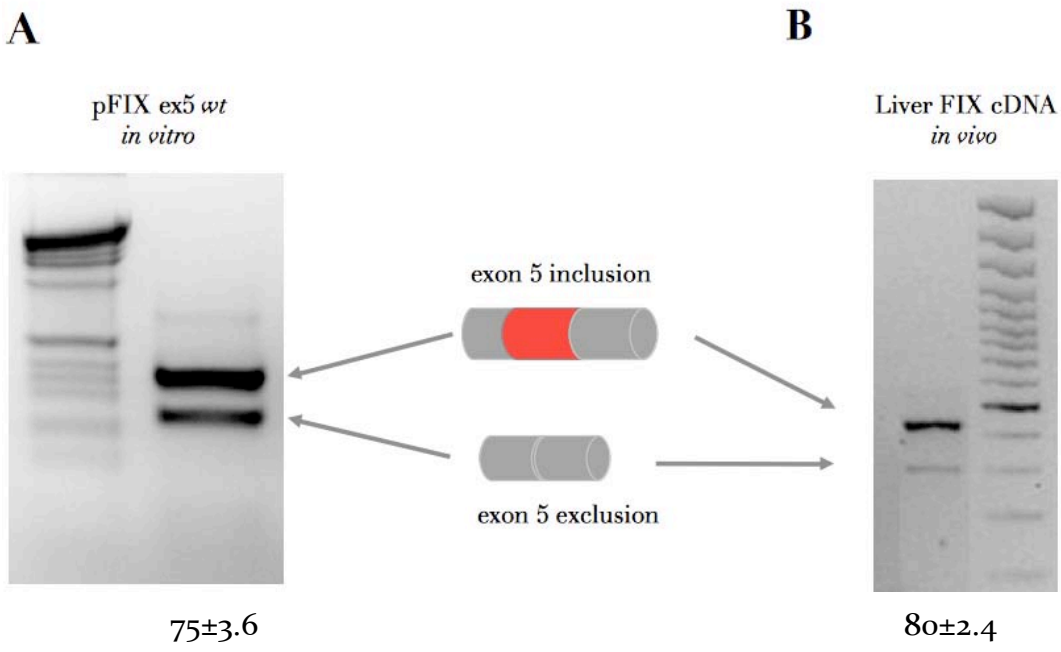

Figure 2.2.2. **Comparison between the levels of *F9* exon 5 splicing *in vivo* and using a hybrid-minigene system.**
A, Hela cells were transfected with 0.5 µg of pFIX ex5 *wt* minigene and splicing pattern evaluated by RT-PCR with alpha2,3 and bra2 primers. Details of the minigene are shown in Fig. 2.2.3A. Amplified products were run on 2% agarose gel and the analysis of PCR products showed 2 bands: the upper one corresponds to exon 5 inclusion and the lower one to exon 5 exclusion. The percentage of exon 5 inclusion (±SD) is indicated.
B, Human liver *F9* cDNA *in vivo* amplified using FIX140 and FIX279 primers.

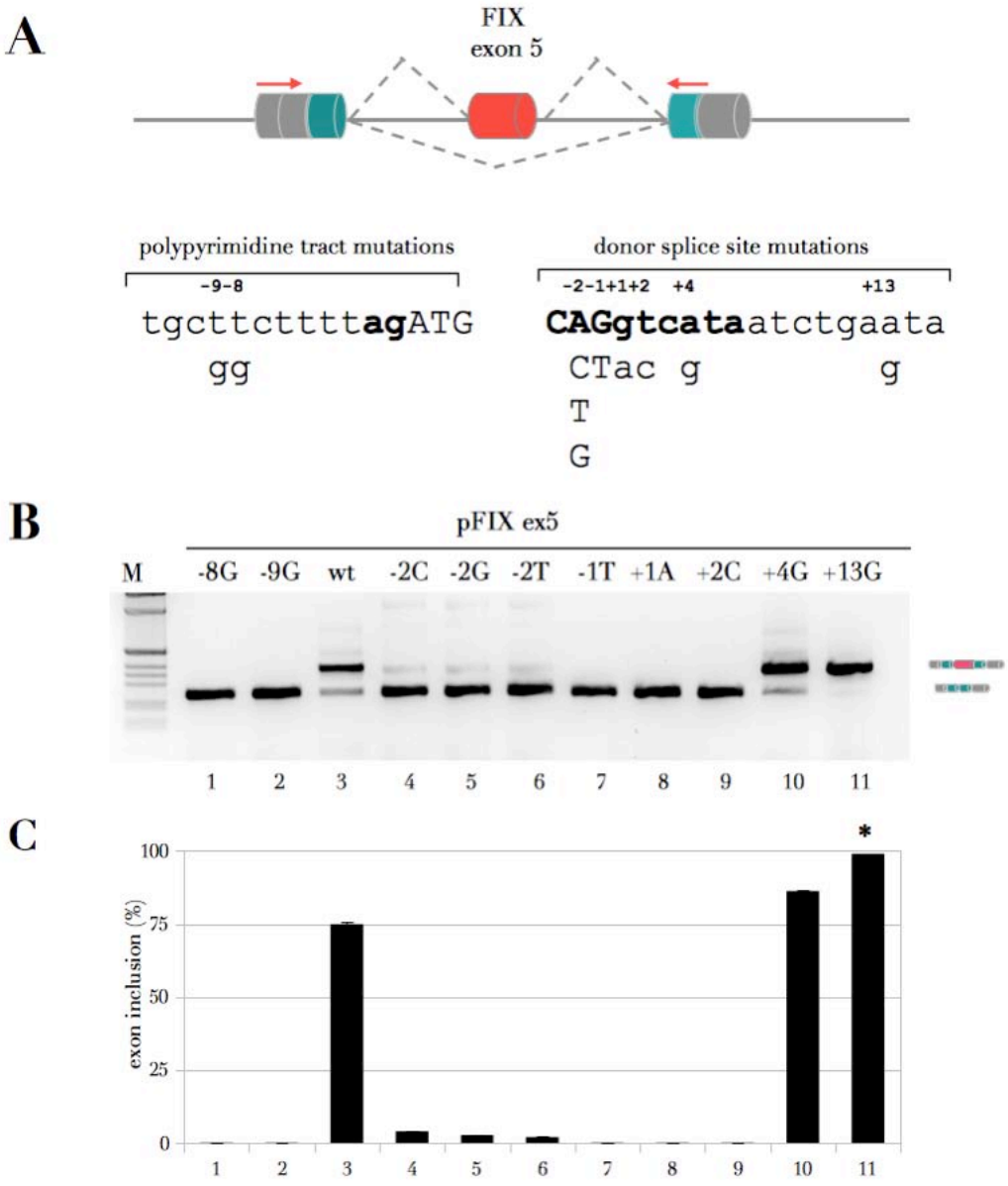

***Figure 2.2.3*. Effect of polypyrimidine tract and donor splice site mutations on splicing of *F9* exon 5.**

A, Schematic representation of pFIX ex5 minigene. Dashed lines (- - -) illustrate possible splicing outcome. Arrows represent primers for PCR amplification. Sequence variation (below wild-type sequence) and position (above) are shown; (-) or (+) denotes position from exon/intron junction. Uppercase letters, exonic sequences; lowercase, intronic sequences. Factor IX exon 5 donor splice site sequence and acceptor site are bolded.

B, Analysis of spliced transcripts. HeLa cells were transfected with 0.5 µg of pFIX ex5 *wt* or mutant minigenes and splicing pattern evaluated by RT-PCR with alpha2,3 and bra2 primers. Amplified products were resolved on a 2% agarose gel. The identity of the bands is indicated on the right side of the panel. M, molecular weight marker (1kB, Invitrogen).

C, Quantification of exon 5 inclusion of 5' splice site mutations. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate. *, cryptic site usage, see section 2.6.

Section 2.3. ***Effect on splicing of modified U1 snRNAs complementary to donor splice site mutations.***

The *wt* donor splice sites of *CFTR* exon 12 and *F9* exon 5 show several mismatches with the 5' end of U1 snRNA. In fact, *CFTR* exon 12 donor splice site has an uridine at position +4, which differs from consensus adenine and a G in position -3 (Fig. 2.3.1A). *F9* exon 5 donor splice site presents three mismatches: a C, a U and an A at positions +3, +5 and +6, respectively (Fig. 2.3.4A). To understand if 5' ss mutations impair their interaction with U1 snRNA, I have evaluated the role of U1 snRNA in exon recognition by introducing compensatory changes within the 5'-tail of U1 snRNA. I created several U1 snRNAs with perfect complementarity to each *wt* or mutated 5' ss sequence, as illustrated in Fig. 2.3.1 for *CFTR* exon 12 and in Fig. 2.3.4 for *F9* exon 5.

2.3.1. Effect of modified U1 snRNA on CFTR exon 12 splicing

In the first experiment, I evaluated the role of the natural *CFTR* exon 12 +4 A>T variation on splicing. This was done through co-transfection of U1CFwt: this U1 has an adenine complementary to the uridine present at position +4 of *wt* CF exon 12 donor site (Fig. 2.3.1B). U1CFwt was co-transfected with *CFTR* exon 12 *wt* or mutant minigenes as illustrated in Fig. 2.3.2. Co-transfection of U1CFwt with pCF ex12 *wt* improved exon inclusion. The percentage of exon inclusion increased from approximately 80% to 95% (Fig 3.2A and B, lane 1 and 3). This effect is specific for the compensatory changes introduced, since over-expression of normal U1 snRNAs (U1 wt) did not change the splicing pattern (Fig. 2.3.2A and B, lanes 1-3). U1CFwt also exhibited a positive effect on the two exonic mutations 17A and 52T, rising exon inclusion levels from 10% to ~32% and ~45%, respectively (Fig. 2.3.2A and B, lanes 4-7). On the contrary, the

complete exon skipping of the mutations -1A, -1T +3C and +5A, was not affected by co-transfection of U1CFwt (Fig 3.2, lanes 8-13 and 16-17). The only donor site mutant recovered in part by U1CFwt was the +3G mutant (Fig.3.2A and B, lane 15). In this case, the U1CFwt forms a non-Watson-Crick bond between the mutant guanidine in position +3 on the pre-mRNA and the corresponding uridine of the U1CFwt (Fig. 2.3.1A). This complementarity is possibly involved in the positive response observed. On the contrary, the other 5' ss variants maintained a mismatch between the corresponding mutation site on pre-mRNA and U1CFwt. In these cases the compensatory change in position +4 was not sufficient for splicing rescue. These data indicate that the compensatory change in position +4 of U1CFwt improved exon inclusion of the *wt* context, of the exonic mutations and only partially of the +3G variant.

Subsequently, I tested additional modified U1 snRNAs complementary to each 5' ss variants, as listed in Fig. 2.3.1C. These modified U1 snRNAs have the compensatory T at position +4 as in U1CFwt along with an additional nucleotide substitution to compensate each mutation. For example, U1CF -1A has two nucleotide substitutions in comparison to the normal U1 snRNA: an adenine (instead of the normal uridine) to complement the natural uridine at position +4 of *CFTR* exon 12 5' ss and an uridine to compensate the mutant adenine at position -1 (Fig. 2.3.1C). Co-transfection experiments showed that compensatory U1CF -1A, -1T, +3C, +3G and +5A enhanced exon inclusion of their corresponding mutant minigenes with different efficiency: to 65.7%, 18.2%, 33.3%, 73.2% and 10.1%, respectively (Fig. 2.3.3A and B, lanes 5, 7, 9, 11 and 13). Thus, these data indicate that, in *CFTR* exon 12, defective binding of U1 snRNA is involved in exon skipping induced by the -1A, -1T, +3C, +3G and +5A mutations.

2.3.2. Effect of modified U1 snRNAs on FIX exon 5 splicing

I tested a modified U1 snRNA fully complementary to the F9 exon 5 wt donor splice site. U1FIXwt carries three base changes within its 5'-tail which base pairs at positions +3, +5 and +6 of F9 exon 5 5' ss (Fig. 2.3.4A). Co-transfection of U1FIXwt with *F9* exon 5 *wt* minigene improved exon inclusion, from 75% to nearly complete inclusion. Over-expression of U1 wt did not change *F9* exon 5 *wt* splicing pattern (Fig. 2.3.5A and B, lanes 1-3). Co-transfection of U1FIXwt with -2 mutant minigenes also led to complete exon inclusion, regardless of the base change under analysis (Fig. 2.3.5 B and C, lanes 3-8). In contrast, U1FIXwt did not revert exon skipping caused by mutations -1T, +1A and +2C (Fig. 2.3.5 B and C, lanes 10-15). Furthermore, I also explored the effect of U1FIXwt on exon 5 skipping caused by the PPT mutations -8G and -9G. As shown in Fig. 2.3.5C and D, U1FIXwt improved exon inclusion greatly, reaching ~75% for the PPT mutant minigenes. Afterwards, I engineered additional modified U1 snRNAs, complementary to each *F9* exon 5 donor site mutation (Fig. 2.3.4C). Upon co-transfection, U1FIX -2C, -2G and -2T rescued exon inclusion of their corresponding mutant minigenes (Fig. 2.3.6, lanes 3-8), whereas U1FIX -1T, +1A and +2C failed to affect the splicing pattern (Fig 3.6, lanes 9-15). The exon skipping observed for the three mutants in position -2 was rescued by the corresponding compensatory U1 snRNAs (-2C, -2G, -2T), which are fully complementary to the mutant sites and also by U1FIXwt, which has a mismatch in position 2 (Fig. 2.3.5 and Fig. 2.3.6, lanes 3-9). Thus, I investigated which positions of U1 snRNA might be preferentially required to rescue the -2 mutants. In order to do so, I created two new modified U1 snRNAs: U1FIX 3C to compensate the position 3, and U1FIXA5T6 to compensate the positions +5 and +6 (Fig. 2.3.7A). Co-transfection experiments with -2 mutant minigenes showed that U1FIX 3C promoted exon inclusion (Fig. 2.3.7 B and C, lanes 6-9). Conversely, U1FIX A5T6 did not increase exon 5

inclusion of these mutants (Fig. 2.3.7B and C, lanes 10-12).Although U1FIX 3C was capable of promoting exon inclusion in -2 mutant context, its effect was not as efficient as the one of U1FIXwt (compare Fig. 2.3.5A and B. lanes 4-9 with Fig. 2.3.7A and B, lanes 6-9). Therefore, substitutions A and T at positions +5 and +6 are by themselves inactive, although required along with substitution 3C in order to achieve full exon inclusion in -2 mutants.

Taken together, these results revealed that loading of U1 snRNAs complementary to the mutated donor site can recover several exon skipping-causing mutations. This was particularly effective for the exonic mutations (17A and 52T) in *CFTR* exon 12 and for the PPT mutations (-8G and -9G) in *F9* exon 5. In the case of donor site mutants, the rescue efficiency depends on the type of mutation. In fact, exon skipping caused by donor site mutations -1A, -1T, +3C, +3G and +5A within *CFTR* exon 12 and mutations -2C, -2G and -2T within *F9* exon 5 were corrected (although with variable efficacy) while *F9* exon 5 -1T, +1A and +2C failed to be rescued.

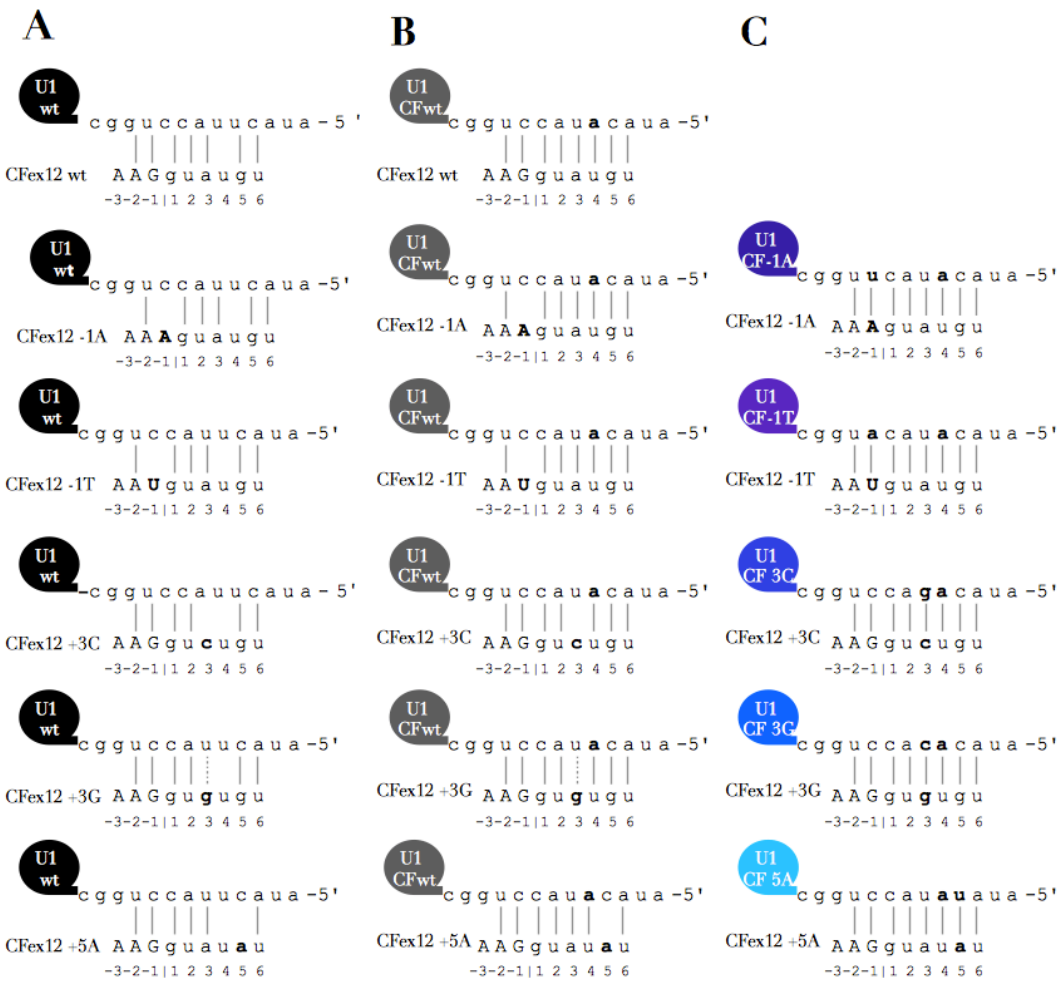

Figure. 2.3.1. **Complementary between modified U1-snRNAs and *CFTR* exon 12 donor site mutants.**

A, Base pairing of *wt* and mutant *CFTR* exon 12 donor sites with normal U1 snRNA (U1wt).

B, Base pairing of *CFTR* exon 12 *wt* and mutant donor sites with U1CFwt. U1CFwt bears an adenine (bold, U1 5'-tail) which base pairs with a guanidine present at position +4 of *wt CFTR* exon 12 donor site. This substitution allows the formation of a Watson-Crick bond.

C, Complementarity of mutant *CFTR* exon 12 donor sites with corresponding compensatory U1 snRNA. Each compensatory U1 contains an adenine at position +4 plus an additional change (bold, U1 5'-tail) for each *CFTR* exon 12 donor site mutation (bold, donor site sequence). Straight line, Watson-Crick bond; Dashed line, non Watson-Crick bond. Bolded bases within U1 5' end represent introduced changes within normal U1 snRNA 5' end: bolded bases within donor site sequences represent mutations.

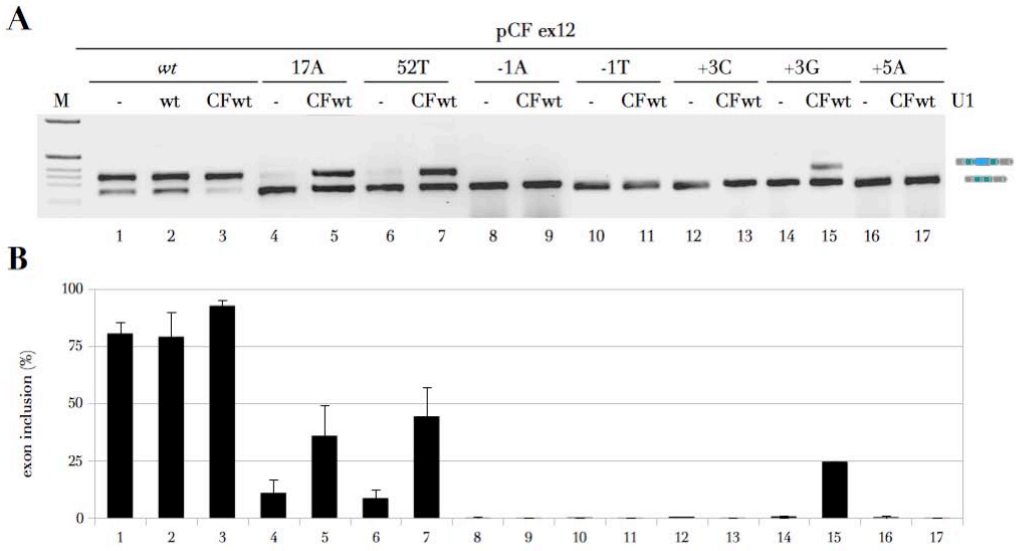

Figure 2.3.2. **Effect of U1CFwt on *CFTR* exon 12 exonic and donor site mutants**
A, Splicing pattern analysis. *CFTR* exon 12 *wt*, exonic (17A and 52T) and donor site mutant minigenes were transiently transfected into HeLa cells and the total RNA analyzed by RT-PCR using alpha2,3 and bra2 primers. PCR amplification products were separated on 2% agarose gel and the two splicing possibilities are indicated on the right of the gel. U1CFwt is able to correct exon skipping due to exonic mutations 17A and 52T and donor site mutation +3G, but is ineffective in other donor site mutant contexts.
B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

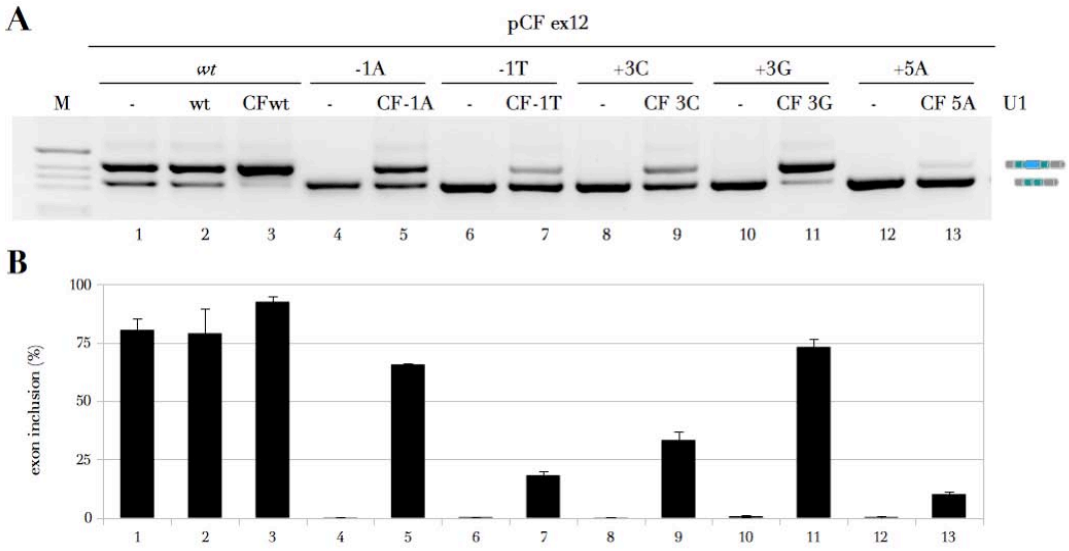

Figure 2.3.3. **Effects of the compensatory U1 snRNAs on *CFTR* exon 12 donor site mutants**
A, Splicing pattern analysis. The minigenes were transiently transfected into HeLa cells and the total RNA analyzed by RT-PCR using alpha2,3 and bra2 primers. PCR amplification products were separated on 2% agarose gel and the two splicing possibilities are indicated on the right of the gel. Splicing defects caused by all donor site mutations are corrected in presence of a compensatory U1 snRNA, although with variable efficacy .
B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

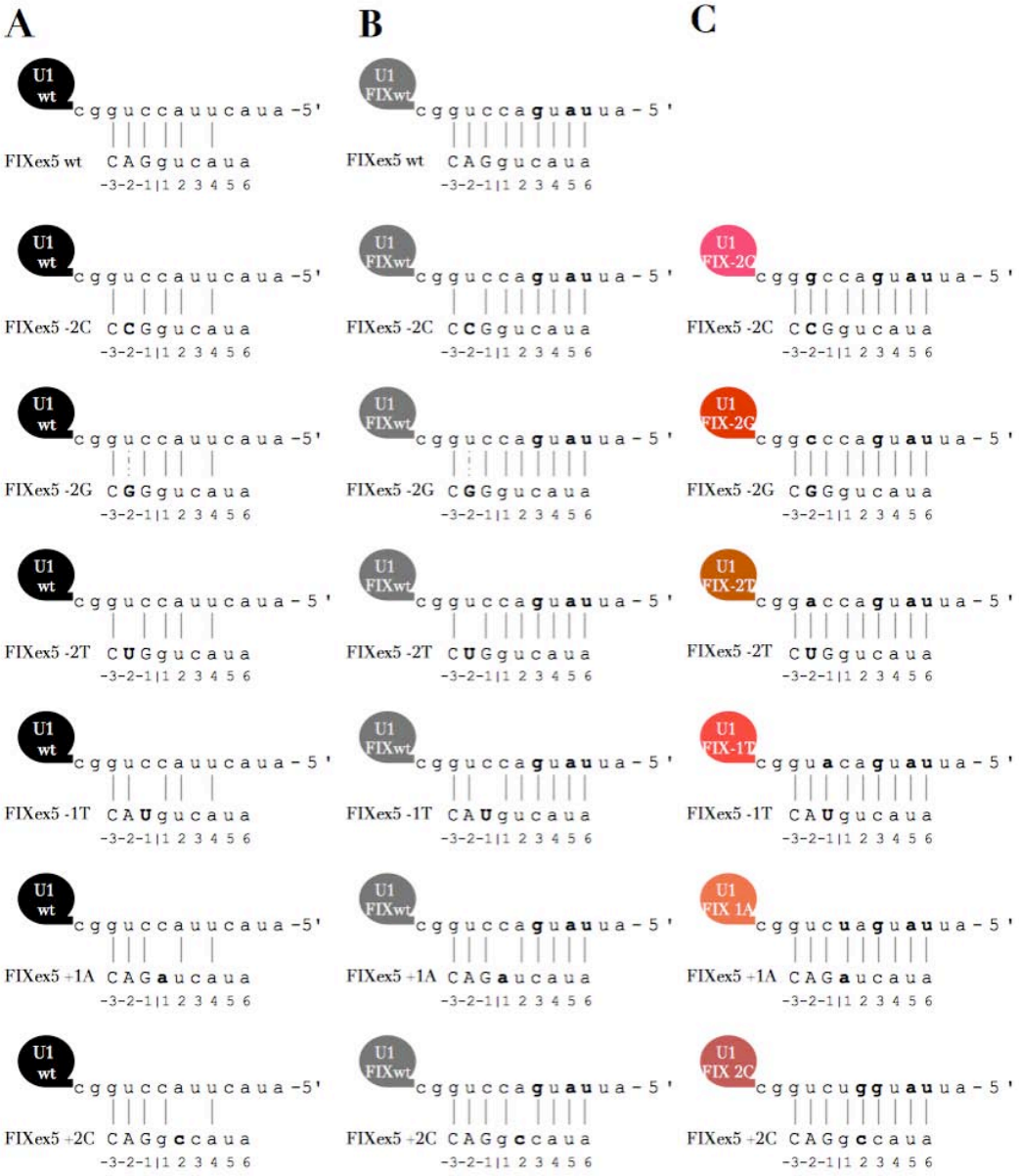

Figure 2.3.4. **Complementary between modified U1 snRNAs and *F9* exon 5 donor site mutants.**
A, Base pairing of *F9* exon 5 *wt* and mutant donor sites with normal U1 snRNA (U1 wt).
B, Base pairing of *F9* exon 5 *wt* and mutant donor sites with U1FIXwt. U1FIXwt was created by introducing three substitutions (a cytosine, an adenine and a thymidine; bold, U1 5'-tail) within normal U1 snRNA to increase base pairing to *F9* exon 5 *wt* donor site. These substitutions allow the formation of Watson-Crick bonds.
C, Complementarity of *F9* exon 5 mutant donor sites with corresponding compensatory U1 snRNA. Each compensatory U1
contains the aforementioned substitutions for U1FIXwt plus an additional change (bold, U1 5' end) for each *F9* exon 5 donor site mutation (bold, donor site sequence).
Straight line, Watson-Crick bond; Dashed line, non Watson-Crick bond. Bolded bases within U1 5'-tail represent introduced changes within normal U1 snRNA 5' end: bolded bases within donor site sequences represent mutations.

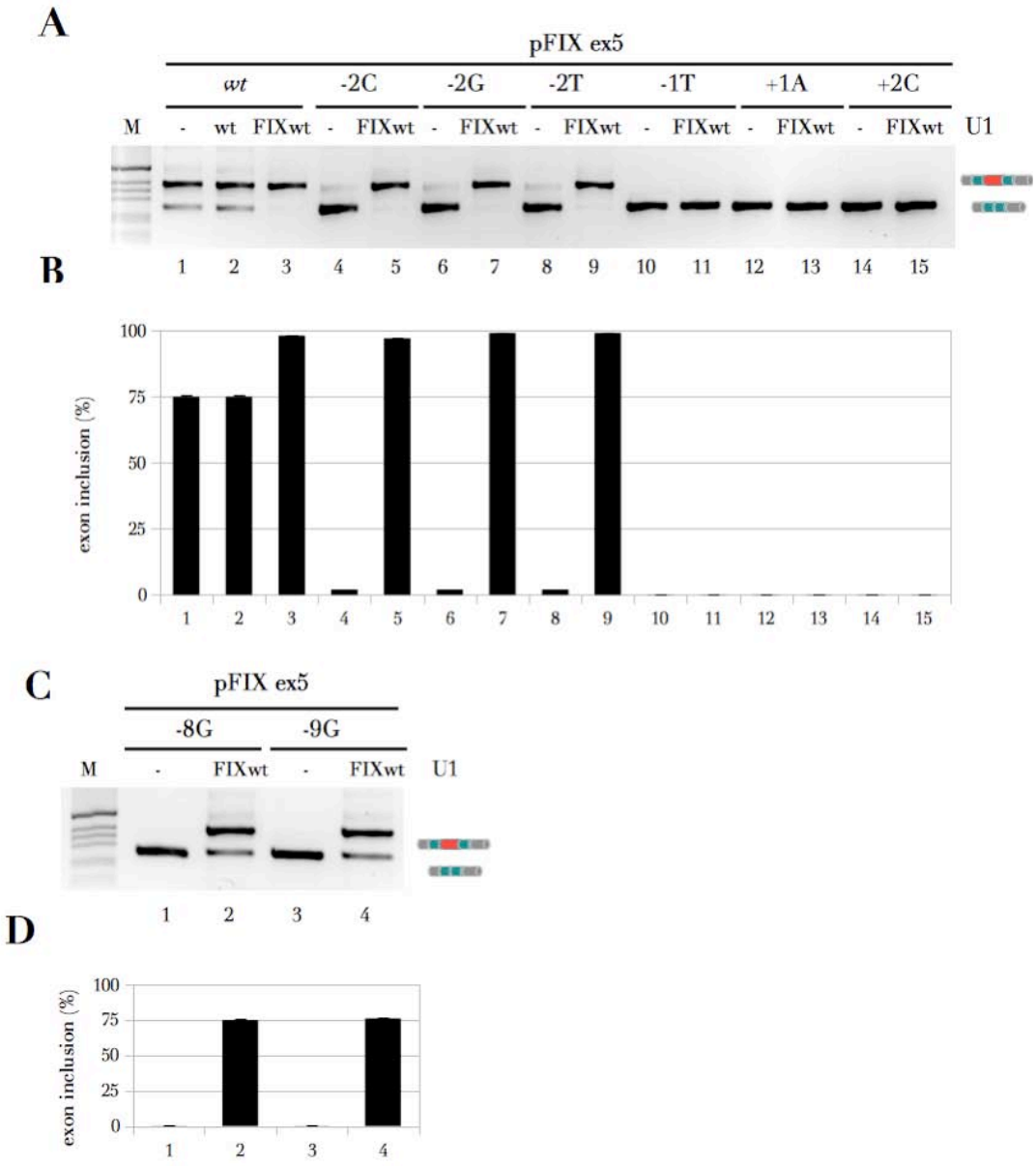

***Figure 2.3.5. F9* exon 5 mutant minigenes co-transfected with U1FIXwt.**

A, Splicing pattern analysis of the RT-PCR products derived from RNA of transfected HeLa cells, separated on 2% agarose gel. *wt* and donor splice site mutant *F9* exon 5 minigenes were transfected alone (0.5 µg) or with U1FIXwt (0.5 µg) and the splicing pattern evaluated with alpha2,3 and bra2 primers. pFIX ex5 *wt* was also transfected with U1 wt as a control (lane 2). U1FIXwt is able to correct exon skipping caused by mutations at position -2, but it is inactive in other donor site mutant contexts.

B, Quantification of splicing pattern analysis in A. Exon 5 inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate

C, Splicing pattern analysis of RT-PCR products derived from *F9* exon 5 polypyrimidine tract (PPT) mutant minigenes transfected either alone or with U1FIXwt, as in B. U1FIXwt is able to correct exon skipping caused by the PPT mutations -8G and -9G, rising *F9* exon 5 inclusion levels greatly.

D, Quantification of splicing pattern analysis in C. Exon 5 inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate

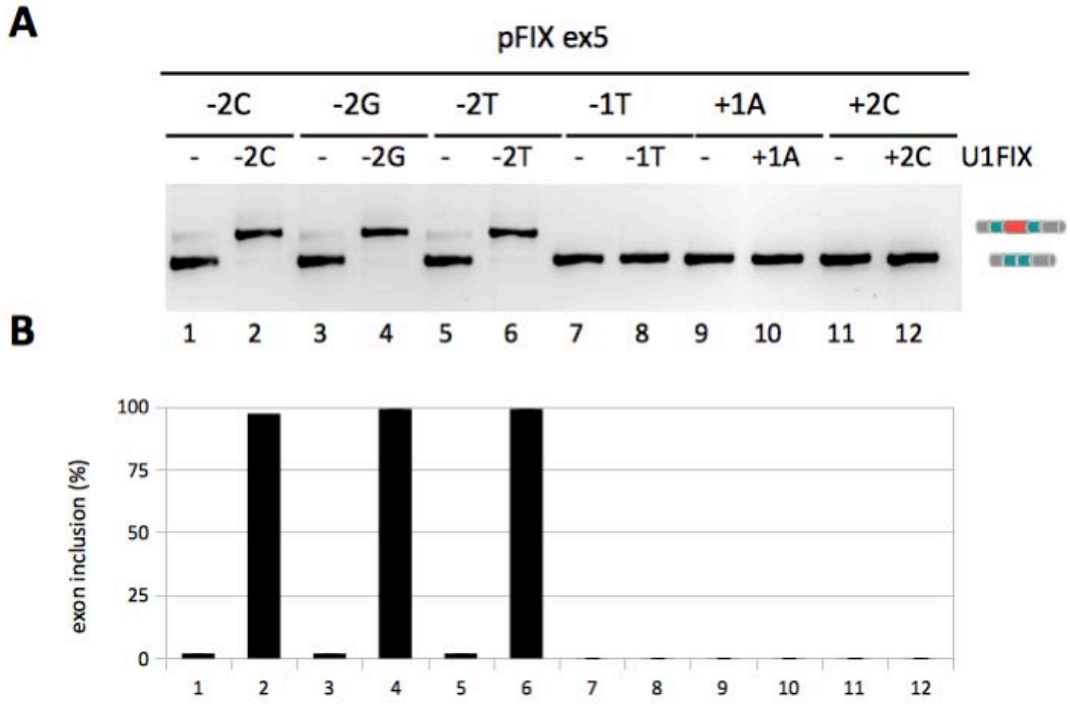

***Figure 2.3.6. F9* exon 5 donor site mutant minigenes co-transfected with compensatory U1FIX snRNAs.**

A, Splicing pattern analysis of the RT-PCR products derived from RNA of transfected HeLa cells, separated on 2% agarose gel. Donor splice site mutant *F9* exon 5 minigenes were transfected alone (0.5 µg) or with their corresponding compensatory U1FIX (0.5 µg) and the splicing pattern evaluated with alpha2,3 and bra2 primer. Compensatory U1FIX for mutations at position -2 were able to to correct exon skipping. Compensatory U1FIX for mutations -1T, +1A and +2C did not change splicing pattern.

B, Quantification of splicing pattern analysis in A. Exon 5 inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

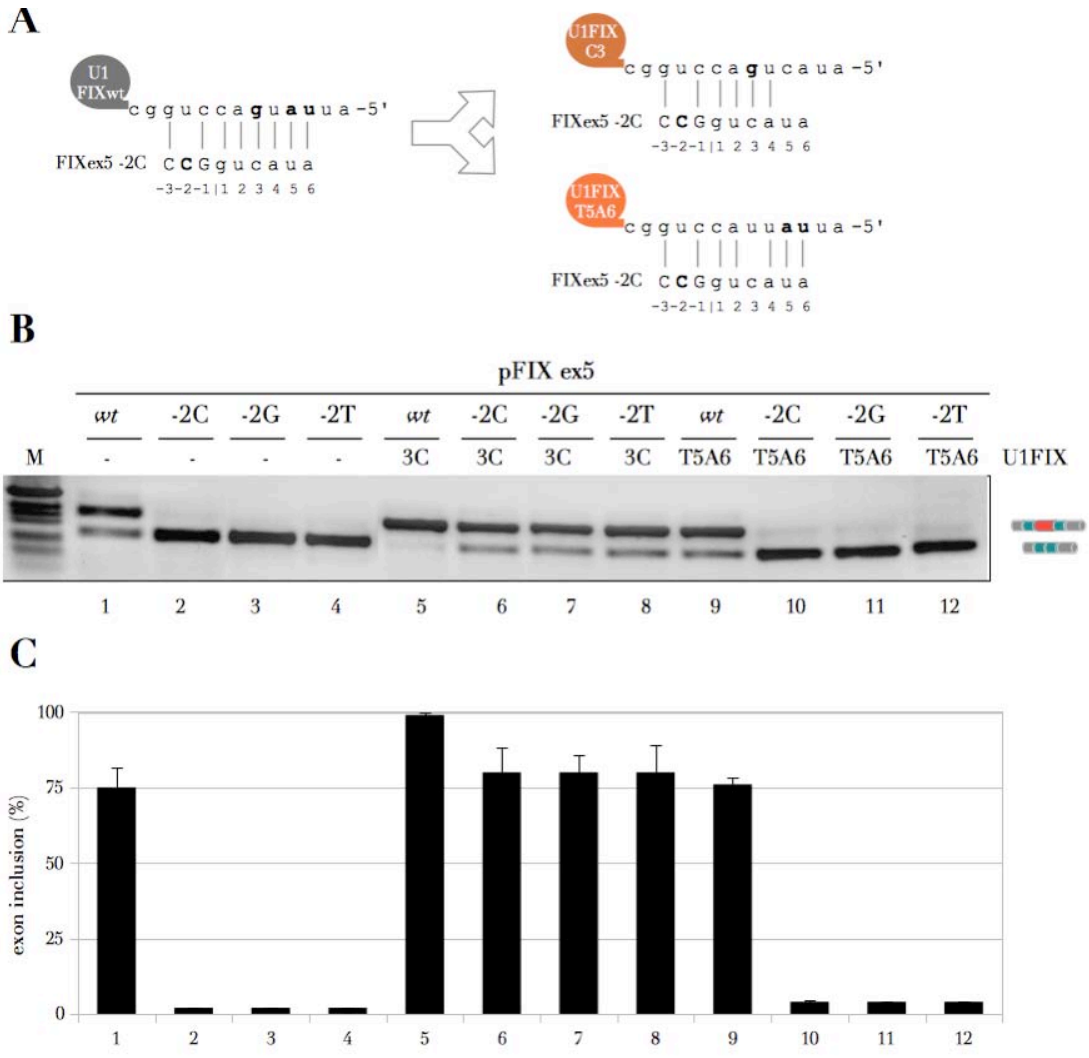

Figure 2.3.7. **Effect of U1FIX 3C and T5A6 on mutants at position -2.**
A, Complementarity between modified U1 snRNAs and *F9* exon 5 -2C mutant donor site. U1FIX 3C base pairs to a cytosine at position +3 while U1FIX T5A6 base pairs with a thymidine at position +5 and an adenine at position +6.
B, Splicing pattern analysis. *F9* exon 5 *wt* and mutant -2C, G and T minigenes were transfected into HeLa cells with the indicated modified U1 snRNAs and splicing pattern was evaluated by RT-PCR with alpha2,3 and bra2 primers. Amplified products were separated in electrophoresis on 2% agarose gel.
C, Quantification of exon 5 inclusion. Exon 5 inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate

Section 2.4. ***Effect of Exon Specific U1 snRNAs (ExSpeU1s) on* CFTR *exon 12 mutations***

To gain further insight into U1 snRNP role on exon recognition and to identify a possible unique U1 snRNA acting on multiple splicing defect on the same exon, I prepared a set of modified U1 snRNAs, we named exon-specific U1 snRNA (ExSpeU1). These ExSpeU1s have a modified 5'-tail allowing for binding to intronic sequences in the proximity of the 5' ss (Fig. 2.4.1 and 5.1). Cohen and colleagues (1994) have shown that U1 snRNAs base pairing at distance from the 5' ss can rescue exon inclusion induced by artificial donor site mutations. Based on these observations, I used a similar strategy for correcting the splicing defects induced by the *CFTR* exon 12 and *F9* exon 5 mutants. The list of all modified U1 snRNAs and their binding regions is schematically reported in Table 2.4.1 and Figure 4.1, for *CFTR* exon 12 and in Table 2.5.1 and Figure 5.1, for *F9* exon 5. This section reports the analysis on *CFTR* exon 12 whereas the experiments with *F9* exon 5 are described in the next section. I produced five ExSpeU1s for *CFTR* exon 12 namely U1CF 1, 9, 11, 15 and 33, whose target sequences (listed in Table 2.4.1) spanned from position +1 to position +44, downstream of donor site (Fig. 2.4.1). These ExSpeU1CFs were co-transfected into HeLa cells with *CFTR* exon 12 *wt*, exonic and donor site mutant minigenes and the resulting splicing pattern are shown in Fig. 2.4.2,

U1 snRNA	Target sequence (5' to 3')	Length
CF 1	**guaugu***ucuuuga*	13 bp
CF 9	*uuugaauaccuuacu*	15 bp
CF 11	*ugaauaccuuacuuau*	16 bp
CF 15	*uaccuuacuuauaa*	14 bp
CF 33	*caugcuaaaauaa*	13 bp

***Table 2.4.1.* List of exon-specific U1CF snRNAs created, their target sequence and length.**
The table lists a set of ExSpeU1CFs tested on *CFTR* exon 12 exonic and donor site mutations. The specific sequence of intronic flanking regions which 5'-tails of these ExSpeU1CFs are targeted to are reported in the second column, and their length are shown in the third. Donor splice site sequence is bolded.

4.3 and 4.4. In *wt* context, ExSpeU1CF 1, 9, 11 and 15 improved exon recognition, rising *CFTR* exon 12 inclusion levels from 80% to almost 100% (Fig. 2.4.2, lanes 1-6). Similarly, ExSpeU1CF 1, 9, 11 and 15 also stimulated exon 12 inclusion in 17A and 52T mutant context, although with variable efficiency (Fig. 2.4.2, lanes 8-12 and 15-18). Among all the tested ExSpeU1CFs, U1CF 11 showed the greatest stimulating effect, rising exon inclusion up to ~75% for both exonic mutations (Fig. 2.4.2, lanes 11 and 17). U1CF 33 had an opposite effect, inhibiting exon 12 inclusion both in *wt* minigene and in the two exonic mutants. In *wt* context, U1CF 33 reduced *CFTR* exon 12 inclusion from 80% to 50% (Fig. 2.4.2, compare lane 1 to 7). This inhibitory effect was also observed for 17A and 52T mutants,

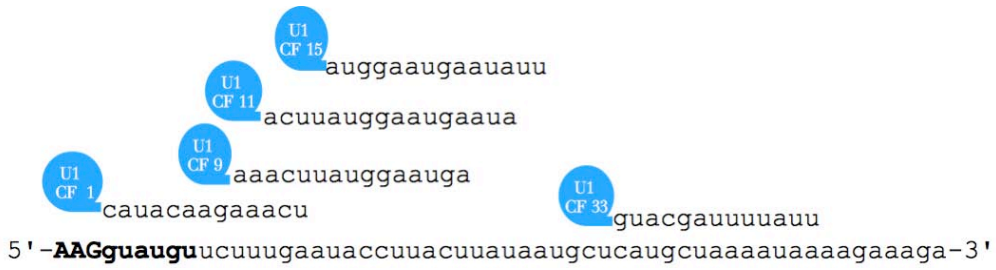

Figure 2.4.1. **Binding regions of Exon-Specific U1CF snRNAs**
Schematic representation of the positions where the ExSpeU1CFs bind. Exonic sequences are represented in capital letters and sequences in lowercase are part of the IVS12 flanking region. *CFTR* exon 12 donor splice site sequence is bolded.

although less evident, likely because these mutants presented lower basal levels (~10%) of exon inclusion (Fig. 2.4.2, lanes 13 and 19). Figures 4.3 and 4.4 show the effect of U1CFs on the 5' ss mutants. In these cases, co-transfection of the ExSpeU1CFs showed variable impact on exon inclusion depending on the minigene tested (Fig. 2.4.3 and 4.4). ExSpeU1 CF1 reverted exon skipping in -1A and -1T minigenes (Fig. 2.4.4, lanes 3 and 15) inducing ~ 50% of exon inclusion but had minimal or null effect in the +3C, +3G and +5A mutants (Fig. 2.4.3, lanes 3 and 9 and Fig. 2.4.4, lane 9). ExSpeU1CF 9 increased exon 12 inclusion in +3G and +5A mutants to ~28% and 10% respectively (Fig. 2.4.4, lanes 10 and 16) but had no effect on the other 5' ss mutations. ExSpeU1CF 15 did not affect

the splicing pattern of all tested 5' ss mutants (Fig.4.3, lanes 6 and 12; Fig. 2.4.4, lanes 6, 12, and 18).

Interestingly, ExSpeU1CF 11 exhibited a marked enhancement of exon 12 inclusion in all 5' ss mutants (Fig. 2.4.3, lanes 5 and 11; Fig. 2.4.4, lanes 5, 11 and 17). Specifically, U1CF 11 induced ~50% of exon inclusion in -1A, +3C and +3G mutants, ~40% in +5A mutant and ~20% in -1T mutant. Therefore U1CF 11 emerged as an exon-specific U1 snRNA (ExSpeU1) able to correct *CFTR* exon 12 skipping caused by exonic and donor site mutations.

The mutation +3G creates a new GU pair within the *CFTR* exon 12 donor site whose sequence becomes AAG|gtgtgt (Table 2.1.1). Since this new GU pair is located next to the canonical one, I confirmed with direct sequencing that the correct GU pair is used upon co-transfection with ExSpeU1CF 11.

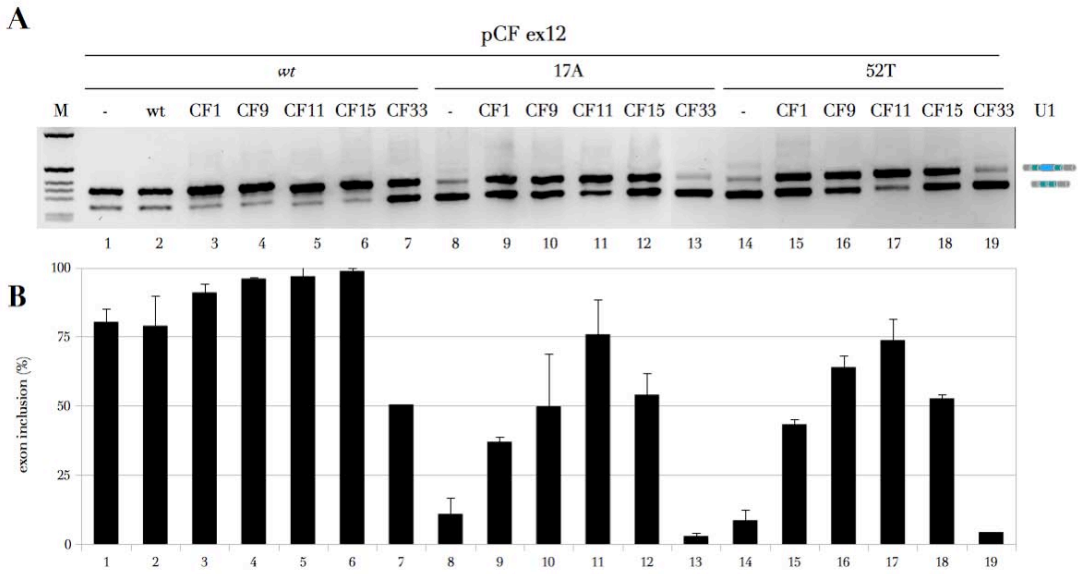

Figure 2.4.2. **Effects of exon-specific U1CF snRNAs on splicing pattern of pCF ex12 *wt*, 17A and 52T minigenes**

A, Splicing pattern of *CFTR* exon 12 *wt*, 17A and 52T minigenes co-transfected with plasmids encoding for ExSpeU1CF snRNAs (U1CF 1, 9, 11, 15 and 33) is visualized on a 2% agarose gel after RT-PCR. U1CF 1, 9, 11 and 15 reduced *CFTR* exon 12 skipping indicating that their binding near 5 'ss favoured exon definition. On the contrary, co-transfection with U1CF 33 reduced exon inclusion in the three minigene contexts.

B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

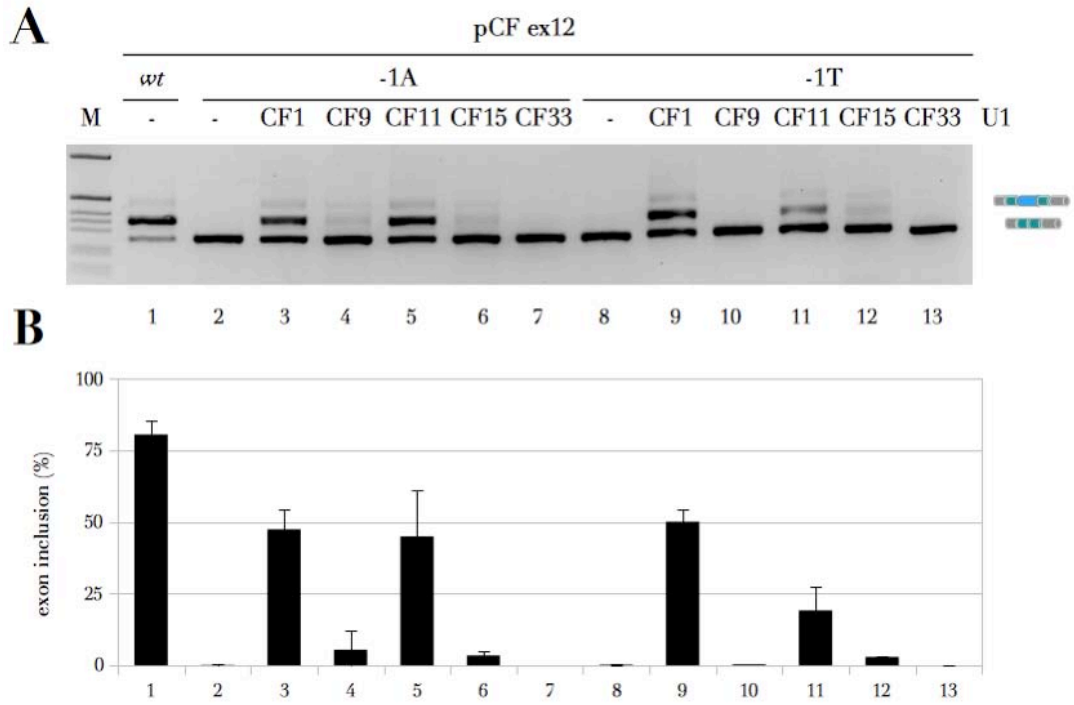

Figure 2.4.3. **Effects of exon-specific U1CF snRNAs on splicing pattern of pCF ex12 -1A and -1T minigenes**

A, Splicing pattern of *CFTR* exon 12 -1A and -1T minigenes co-transfected with plasmids encoding for ExSpeU1CFs sRNAs (U1CF 1, 9, 11, 15 and 33) is visualized on a 2% agarose gel after RT-PCR. U1CF 1 and 11 reduced *CFTR* exon 12 skipping in both mutant contexts, while U1CF 9, 15 and 33 showed negligible or no effect.

B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

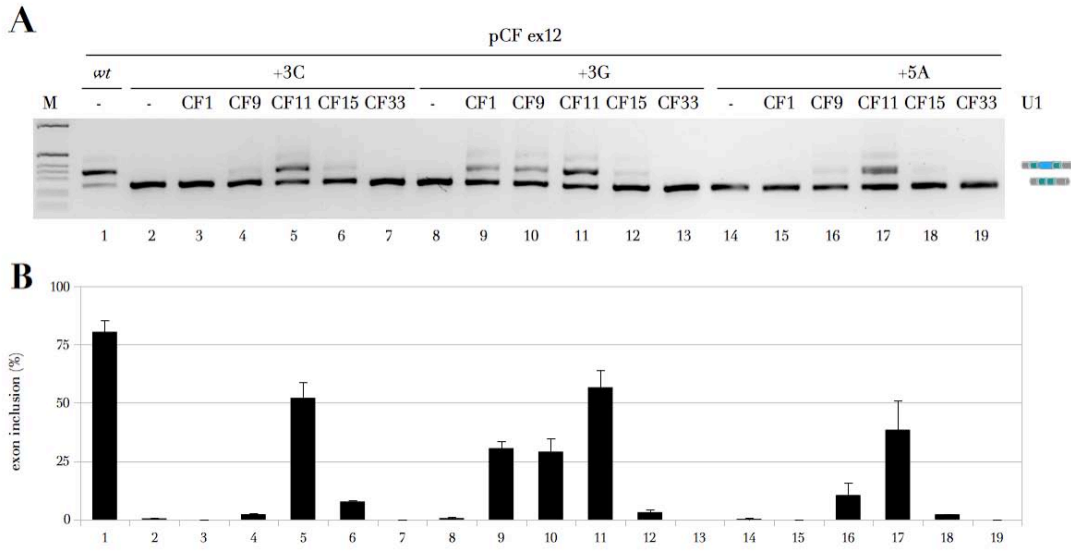

Figure 2.4.4. **Effects of exon-specific U1CF snRNAs on splicing pattern of pCF ex12 +3G, +3C and +5A minigenes**

A, Splicing pattern of *CFTR* exon 12 +3G, +3C and +5A minigenes co-transfected with plasmids encoding for ExSpeU1CF snRNAs (U1CF 1, 9, 11, 15 and 33) is visualized on a 2% agarose gel after RT-PCR. U1CF 11 reduced *CFTR* exon 12 skipping in all three mutant contexts. In +3G mutant context, U1CF 1 and 9 stimulated exon inclusion though less efficiently as U1CF 11. U1CF 15 and showed negligible or no effect in all three mutant contexts.

B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

Section 2.5. ***Effect of Exon-Specific U1 snRNAs on* F9 *exon 5 mutations.***

To identify a unique U1 snRNA correcting the splicing defects of *F9* exon 5 mutations, I first tested different *F9* exon-specific U1 snRNAs through co-transfection experiments on the pFIX ex5 -2C minigene. A total of eleven *F9* exon 5-specific U1 snRNAs (ExSpeU1FIX) binding in the proximity of the donor splice site were tested (Fig. 2.5.1). Analysis of the splicing pattern revealed that the majority of U1FIXs were able, although with different efficiency, to correct exon 5 skipping (Fig. 2.5.2). In particular, ExSpeU1FIX 1, 9, and 10 rescued very efficiently the aberrant splicing of the -2C mutation, leading to a percentage of inclusion of about 95-100% (Fig. 2.5.2, lanes 5, 7 and 8).

U1 snRNA	Target sequence (5' to 3')	Length
FIX -7	*ccag***caggtca**	11 bp
FIX 1	**gtcata***atct*	9 bp
FIX 7	*tctgaataaga*	13 bp
FIX 9	*tgaataaga*	9 bp
FIX 10	*gaataagat*	9 bp
FIX 13	*taagatttt*	9 bp
FIX 16	*gatttttta*	9 bp
FIX 22	*ttaaagaaa*	9 bp
FIX 33	*ctgtatcta*	9 bp
FIX 38	*ctgaaactt*	9 bp
FIX 63	*aacctacta*	9 bp

Table 2.5.1. **List of exon-specific U1FIX snRNAs created, their target sequence and length.**
The table lists a set of ExSpeU1FIX snRNAs tested on exonic *F9* exon 5 -2C mutation. The specific sequence of intronic flanking regions which 5'-tails of these ExSpeU1FIXs are targeted to are reported in the second column, and their length are shown in the third. Donor splice site sequence is bolded.

Generally, the percentage of exon 5 inclusion progressively decreased as binding of the ExSpeU1FIXs moved downstream (U1FIX 13, 16, 22, 33, 38 and 63; Fig. 2.5.2, lanes 9-14). Base pairing of ExSpeU1FIXs between position +1 and +18 had the strongest stimulatory effect on -2C mutant exon 5 inclusion whereas downstream binding progressively decreased the efficacy. Of the three most active ExSpeU1FIXs, I chose U1FIX 9 for further analysis on the other *F9*

mutants: the two additional synonymous mutations at position -2, the mutations -1T, +1A and +2C and the PPT mutations -8G and -9G. Co-transfection of ExSpeU1FIX 9 completely rescued the exon skipping caused by -8G and -9G PPT mutations and by -2 (-2G and -2T) mutations (Fig. 2.5.3, lanes 1-10). However, the splicing defects present in the -1T, +1A and +2C minigenes were not corrected by co-transfection of ExSpeU1FIX 9 (Fig. 2.5.3, lanes 11-16). Interestingly, the *F9* exon 5 mutants rescued by U1FIX 9 were also similarly rescued by U1FIXwt (Fig. 2.3.5) or U1FIX 3C (Fig. 2.3.7). Exon skipping due to mutations -1T, +1A and +2C could not be reverted by modified U1 snRNAs, either directed to the mutant 5' ss or loaded at a certain distance from the 5' ss.

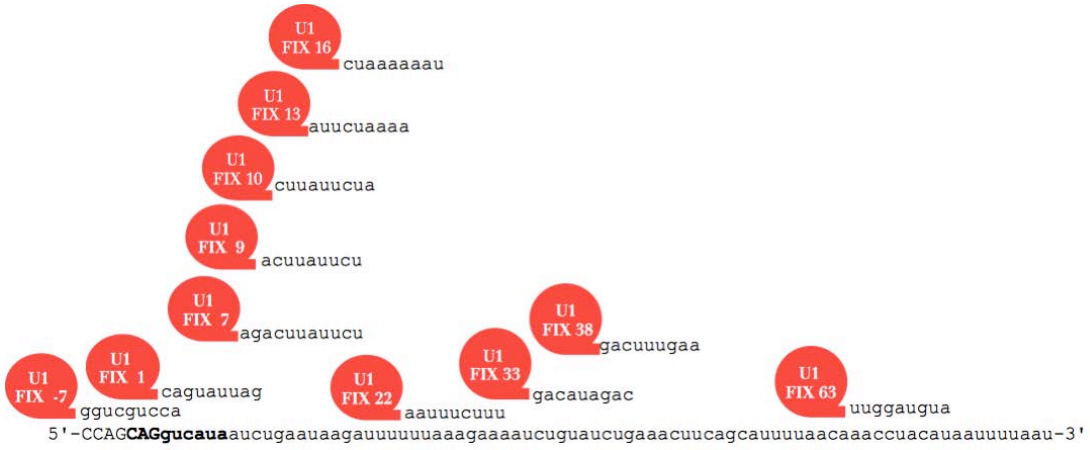

Figure 2.5.1. **Binding regions of exon-specific U1FIX snRNAs**
Schematic representation of the positions where ExSpeU1FIX bind. Exonic sequences is represented in uppercase letters and part of the IVS5 flanking region in lowercase letters. *F9* exon 5 donor splice site is bolded.

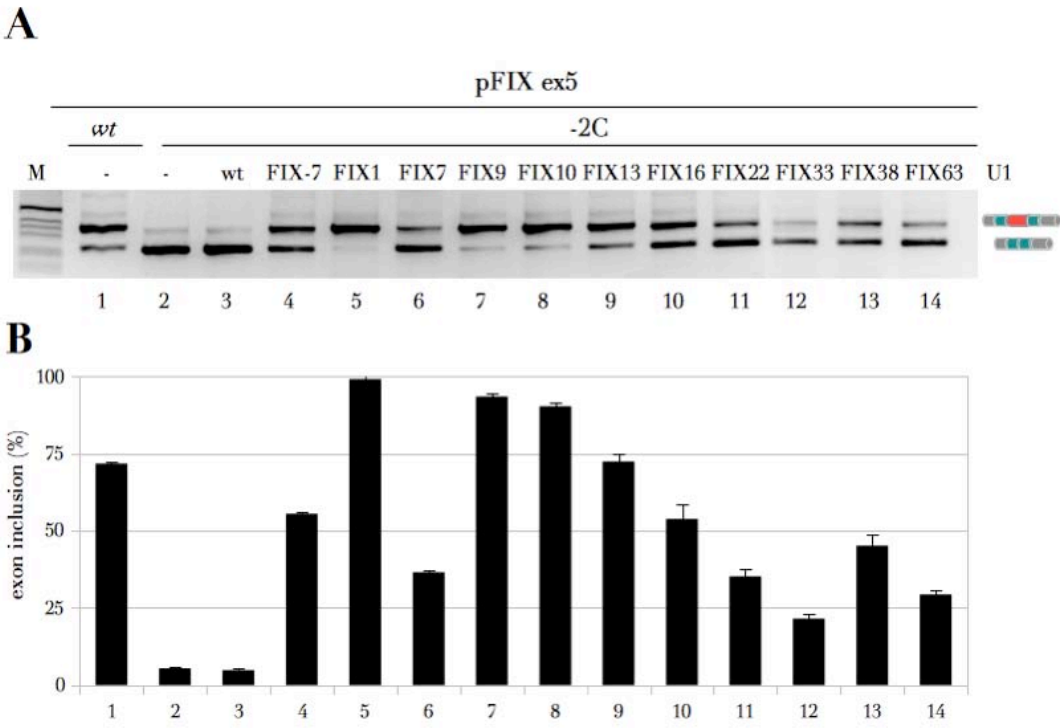

Figure 2.5.2. **Effects of exon-specific U1FIX snRNAs on splicing pattern of pFIX ex5 -2A>C minigene**

A, Splicing pattern of *F9* exon 5 -2C minigene co-transfected with plasmids encoding for ExSpeU1FIX snRNAs (U1FIX -7, 1, 7, 9, 10, 13, 16, 22, 33, 38, 63) is visualized on a 2% agarose gel after RT-PCR. All ExSpeU1FIXs reduced *F9* exon 5 skipping indicating that their binding near 5' ss favoured exon definition. This stimulatory activity loose strength as ExSpeU1FIXs binding sites move away from 5' ss. U1FIX 1, 9 and 10 exhibited the highest stimulatory activity, which is clearly evidenced from the histogram in B.

B, Quantification of exon 5 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

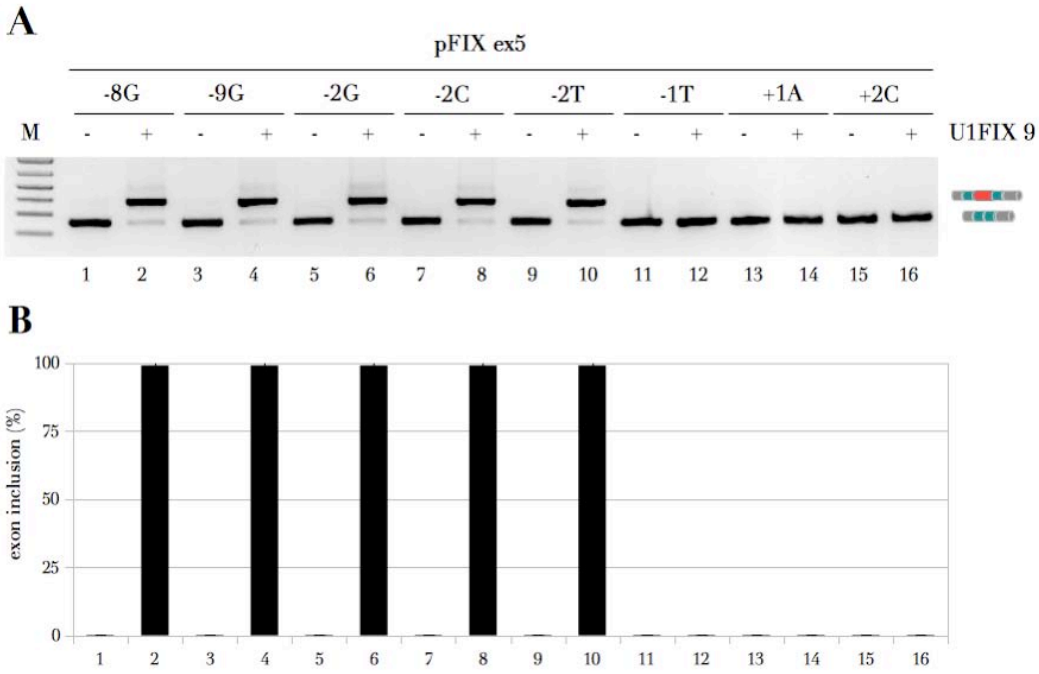

Figure 2 5.3. **Effect of ExSpeU1FIX 9 on splicing pattern of polypyrimidine tract and donor site mutations in *F9* exon 5**

A, Splicing pattern of polypyrimidine tract (PPT; -8G, -9G) and donor site (-2G, -2C, -2T, -1T, +1A, +2C) mutant minigenes co-transfected with ExSpeU1FIX 9 is visualized on a 2% agarose gel after RT-PCR. U1FIX 9 promotes exon inclusion in presence of mutations at the PPT and position -2 at the donor splice site of *F9* exon 5. Exon skipping caused by donor site mutations -1T, +1A and +2C is not corrected by ExSpeU1FIX 9.

B, Quantification of exon 5 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

Section 2.6. ***Analysis of* F9 *exon 5 +13G mutation and the effect of compensatory U1 snRNAs.***

As previously shown in Section 2.1, the naturally occurring mutation +13G in the intronic region flanking *F9* exon 5 creates a cryptic splice site downstream the canonical site (Table 2.1.2 and Fig. 2.2.3, lane 9). To test whether preferential usage of natural 5' ss can be enhanced using modified U1 snRNAs, I performed co-transfection experiments with ExSpeU1FIX 9 (Table 2.5.1 and Fig. 2.5.1), whose 5'-tail binds between position +9 and +17 overlapping the mutation. Interestingly, the A to G mutation did not disrupt the complementarity between the ExSpeU1FIX 9 and the target sequence, which covers the region where the cryptic 5' ss is located (Fig. 2.6.1). To evaluate the efficacy of ExSpeU1FIX 9 and quantify the percentage of exon inclusion, I carried out a denaturing capillary electrophoresis of fluorescent-labelled RT-PCR products (Fig. 2.6.2). As shown in the graphic, the experimental procedure quantified the amount of normal and aberrant transcripts produced in the presence or absence of the ExSpeU1FIX 9. The mutant +13G minigene produced a low amount (about 3%) of normal transcript (peak 2) and exhibited prevalent usage of the cryptic 5' ss (peak 3). Co-transfection with ExSpeU1FIX 9 increased the level of normal transcript up to about 41%. As a control, co-transfection of U1wt did not affect the splicing pattern (Fig. 2.6.2).

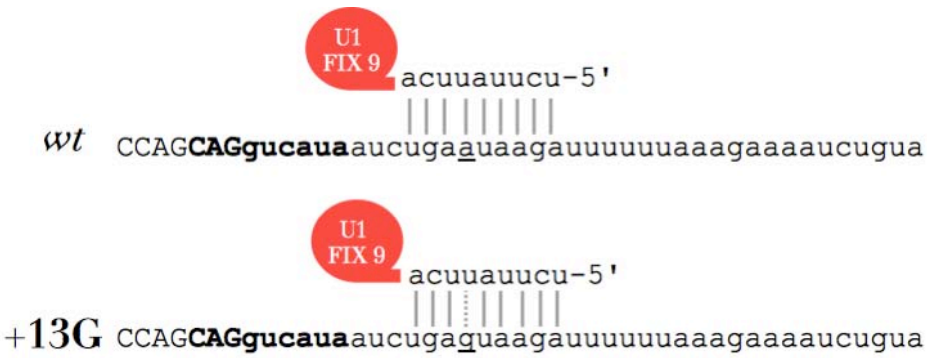

Figure 2.6.1. **Complementarity between U1FIX 9 and its target region for *F9* exon 5 *wt* and mutant +13G.**

Uppercase letters represent exonic sequence and lowercase letters intronic sequence. *F9* exon 5 donor splice site is bolded; position +13 is underlined. Vertical lines symbolize Watson-Crick (continuous) and non-Watson-Crick (dashed) bonds.

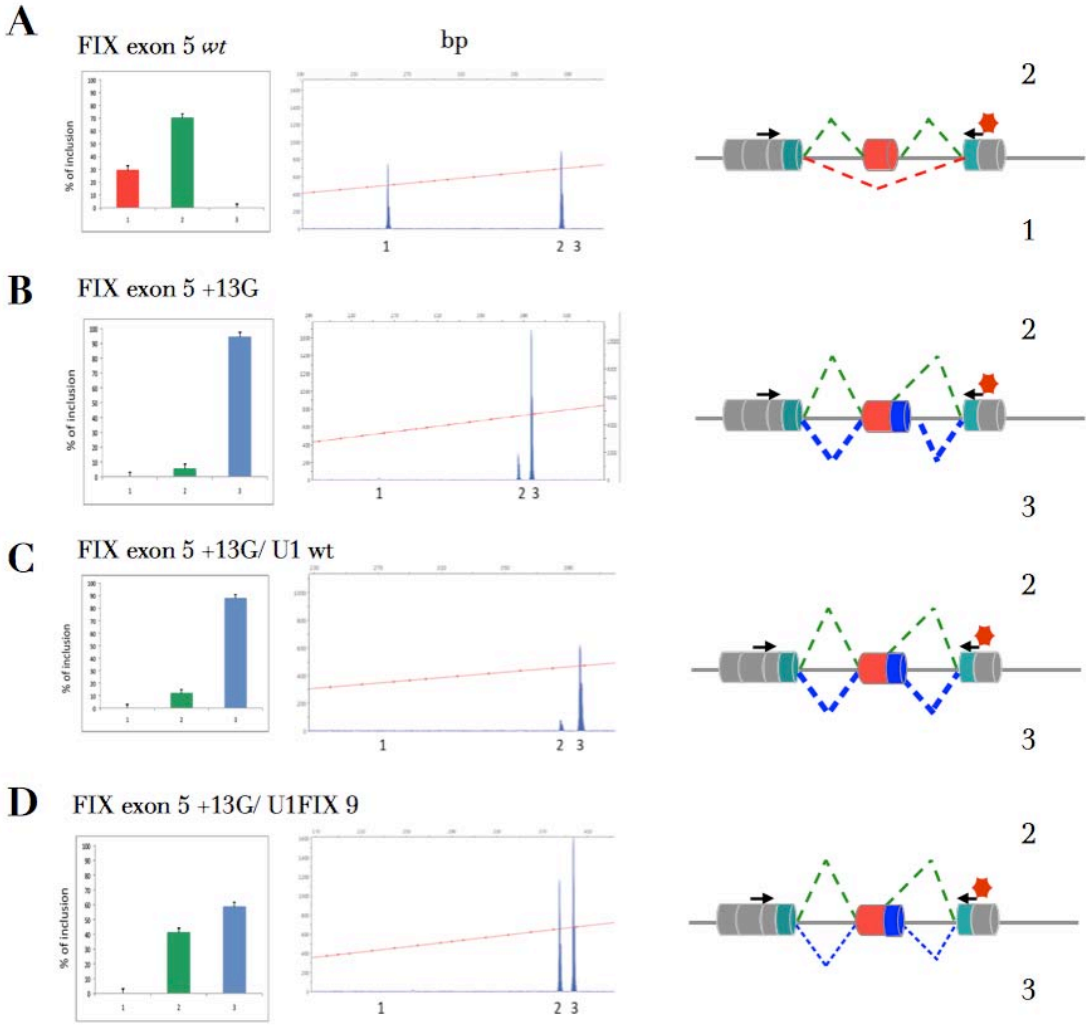

Figure 2.6.2. **Splicing pattern analysis of *F9* exon 5 +13G mutant minigene co-transfected with exon-specific U1FIX 9 through capillary electrophoresis**

Quantification graphic of normal and aberrant splicing after separation on a denaturing capillary system (Automated ABI-3100) of fluorescently labeled RT-PCR products obtained from total RNA of HeLa cells expressing:

A, pFIX exon 5 *wt*;

B, pFIX exon 5 +13G

C, pFIX exon 5 +13G co-transfected with U1wt

D, pFIX exon 5 +13 G co-transfected with UFIX 9

The pattern of splicing was evaluated by RT-PCR using alpha2,3 and bra2^{FAM} primers. Bra2^{FAM} primer was fluorescently labeled. The amplified PCR products were then evaluated by capillary electrophoresis in the presence of a ROX standard.

On the middle panel of the figure the electrophoresis spectra of 1 uL of 1:100 diluted RT-PCR reaction is shown. Each peak represent a transcript amplified using the above mentioned oligo. Peak 1 is a transcript of 247 bp (which does not include *F9* exon 5) while peak 2 is a transcript of 375 bp (which includes *F9* exon 5) and peak 3 is 13 bp longer than peak 2 (388 bp) due to the use of the cryptic site at position +13.

On the right side, the quantification of the relative abundance of the peaks is illustrated. On the left side, splicing pattern analysis with corresponding outcomes are depicted. Red dashed lines, exon 5- transcript (peak 1); green dashed lines, exon +5 transcript (peak 2); blue dashed lines, cryptic donor site usage (peak 3).

Section 2.7. ***Identification of intronic splicing silencers (ISS) downstream the 5′ ss of*** **F9** ***exon 5 and*** **CFTR** ***exon 12.***

The exon-specific U1FIXs that bind between position +1 to +18 downstream of *F9* exon 5 donor site showed the most significant effect on exon inclusion (Fig. 2.5.2). Similarly, the ExSpeU1CF 9, 11 and 15, binding from position +1 to +28, modulated exon inclusion in the *CFTR* exon 12 system (Fig. 2.4.2). The positive effect of these ExSpeU1s on exon processing might be due to their direct interference with intronic splicing regulatory elements located downstream the splice site. To test this hypothesis, I engineered several mutant minigenes with a set of consecutive intronic deletions downstream the donor sites, as illustrated in Fig. 2.7.1 and Fig. 2.7.2 for *CFTR* exon 12 and in Fig. 2.7.3 for *F9* exon 5. Concerning the *CFTR* exon 12 minigene, I introduced six deletions in the *wt* context: from position +12 to +15 (Δ12-15), from position +17 to +23 (Δ17-23), from position +24 to +34 (Δ24-34), from position +34 to +44 (Δ34-44), from position +45 to +55 (Δ45-55) and from position +55 to +65 (Δ55-65). These mutants were evaluated upon transfection into HeLa cells and the percentage of exon 12 inclusion was analysed by RT-PCR. In comparison to the *wt* CFTR exon 12, the Δ12-15 and Δ17-23 mutants increased the percentage of exon inclusion. In fact, the percentage of exon inclusion was increased from 80% to 100% (Fig.7.1B,C, lanes 2 and 3). The other deletion mutants did not affect the splicing pattern (Fig. 2.7.1B,C lanes 4-7). Similarly, deletion of Δ12-15 and Δ17-23 in 17A context induced exon inclusion (Fig.7.1D, lanes 2 and 3) whereas the other deletions had no effect on pre mRNA splicing (Fig.7.1D,E lanes 4-7). In this case, the 10% exon inclusion of 17A was increased to 50% for Δ12-15 and 25% for Δ17-23. Therefore, these observations suggest that, in the *wt* and 17A contexts, the intronic sequence from position +12 to +23 acts as a weak intronic splicing silencer (ISS). However, when the Δ12-15 and the Δ17-23 were evaluated in the context of donor splice site mutants, the deletions did not

modify the splicing outcome (Fig. 2.7.1D and E). Indeed, the -1A and +3G mutants showed complete exon skipping and the presence of the Δ12-15 or Δ17-23 deletions did not change the splicing pattern (Fig. 2.7.2). These data indicate that the ISS in position +12 to +23 downstream to the 5' ss can

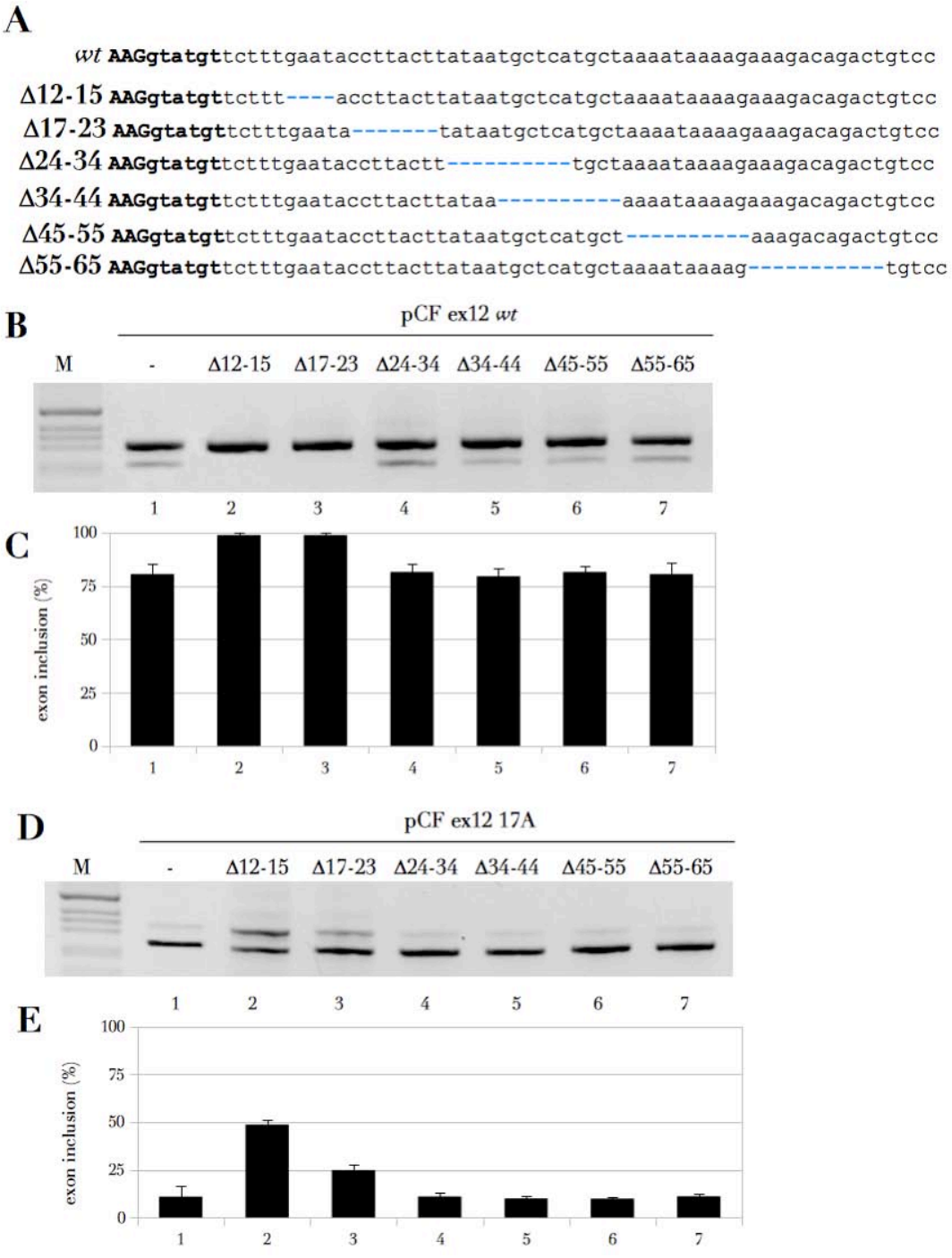

Figure 2.7.1. **Serial intronic deletion mutagenesis analysis downstream of *CFTR* exon 12 donor splice site**

A, *CFTR* exon 12 *wt* donor splice site sequence (bolded) and downstream intronic region. Serial intronic deletion mutants sequences are shown below. These mutants were created by overlap extension PCR, using specific oligonucleotides complementary to up- and downstream sequences of the deleted region. Uppercase letters, exonic sequence; lowercase letters, intronic sequence. Dashes represent the deleted bases.

B, Splicing pattern analysis of pCF ex12 *wt* deletion mutants (Δ12-15, Δ17-23, Δ24-34, Δ34-44, Δ45-55 and Δ55-65). These minigenes were transfected into HeLa cells and subsequently RT-PCR was performed on total RNA. PCR products are visualized after 2% agarose gel electrophoresis. In this *wt* context, deletion of bases +12 to +23 downstream of *CFTR* exon 12 donor splice site promoted exon inclusion (lanes 2 and 3). Deletions from base +24 to +65 did not affect *wt* splicing pattern (lanes 4-7).

D, Splicing pattern analysis of pCF ex12 17A deletion mutants (Δ12-15, Δ17-23, Δ24-34, Δ34-44, Δ45-55 and Δ55-65). These minigenes were transfected into HeLa cells and subsequently RT-PCR was performed on total RNA. PCR products are visualized after 2% agarose gel electrophoresis. In this 17A context, deletion of bases +12 to +23 downstream of *CFTR* exon 12 donor splice site promoted exon inclusion (lanes 2 and 3). Deletions from base +24 to +65 did not affect *wt* splicing pattern (lanes 4-7).

C and E, Quantification of exon inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD.

modulate the *CFTR* exon 12 splicing in the *wt* context and in the two exonic variants but not when the 5' ss donor site is mutated. Concerning the FIX exon 5, I engineered a set of consecutive 3 or 4 bp deletions, starting from the sequences targeted by the most effective ExSpeU1FIXs, U1FIX 9 and 10. I deleted the sequences spanning from base +8 to +10 (Δ8-10), from +11 to +13 (Δ11-13), from +15 to +19 (Δ15-19) and from +19 to +24 (Δ19-24) as illustrated in Fig. 2.7.3A. The four resulting mutant minigenes were transfected into HeLa cells and the percentage of exon 5 inclusion was analysed by RT-PCR (Fig. 2.7.3B). The data showed that deletions from base +11 to +13 and from base +15 to +19 significantly increased the percentage of exon 5 inclusion in the *wt* context (Fig. 2.7.3B, lanes 3 and 4). In contrast, deletions from base +8 to +11 and from base +19 to +24 had no effect on the splicing outcome (Fig. 2.7.3B, lanes 2 and 5). These results indicated that the sequence *gaauaagauu* downstream to the exon 5 donor splice site acts as an intronic splicing silencer (ISS). Altogether these mutagenesis experiments demonstrated the presence of intronic splicing silencer elements downstream to the 5' ss in both *CFTR* exon 12 and *F9* exon 5. They are located in a position that corresponds to the target sequence of most of the active ExSpeU1s.

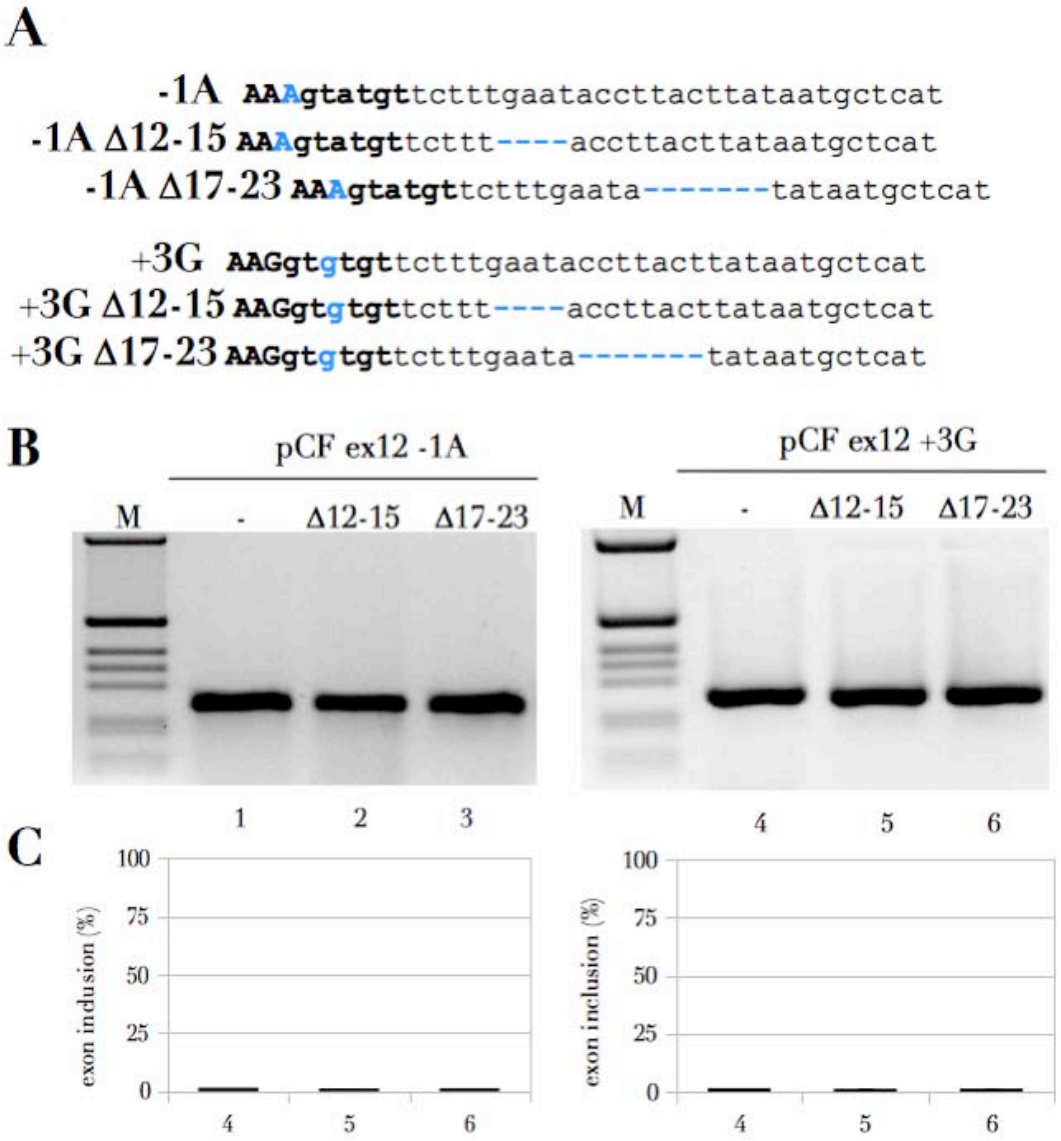

Figure 2.7.2. **Effect of deletion mutants Δ12-15, Δ17-23 in presence of *CFTR* exon 12 donor splice site mutations -1A and +3G**

A, *CFTR* exon 12 donor splice site region in presence of mutation -1A or +3G (coloured) and downstream intronic region. Deletion mutants Δ12-15 and Δ17-23 are shown below. These mutants were created by overlap extension PCR, using specific oligonucleotides complementary to up- and downstream sequences of the deleted region. Donor splice site sequence is bolded. Uppercase letters, exonic sequence; lowercase letters, intronic sequence. Dashes represent the deleted bases.

B, Splicing pattern analysis of pCF ex12 -1A and +3G deletion mutants Δ12-15 and Δ17-23. These minigenes were transfected into HeLa cells and subsequently RT-PCR was performed on total RNA. PCR products are visualized after 2% agarose gel electrophoresis. Deletions Δ12-15 and Δ17-23 did not affect splicing pattern of neither mutant -1A nor +3G. (lanes 2-3, 5-6).

C, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

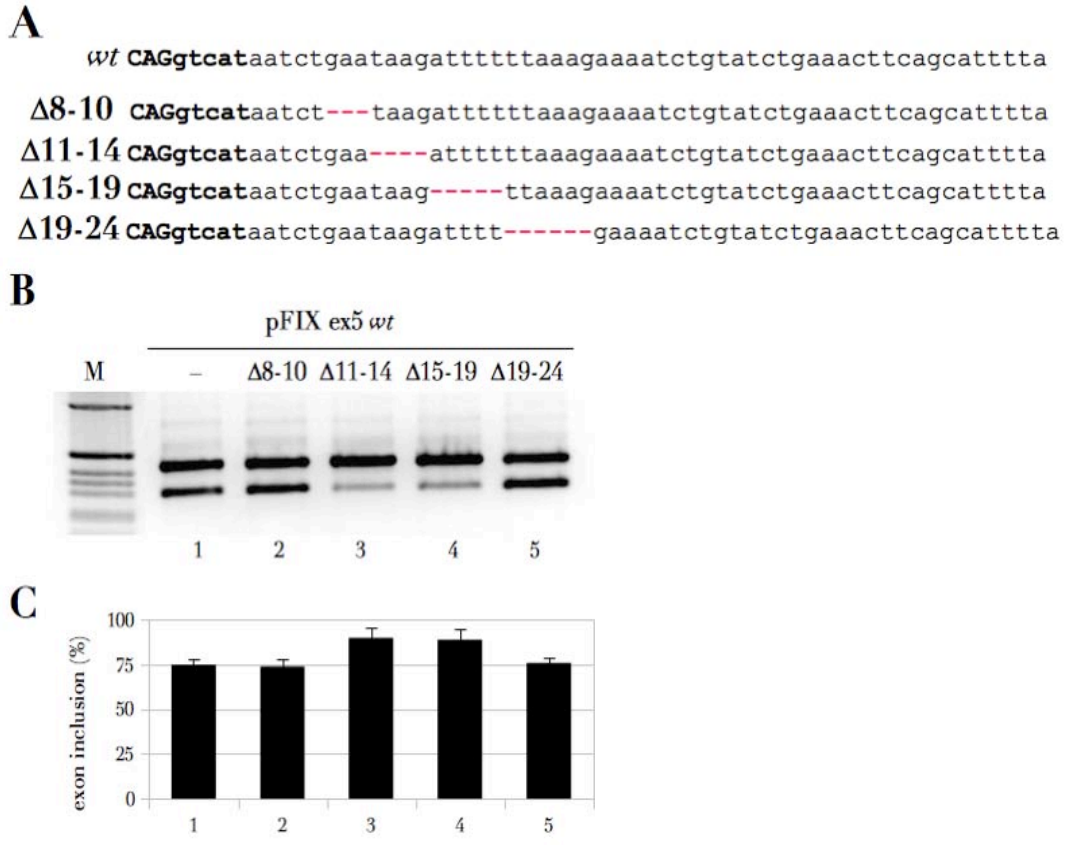

Figure 2.7.3. **Serial intronic deletion mutagenesis analysis downstream of *F9* exon 5 donor splice site**

A, *F9* exon 5 *wt* donor splice site sequence (bolded) and downstream intronic region. Serial intronic deletion mutants sequences are shown below. These mutants were created by overlap extension PCR, using specific oligonucleotides complementary to up- and downstream sequences of the deleted region. Uppercase letters, exonic sequence; lowercase letters, intronic sequence. Dashes represent the deleted bases.

B, Splicing pattern analysis of pFIX ex5 *wt* deletion mutants (Δ8-10, Δ11-14, Δ15-19 and Δ19-24). pFIX ex5 *wt* and the deletion mutant minigenes were transfected into HeLa cells and subsequently RT-PCR was performed on total RNA. PCR products are visualized after 2% agarose gel electrophoresis.

Deletion of bases +11 to +24 downstream of *F9* exon 5 donor splice site promoted exon inclusion (lanes 3 and 4). In contrast, deletions from base +8 to +10 and +19 to +24 did not affect *wt* splicing pattern (lanes 2 and 5).

C, Quantification of exon 5 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

Section 2.8. ***Co-transfection of ExSpeU1s with ISS deletion mutants in* CFTR *exon 12 minigenes.***

To verify that the splicing rescue mediated by the ExSpeU1s is due to their direct base-pairing with the target intronic sequences, I performed co-transfection experiments of selected ExSpeU1s along with intronic deletion mutants. In the previous section (Section 2.7), I have shown that deletion of sequences ranging from positions +12 to +15 downstream of the 5' ss improved *CFTR* exon 12 inclusion in the *wt* context and in the exonic mutant 17A (Fig. 2.7.1B and D, lanes 2 and 3). Deletion Δ12-15 abolished most of the ExSpeU1CF 11 complementarity, but preserved the complementarity of ExSpeU1CF 15, as shown in Fig. 2.8.1A. These features make pCF ex12 17A Δ12-15 the ideal minigene candidate to test whether ExSpeU1s activities depend on base pairing to their target sequences. Therefore, I co-transfected pCF ex12 17A Δ12-15 minigene with ExSpeU1CF 11 and U1CF 15 (Fig. 2.8.1).

ExSpeU1CF 11 did not change the splicing pattern of pCF ex12 17A Δ12-15 minigene pre-mRNA (Fig.8.1B, lane 5) whereas ExSpeU1CF 15 stimulated exon 12 inclusion (Fig. 2.8.1B, lane 6). These data confirm that binding of ExSpeU1CFs to their target sequences induce exon inclusion.

In comparison to the other ExSpeU1CFs, ExSpeU1CF 33 had the opposite effect on exon processing. In fact, U1CF 33 inhibits *CFTR* exon 12 inclusion (Fig. 2.4.2). To test whether its splicing inhibition was dependent on base pairing to +33 to +44 sequences, I co-transfected U1CF 33 with *CFTR* exon 12 *wt* deletion mutant Δ34-44. U1CF 33 still promoted exon skipping (Fig. 2.8.2, compare lanes 2 and 4). Therefore, the inhibiting effect on exon inclusion displayed by ExSpeU1CF 33 does not depend on its target intronic sequences. Subsequently, to understand if the splicing inhibition played by ExSpeU1CF 33 was specific to *CFTR* exon 12 I co-transfected U1CF 33 with the unrelated pFIX ex5 *wt* minigene. The pFIX exon 5 *wt* shares with pCF ex12 *wt* the pTB minigene

context, and the two minigenes differ in the central region corresponding to the exons and to approximately 300 bp of flanking introns. Upon co-transfection neither ExSpeU1CF 11 nor U1CF 33 induced changes on splicing pattern of *F9* exon 5 *wt* (Fig. 2.8.3A). Consequently, the ExSpeU1CF 33 effect is restricted to the *CFTR* exon 12. ExSpeU1CF does not play a general inhibitory effect on exon processing. The observed inhibition of *CFTR* exon 12 inclusion could be due to specific binding of U1CF 33 to intronic or exonic *CFTR* exon 12 sequences.

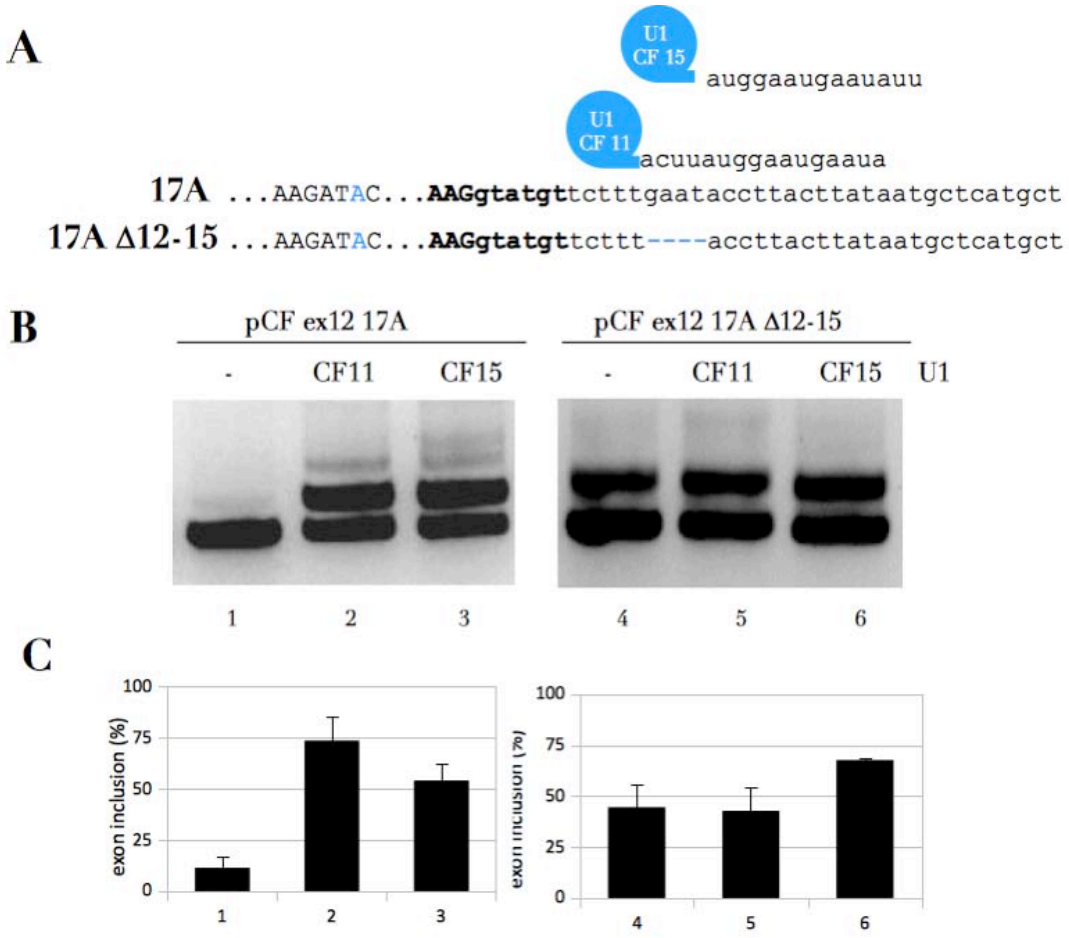

Figure 2.8.1. **Exon-specific U1CF snRNAs act through base pairing to their target sequence**
A, Sequence of *CFTR* exon 12 17A exonic mutant and deletion mutant 17A Δ12-15 deletion mutant, donor splice site sequence (bolded) and downstream intronic region. Target sequence o ExSpeU1CF 11 is partially deleted in Δ12-15 mutant. On the other hand, thymidine at position 15 is replaced by thymidine at position +11, thus maintaining ExSpeU1CF 15 target sequence. Uppercase letters, exonic sequence; lowercase letters, intronic sequence. Mutant adenine is coloured and dashes represent the deleted bases in the intronic deletion mutant
B, Splicing pattern analysis of pCF ex12 17A and 17A Δ12-15. These minigenes were transfected into HeLa cells and subsequently RT-PCR was performed on total RNA. PCR products were visualized after 2% agarose gel electrophoresis. U1CF 11 positive effect on 17A mutant is lost when bases +12 to +15 are absent, as it is the case in the pCF ex12 17A Δ12-15 minigene (compare lane 2 and 5). U1CF 15 target sequence is preserved in the pCF ex12 17A Δ12-15 minigene and consequently it showed the same activity as in the pCF ex12 17A minigene (compare lanes 3 and 6).
C, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

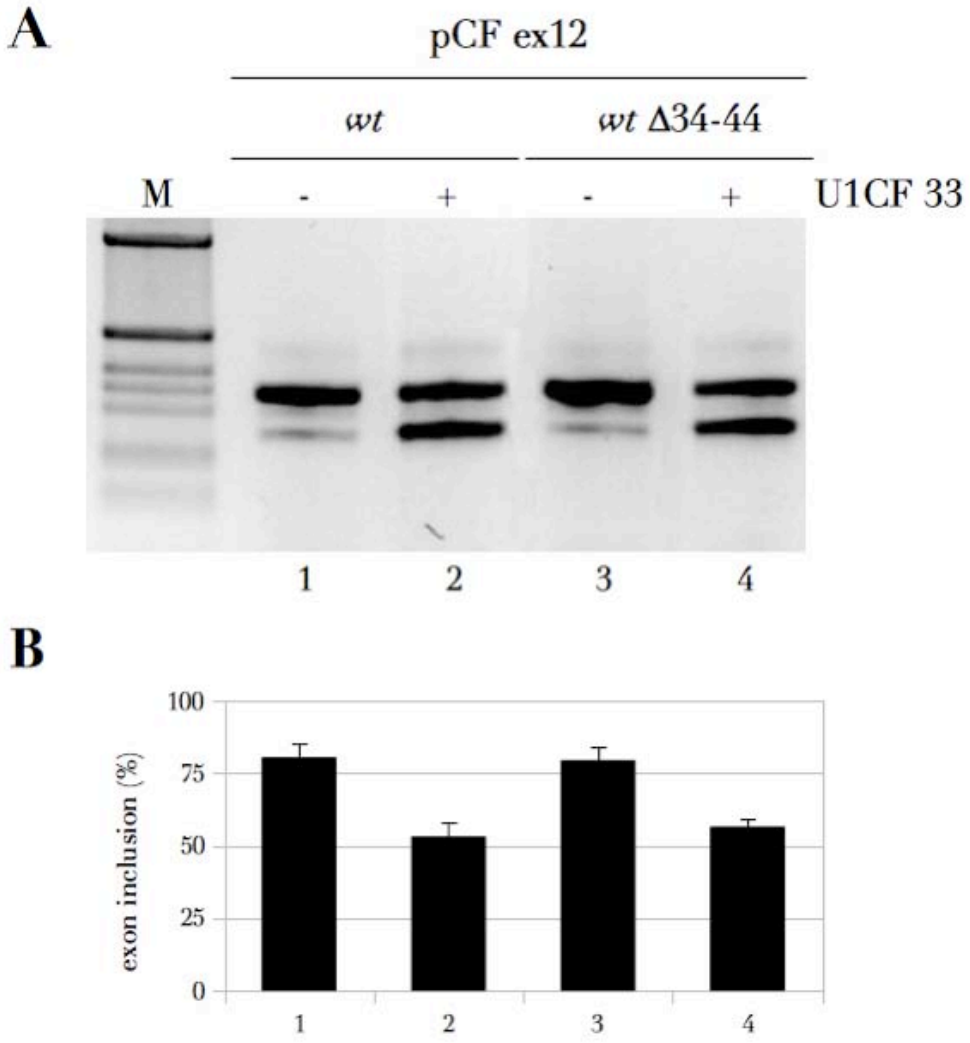

Figure 2.8.2. **Inhibitory activity of U1CF 33 on splicing is independent of base pairing to region +34-+44 downstream of *CFTR* exon 12 5'ss**

A, Splicing pattern analysis of pCF ex12 *wt* (lanes 1 and 2) and *wt* Δ34-44 (lanes 3 and 4) transfected either alone or in presence of a plasmid encoding for U1CF 33 (see Table 2.4.1 and Fig. 2.4.1 for its target sequence and location). RT-PCR was performed on total RNA as describe before and splicing products were resolved on a 2% agarose gel electrophoresis. Deletion of U1CF 33 target sequence did not abolish its inhibitory activity on *CFTR* exon 12 inclusion as it promoted exon skipping in *wt* context (compare lane 2 and 4).

B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

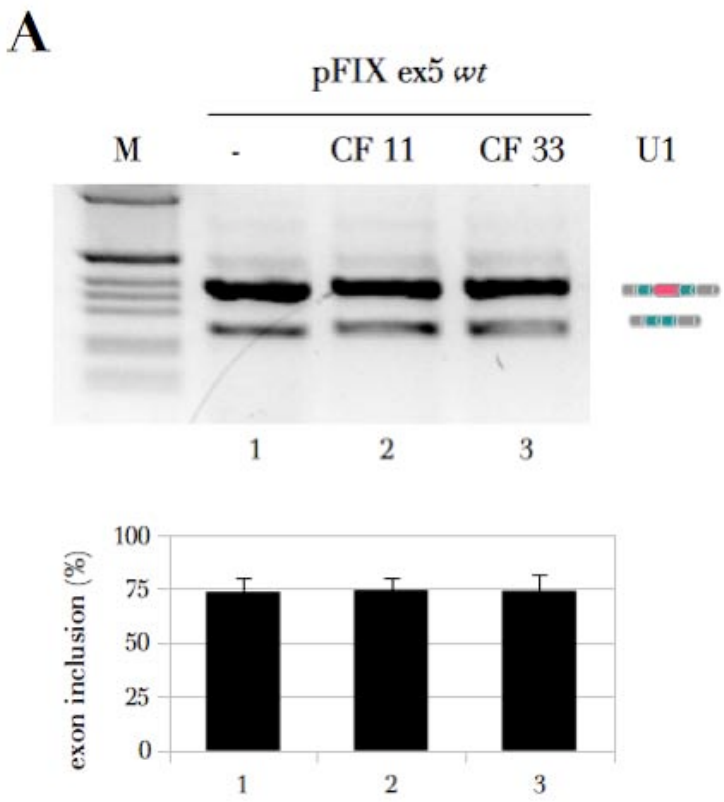

***Figure 2.8.3*. U1CF 33 did not affect pFIX ex5 *wt* splicing pattern**
A, Splicing pattern of *F9* exon 5 *wt* minigene transfected alone (lane 1) or with plasmids encoding for U1CF 11 (lane 2) or U1CF 33 (lane 3) is visualized on a 2% agarose gel after RT-PCR. Neither U1CF 11 nor U1CF 33 induced changes in *F9* exon 5 *wt* splicing pattern. Since pCF ex12 and pFIX ex5 minigenes share the pTB construct backbone, this result indicates that U1CF 33 splicing inhibitory activity is specific for the *CFTR* exon 12 fragment.
B, Quantification of exon 5 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

Section 2.9. ***HnRNP A1/A2 binds to ISS in* F9 *exon 5 but not to* CFTR *exon 12 ISS.***

Deletion analysis in *F9* exon 5 and *CFTR* exon 12 showed an intronic splicing silencer located downstream of the donor splice sites. Interestingly, the *s*urvival of *m*otor *n*euron *2* (*SMN2*) exon 7 has a similar intronic splicing silencer downstream its 5'ss (Singh et al. 2006). As reported in Fig. 2.9.1, *SMN2* ISS shares some sequence similarity with *F9* exon 5 and *CFTR* exon 12. As *SMN 2* exon 7 ISS activity is mediated through binding of hnRNP A1/A2, a well-known *trans*-acting splicing factor (Hua et al. 2008), I tested whether this splicing factor might bind to the ISS in *F9* exon 5 and *CFTR* exon 12. For this purpose, I performed a RNA pulldown assay followed by Western blot (WB) analysis using *CFTR* exon 12 Δ12-15 RNA and the *F9* exon 5 Δ15-19 RNA probes (depicted in Fig. 2.9.1A). I used a polyclonal antibody that recognizes both hnRNP A1 and A2. Pull-down assay revealed that hnRNP A1/A2 binds efficiently to the normal *F9* exon 5 but not to the RNA lacking the ISS with deletion of the bases +15 to +19 (Fig. 2.9.1, lanes 1-3). On the contrary, hnRNP A1/A2 dis not bind to the *CFTR* exon 12 ISS (Fig. 2.9.1B, lanes 4 and 5). Therefore, hnRNP A1/A2 displayed differential binding properties of the two ISS: it bound to the ISS in *F9* exon 5 factor but not the ISS in *CFTR* exon 12 and, the ISS deletion in *F9* exon 5 reduced its binding .

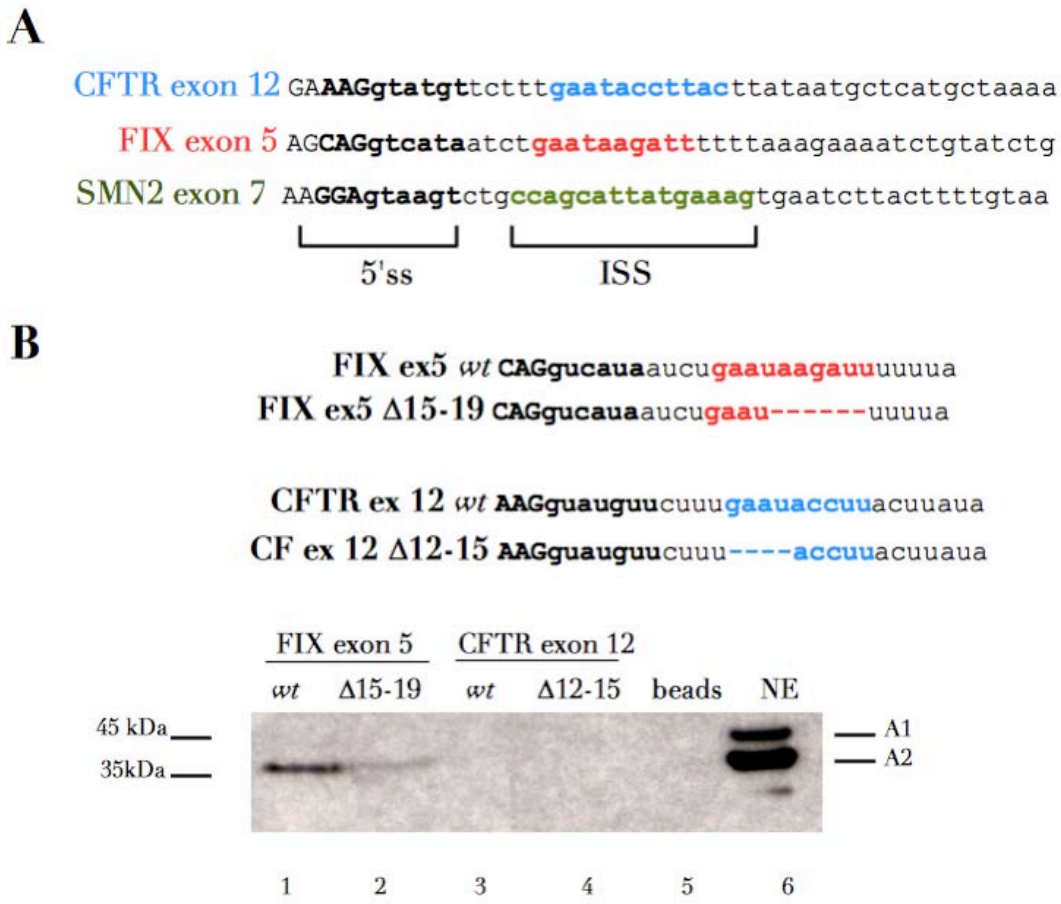

Figure 2.9.1. **hnRNP A1/A2 bind to *F9* exon 5 ISS**

A, Schematic representation of *CFTR* exon 12, *F9* exon 5 and *SMN2* exon 7 donor splice sites and downstream intronic regions. These three unrelated exons have an intronic splicing silencer region (coloured) located right downstream of their donor splice sites (bolded). Uppercase letters are exonic sequences; lowercase letters, intronic sequences.

B, Western blot against hnRNP A1/A2 upon RNA pulldown using *F9* exon 5 *wt*, Δ15-19 and *CFTR* exon 12 *wt* and Δ12-15 probes. hnRNP A1/A2 binds to the intronic splicing silencer located downstream of *F9* exon 5 donor splice site. When bases +15 to +19 are absent, hnRNP A1/A2 binding is impaired. Donor splice sites are bolded, ISS sequences are coloured and dashes represent the deleted bases. NE, nuclear extract.

Section 2.10. ***Fine mapping of* CFTR *exon 12 ISS reveals a potential role of RNA secondary structure on donor site accessibility.***

To further investigate the functional properties of the ISS located in *CFTR* exon 12, I carried out a site-directed mutagenesis of the element (Fig. 2.10.1A). In the pCF ex12 *wt* minigene, I introduced several point mutations across the entire *CFTR* exon 12 ISS, from position +12 to +23 (Fig.2.10.1A). Upon transfection of the ISS mutant minigenes into HeLa cells, I analysed the resulting splicing pattern. Substitutions from positions +13 to +19 clearly altered the splicing pattern (Fig. 2.10.1, lanes 3-10), while no changes were observed for substitutions in position +12 and from position +20 to +23 (Fig. 2.10.1, lanes 3, 11-16). Consistent with the presence of an ISS, substitutions from position +16 to +19 increased *CFTR* exon 12 inclusion from ~80% to almost 100%, (Fig. 2.10.1, lanes 5-10). However, two nearby mutations showed an opposite effect. The substitution of an A for a C at position +13 led to complete *CFTR* exon 12 skipping and the T to C substitution at position +15 reduced exon inclusion to 50% (Fig. 2.10.1 lanes 3 and 4).

To understand why these mutations had a silencing effect distinct from the other nearby substitutions, I considered the possibility that some mutations might affect the RNA secondary structure. To this aim I analysed a stretch of 60 bp covering both upstream and downstream sequences around *CFTR* exon 12 ISS using RNA mfold, a web-based bioinformatic tool designed for RNA secondary structure predictions (Markham et al. 2008). (Fig. 2.10.2). The predicted RNA secondary structure showed a single stem-loop spanning from position -3 to +20 that include the donor splice site. Most of the 5' ss is part of the stem and the core GU of the donor site base pairs to positions +16 and +17. Bases from +12 to +20, which compose most of the *CFTR* exon 12 ISS, are predicted to constitute the rest of the stem-loop structure.

Interestingly, the majority of the mutations that induce exon inclusion are

predicted to disrupt the stem-loop structure, whereas the mutations that induce exon skipping strengthen the stem. Mutation +13C generates a new bond with the G located at position +5, which in turn facilitated the formation of a non Watson-Crick bond between a U in position +6 and a G in position +12. This results in a longer and more stable stem structure which might limit the accessibility of the 5' ss recognition by U1 snRNP (Fig. 2.10.3A). This observation could explain the complete exon skipping detected upon transfection of pCF ex12 +13C minigene (Fig. 2.10.1B, C, lane 3). Mutation +15C was also predicted to affect RNA secondary structure. In this case, the predicted structure consists of a long stem, which bears a bulge between position +2 and +5 of *CFTR* exon 12. The C at position +15 was predicted to interact with the *wt* G at position +5 (Fig. 2.10.3B). This in turn facilitate the formation of three Watson-Crick bonds: between two U in positions +6 and +7 and a two A in position +14 and +13, respective and a between a C at position +8 and a G at position +12.

On the other hand, substitutions at positions +17, +18 and +19 introduced mismatches into the stem, turning the whole stem-loop into a less stable structure (Fig. 2.10.4). Importantly, all these substitutions promoted exon inclusion (Fig. 2.10.1B, lanes 7-10). In position +16, I replaced the *wt* A with either a mutant G or a mutant T (Fig. 2.10.1A and B). Mutation +16T created a mismatch while mutation +16G can form a non-Watson-Crick bond (Fig- 10.4A and B). These observations are consistent with the fact that, although both substitutions promoted exon 12 inclusion, the former had a stronger impact on exon inclusion, as can be distinguished in Fig. 2.10.1B and C (lanes 5 and 6). Finally, bases +20 to +23 were not predicted to be involved in the stem-loop structure (Fig. 2.10.4F), and did not alter the *wt* splicing pattern (Fig. 2.10.1B, lanes 11-16). Altogether these observations indicate that a stem-loop structure at *CFTR* exon 12 5' ss might be involved in splicing regulation.

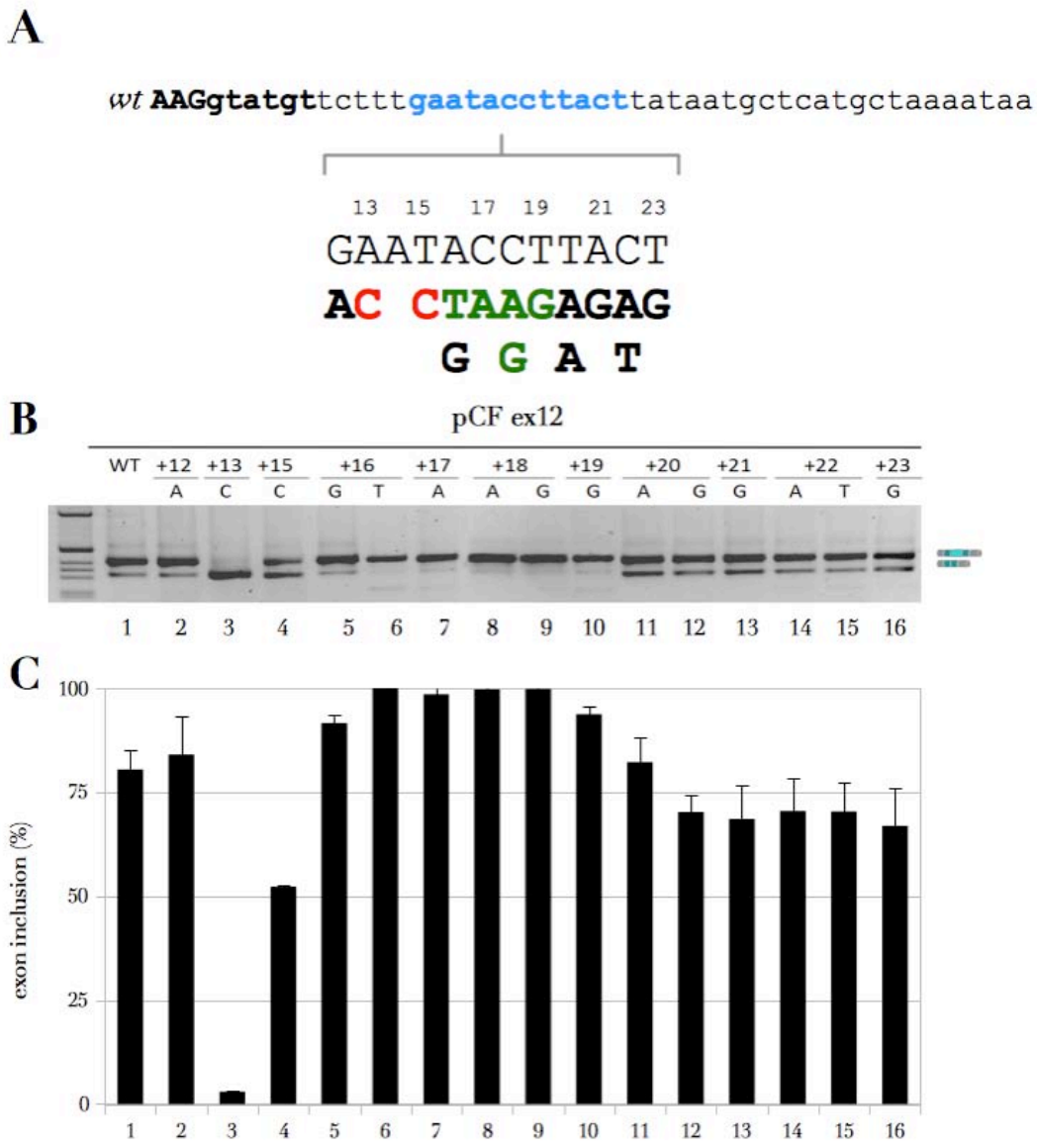

Figure 2.10.1. **Mutagenesis of *CFTR* exon 12 intronic splicing silencer element (ISS)**
A, Schematic representation of the *CFTR* exon 12 *wt* 5' ss and downstream intronic sequences, where ISS is coloured and 5' ss is bolded. Below are shown the point mutations introduced by overlap extension PCR mutagenesis into the ISS. Numbers above ISS indicate base position from 5' ss (bolded). Substitutions in red caused exon 12 skipping, those in green promoted exon inclusion and those bolded did not affect splicing pattern as shown in B.
B, Splicing pattern analysis of *CFTR* exon 12 ISS mutants. ISS mutant minigenes (0.5 ug) were transfected into HeLa cells and subsequently splicing products were analyzed by RT-PCR on total RNA using alpha2,3 and bra2 oligonucleotides. PCR products were loaded onto a 2% agarose gel and visualized after electrophoresis. ISS mutants displayed mixed effects on splicing pattern, ranging from complete (+13C, lane 3) or partial (+15C, lane 4) exon 12 skipping to increased exon inclusion (+16G and T, +17A, +18A and G, +19G; lanes 5-11) or simply showing no effect (+20A and G, +21G, +22A and T, +23G; lanes 12-16).
C, Quantification of exon 12 inclusion of ISS mutants. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

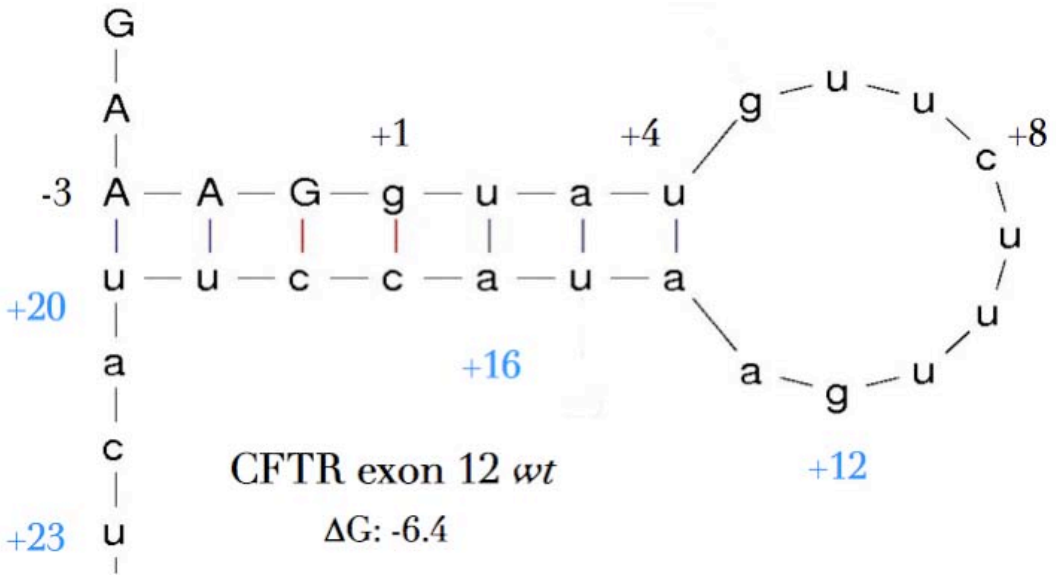

***Figure 2.10.2.*Predicted stem-loop structure for *CFTR* exon 12 *wt* 5' ss and downstream intronic region**

A sequence stretch of 60 bp around *CFTR* exon 12 *wt* donor splice site was subjected to RNA mfold analysis. *CFTR* exon 12 ISS is predicted to form part of an stem-loop structure spanning from base -3 to +20. The invariant GU pair is embedded in the stem structure, forming Watson-Crick bonds with AC pair at positions +16 and +17. Bases +12 to +20, which form part of the *CFTR* exon 12 ISS, are predicted to constitute part of the loop (bases +12 and +13) and the stem (+14 to +20). Bases +21 to +23 would not belong to an RNA secondary structure. Exonic sequences are in upper case letters and intronic sequences in lowercase. Position of bases with respect to *CFTR* exon 12 5'ss is indicated. Coloured numbers indicate the position of the ISS sequence.

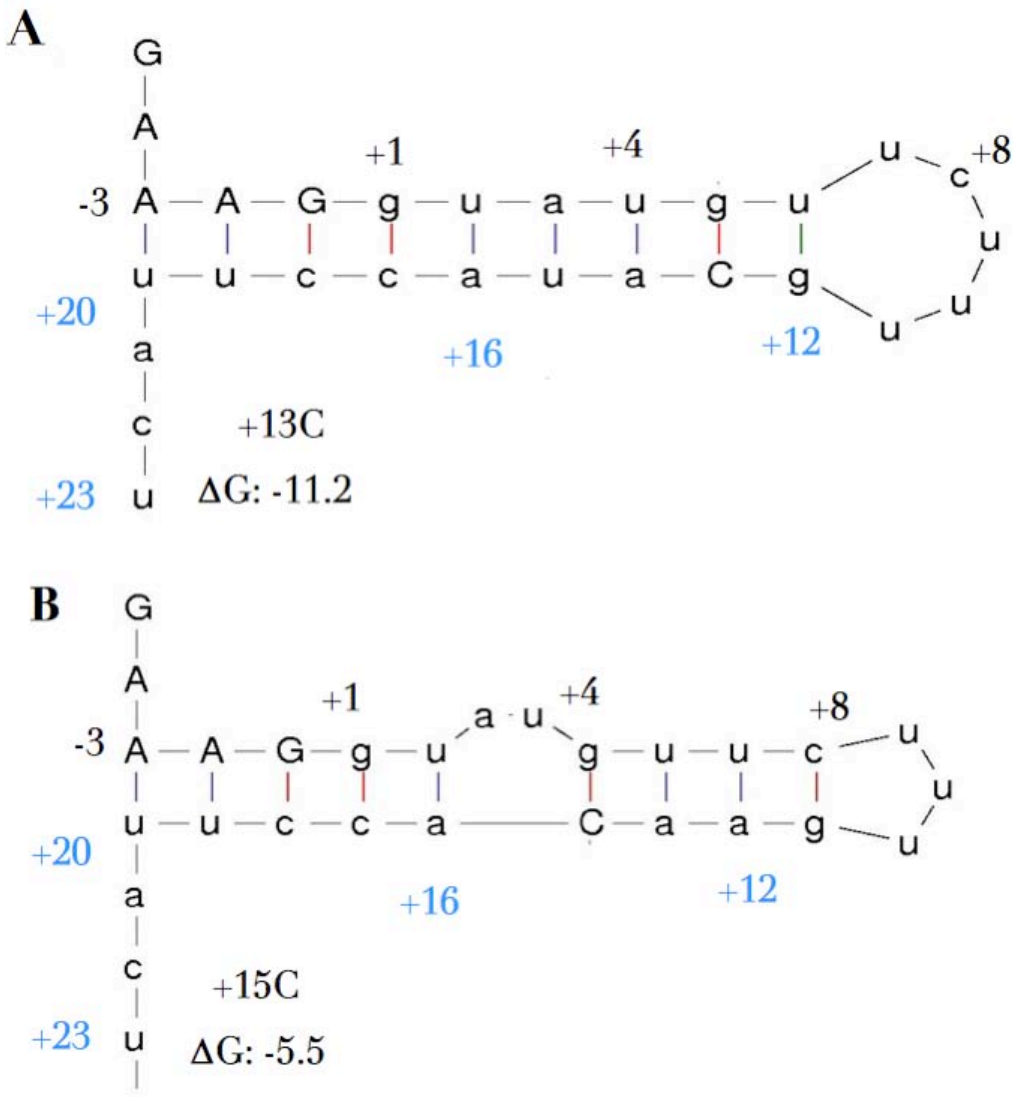

Figure 2.10.3. **Predicted RNA secondary structure for *CFTR* exon 12 mutants +13C and +15C**
A sequence stretch of 60 bp around *CFTR* exon 12 donor splice site including either substitution +13C (**A**), or +15C (**B**) within ISS was subjected to of RNA mfold analysis.
A, The predicted RNA secondary structure for mutant +13C is more stable than *wt* stem-loop(compare with Fig. 2.10.2) since the mutant cytosine at position +13 creates a bond between with the *wt* G at position +5. This in turn facilitates the formation of a non Watson-Crick bond between *wt* G at position +12 with U at position +6.
B, The predicted structure for mutant +13C consists of a long stem which bears a bulge between position +2 and +5 of *CFTR* exon 12. The cytosine at position +15 is predicted to interact with the *wt* guanidine at position +5. This in turn facilitates the formation of three Watson-Crick bonds: between two U in positions +6 and +7 and a two A in position +14 and +13, respective and a between a C at position +8 and a G at position +12.

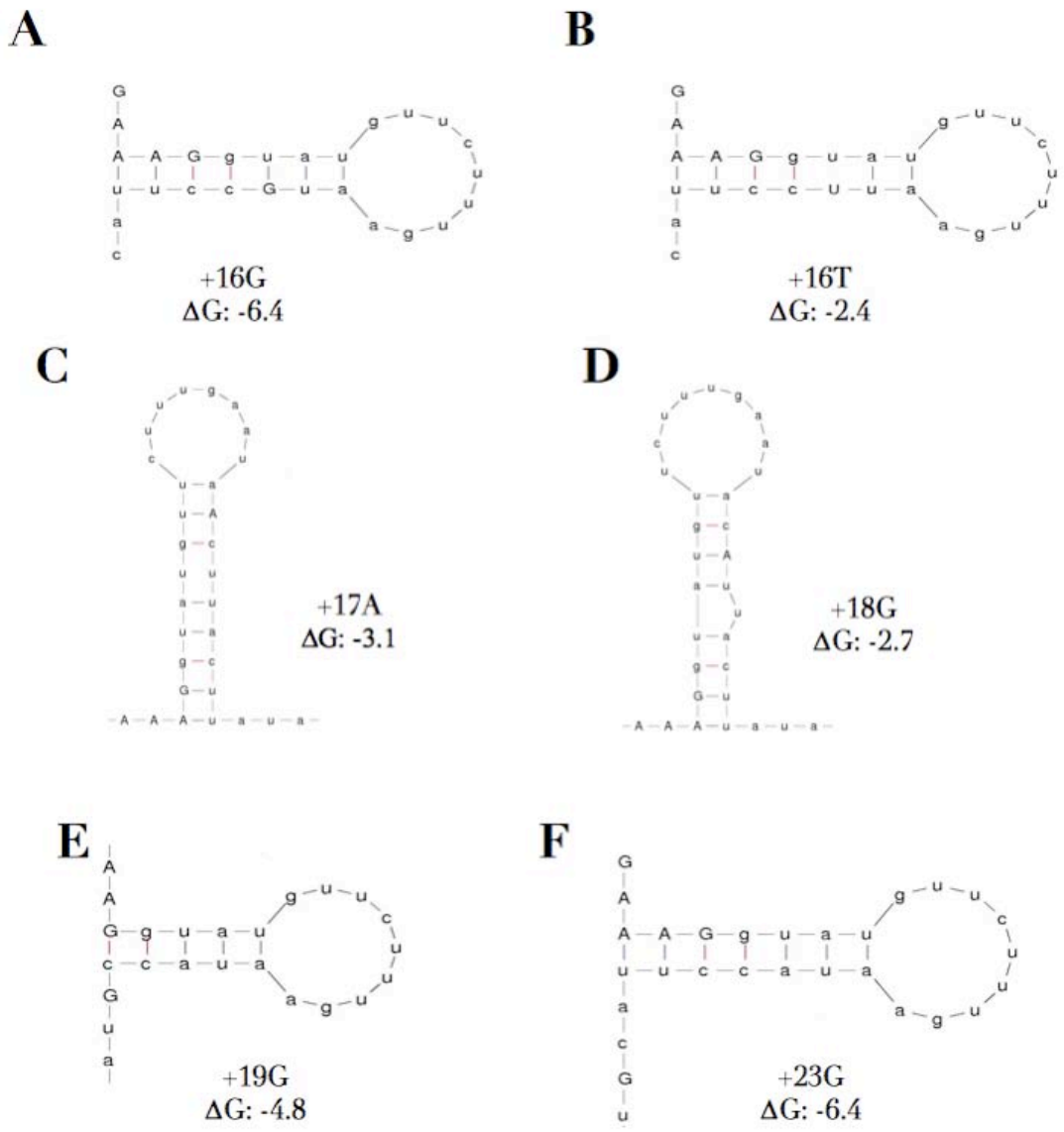

***Figure 2.10.4*. Predicted RNA secondary structure for *CFTR* exon 12 ISS mutants**

A sequence stretch of 60 bp around *CFTR* exon 12 donor splice site including several substitutions within the ISS element was subjected to RNA mfold analysis.

A, Mutation +16G is predicted to maintain *wt* stem-loopstructure since the mutant G is able to form a non Watson-Crick bond with the *wt* U at position +2.

B, **C** and **D**, Mutations +16T, +17A and +18G introduce mismatches into the stem structure and their predicted RNA secondary structure stability is lower than *wt* one.

E, Mutation +19G shortens stem and its predicted stability is lower than *wt* one.

F, Mutation +23G do not alter the predicted *wt* stem-loopstructure.

ISS mutations are shown in uppercase letters.

Section 2.11. ***Effect of antisense oligonucleotides targeted to the*** **F9** ***exon 5 and*** **CFTR** ***exon 12 ISS.***

In *CFTR* exon 12 and *F9* exon 5, most of the enhancing effect of the ExSpeU1s occurred when they target the ISS sequences suggesting that part of their activity could be due to a direct interference with these intronic inhibitory elements.

I first designed 2'-O-methoxyethyl-phosphorothiate antisense oligonucleotides (ASO) (synthesized by Daniela Perrone, University of Ferrara) to target *CFTR* exon 12 and *F9* exon 5 ISSs, namely ASO^{FIX} and ASO^{CFTR} respectively (Fig. 2.11.1A). For each ISS, I designed two ASOs, one of 20 bp and another of 25 bp, which cover the entire ISS sequence. Next I analysed their effect on exon processing in both *F9* exon 5 donor site mutant -2C and *CFTR* exon 12 exonic mutant 17A minigenes in a co-transfection experiment. Two different ASO concentrations, 1 and 10 μM were tested for both mutant minigenes. As a positive control, I also co-transfected, together with ASO 10-27, the *SMN2* exon 7 minigene, pSMN2, which was previously described to promote exon inclusion by targeting *SMN2* exon 7 ISS (Hua et al. 2008).

In control experiments ASO10-27 (10 μM) induced a significant increase in the percentage of *SMN2* exon 7 inclusion, rising it to almost 100% (Fig. 2.11.1B, lanes 1 and 2). In the *F9* exon 5 minigene, neither ASO^{FIX}_{20} nor ASO^{FIX}_{25} affected significantly the splicing pattern of the *F9* exon 5 -2C minigene. In the *CFTR* exon 12 system, the addition of ASO^{CFTR}_{20} did not affect the pattern of splicing of pCF ex12 17A minigene (Fig. 2.11.1C, lanes 2-5). On the other hand, ASO^{CFTR}_{25} at a moderate effect on pCF ex12 17A minigene splicing pattern. At the concentration of 10 μM, ASO^{CFTR}_{25} increased significantly exon 12 inclusion levels of two folds, rising it from 10% to 20% (Fig. 2.11.1C, lane 9). The slight increase in exon 12 inclusion in presence of ASO^{CFTR}_{25} prompted me to test additional concentrations, ranging from 10 to 75 μM (Fig. 2.11.2A). A modest

rise in *CFTR* exon 12 inclusion was observed with increasing ASO^{CFTR}_{25} concentration, peaking at 50 μM (~27%; Fig. 2.11.2A,B lane 5) while adding 75 μM ASO^{CFTR}_{25} did not further improve in *CFTR* exon 12 inclusion (~21%; Fig. 2.11.2A,B lane 6) . Both 50 and 75 μM ASO^{CFTR}_{25} treatments also caused aberrant splicing events, as evidenced by the presence of higher molecular weight bands, which were more pronounced at 75 μM concentration (compare Fig. 2.11.2A, B lanes 5 and 6 with Fig. 2.11.1C lane 10). Sequencing analysis of these high molecular bands confirmed the occurrence of intron retention. Thus, only ASO^{CFTR}_{25} had a positive mild effect on exon inclusion, with an apparent dose-dependent efficacy.

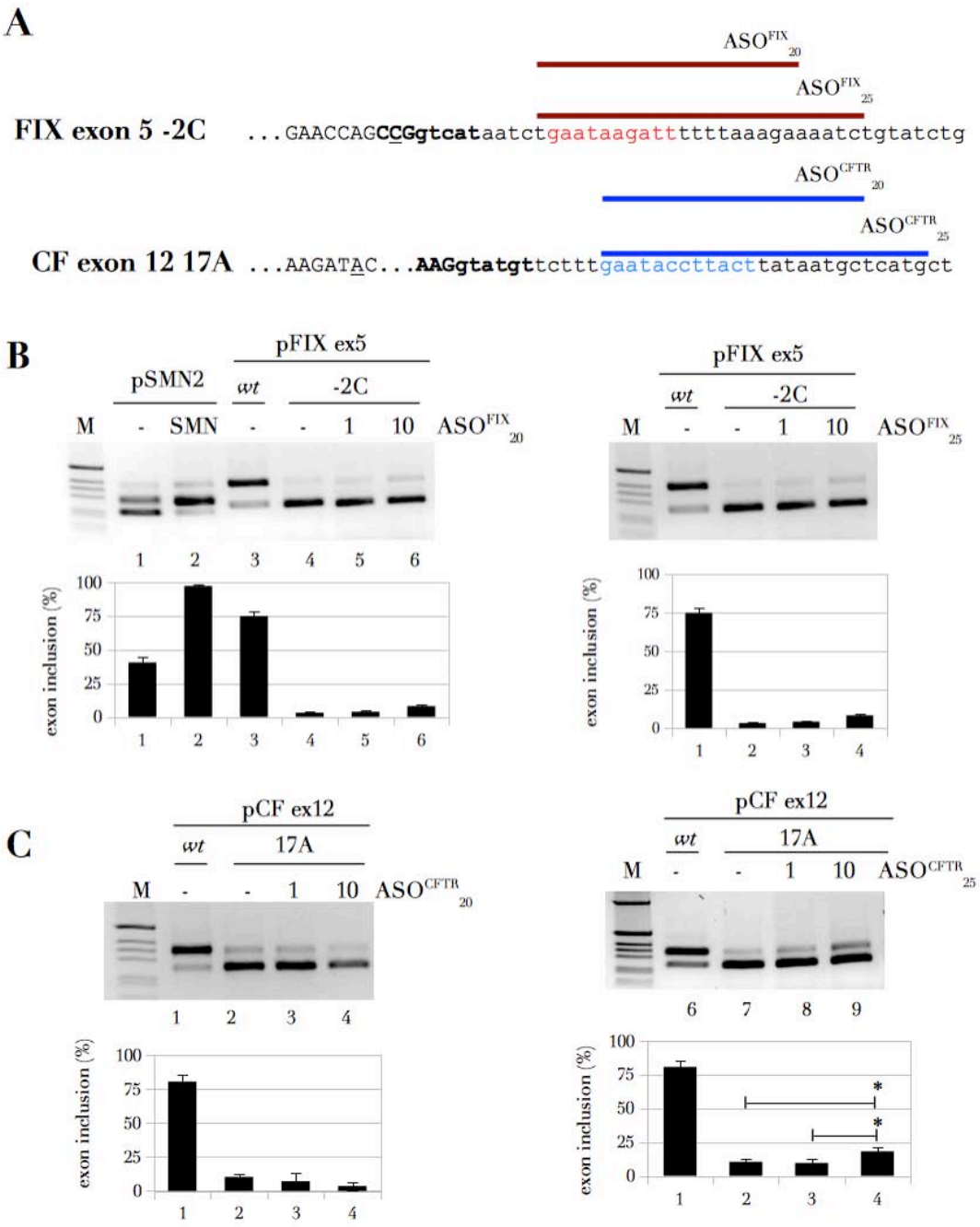

Figure 2.11.1. **Effect of antisense oligonucleotides directed to *F9* exon 5 and *CFTR* exon 12 ISSs**
A, Antisense oligonucleotides (ASO) were designed to target downstream sequences of *F9* exon 5 and *CFTR* exon 12 donor splice sites, (ASO^{FIX} and ASO^{CFTR}, respectively) corresponding to ISSs. Sequences of *F9* exon 5 -2C and *CFTR* exon 12 17A mutants are depicted; donor splice site sequences are bolded, and ISSs are coloured. Mutations -2C and 17A are underlined. Uppercase letters, exonic sequences; lowercase letters, intronic sequences.
B and **C**, Splicing pattern analysis of pFIX ex5 -2C (**B**) and pCF ex12 17A (**C**) minigenes, transfected either alone or with increasing amounts of ASOs (1, 10 and 100 µM) into HeLa cells. Subsequently splicing products were analyzed as described previously. *F9* exon 5 and *CFTR* exon 12 *wt* minigenes were also transfected alone as controls. Only ASO^{CFTR}_{25} induced a modest although significant rise of *CFTR* exon 12 inclusion at 10 µM. M, molecular weight marker (1kB, Invitrogen). Below each agarose gel picture, the histogram of the quantification of exon inclusion is shown. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate. *, $p<0.05$.

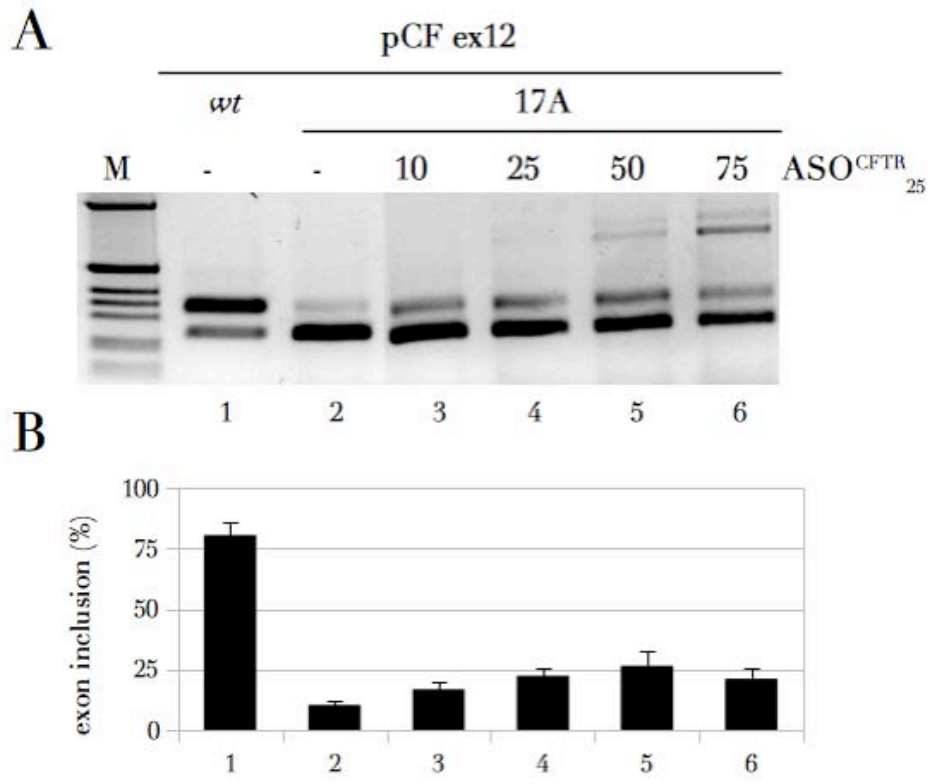

Figure 2.11.2. **ASO$^{CFTR}_{25}$ slightly promoted *CFTR* exon 12 inclusion**
A, Splicing pattern analysis of pCF ex12 17A, transfected either alone or in presence of increasing amounts (10, 25, 50 and 75 µM) of ASO$^{CFTR}_{25}$ into HeLa cells. Subsequently splicing products were analyzed as described previously. *CFTR* exon 12 inclusion levels increased modestly with higher concentrations of ASO$^{CFTR}_{25}$. At 50 and 75 µM some aberrant splicing events were observed (lanes 5 and 6). M, molecular weight marker (1kB, Invitrogen).
B, Quantification of exon 12 inclusion. Exon inclusion percentage was estimated using ImageJ software and is expressed as means ± SD, based on at least three independent experiments done in duplicate.

Section 2.12. ***Use of ExSpeU1FIX 9 to rescue protein production and activity on a* F9 *exon 5 splicing competent cDNA expression vector***

The previously reported results have clearly shown that ExSpeU1s can correct aberrant splicing in the context of *CFTR* exon 12 and *F9* exon 5 hybrid minigenes (see Section 2.7-8). To understand if ExSpeU1s can also rescue protein synthesis, I designed a splicing competent *F9* cDNA expression vector (pSC-FIX exon 5) as illustrated in Fig. 2.12.1. This expression vector includes exon 5 and its flanking intronic sequences inserted in the factor IX cDNA. Correct splicing processing of the upstream and downstream introns flanking exon 5 give rise to full-length *F9* mRNA coding for normal coagulation factor IX. In this system, a unique *Nde*I restriction site is located in the intronic sequences to allow efficient subcloning of exon 5 *wt* or mutant cassettes (see Materials and Methods). As previously shown, U1FIXwt and ExSpeU1FIX 9 rescued splicing at donor site mutants -2C, -2G and -2T and at the PPT mutants -8G and -9G (see Section 2.4). Based on these results, I engineered five mutant *F9* exon 5 splicing competent expression vectors: pSC-FIX exon 5^{-2C} , exon 5^{-2G} and exon 5^{-2T} (bearing donor site mutations -2C, -2G and -2T, respectively) and pSC-FIX exon $5^{PPT\text{-}8G}$ and exon $5^{PPT\text{-}9G}$ (carrying PPT mutations -8G and -9G, respectively). These five constructs were co-transfected with each modified U1 snRNAs into BHK cells, and splicing pattern analysis was carried out by RT-PCR. Amplification of pSC-FIX exon 5^{wt} transcripts using a forward oligonucleotide annealing on *F9* exon 4 (FIXex4F) and a reverse oligonucleotide binding to exon 6 (FIXex6R) showed the presence of two bands, one corresponding to exon 5 inclusion (604 bp) and one to exon 5 skipping (475 bp) (Fig. 2.12.1B, lane 1). In particular, the quantification of the bands intensity revealed ~15% of exon inclusion for pSC-FIX exon 5^{wt}. This percentage is significantly lower than the one found in normal liver (~80%) or the one obtained by the pFIX ex5 *wt* minigene (~75%); (compare Fig. 2.12.1B

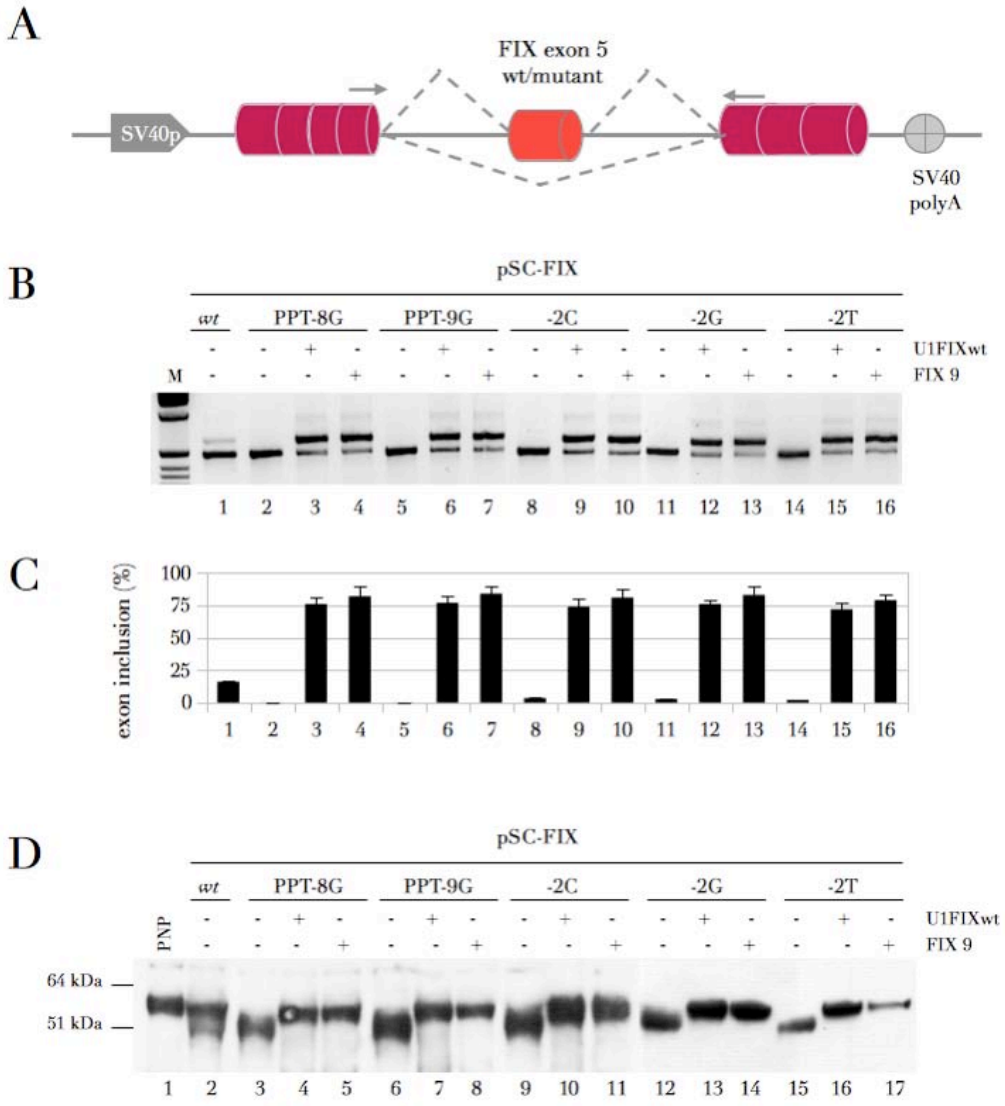

Figure 2.12.1. **Rescue of FIX full-length pre-mRNA and protein secretion by U1FIXwt and ExSpe U1FIX 9 in BHK cells.**

A, Schematic representation of a splicing competent *F9* exon 5 expression vector (pSC-FIX exon 5). Dashed lines (- - -) illustrate possible splicing outcome. Arrows represent primers for PCR amplification. For a detailed description of pSC-FIX exon 5 construction, see Materials & methods.

B, Analysis of spliced transcripts. BHK cells were transfected with 0.5 µg of pSC-FIX exon 5^{wt}, pSC-FIX exon $5^{PPT\text{-}8G}$, pSC-FIX exon $5^{PPT\text{-}9G}$, pSC-FIX exon 5^{-2G}, pSC-FIX exon 5^{-2C} or pSC-FIX exon 5^{-2T} either alone or with U1FIXwt or ExSpe U1FIX 9. Splicing pattern was evaluated by RT-PCR withFIXex4R and FIXex6F primers and amplified products were resolved on a 2% agarose gel. The identity of the bands is indicated on the right side of the panel. In presence of either U1FIXwt or U1FIX 9, exon 5 inclusion was increased notably, reaching similar levels in the three mutant contexts evaluated.

C, Quantification of exon 5 inclusion was estimated using ImageJ software and is expressed as means ± SD.

D, Western Blotting against factor IX protein on supernatant of BHK cells transfected with pSC-FIX exon 5^{wt}, pSC-FIX exon $5^{PPT\text{-}8G}$, pSC-FIX exon $5^{PPT\text{-}9G}$, pSC-FIX exon 5^{-2G}, pSC-FIX exon 5^{-2C} or pSC-FIX exon 5^{-2T} either alone (lanes 2, 3, 6 and 9) or with U1FIXwt (lanes 4, 7, 10) or ExSpeU1FIX 9 (lanes 5, 8, 11). As a control, pooled normal pooled (PNP, lane 1) from healthy patients was used for WB. In presence of U1FIXwt or ExSpeU1FIX 9 secreted full-length factor IX protein is rescued in all three mutants contexts.

with Fig. 2.2.3). Therefore, it can be speculated that context-dependent factors could be involved in the lower efficiency of recognition of *wt* exon 5 observed with the pSC-FIX minigene. As expected, splicing pattern of pSC-FIX exon 5^{PPT-8G}, exon 5^{PPT-9G}, exon 5^{-2C}, exon 5^{-2G} and exon 5^{-2T}minigenes resulted in complete exon 5 skipping (Fig. 2.12.1B, lanes 2, 5, 8, 11, and 14 respectively). In the five mutants, co-transfection of either U1FIXwt or EXSpeU1FIX 9 significantly ameliorated splicing, with ~70-80% of exon included in the final mRNA (Fig. 2.12.1B, lanes 3-4, 6-7, 9-10, 12-13, 16-17). Thus, an important amount of normal full-length *F9* mRNA is produced upon U1FIXwt or ExSpeU1FIX 9 co-transfection. Given that the factor IX protein is normally secreted, it can be recovered from the culture supernatants and quantified. Therefore, I analysed *F9* exon 5 levels in cell culture supernatants using a polyclonal antibody against factor IX protein. In this experimental approach, alternative splicing of exon 5 can give rise to two factor IX proteins: the functionally normal one, which includes exon 5 (~60 kDa), and a shorter one, which is inactive and corresponds to transcripts without exon 5 and lacking the EGF-2 domain (~51 kDa). In fact, skipping of exon 5 maintain in frame the protein-coding sequence leading to a 43 aminoacids shorter product. As control, I used pooled normal plasma (PNP) from healthy patients, which contains normal levels of factor IX protein. As can be appreciated in Fig. 2.12.1D, pSC-FIX $exon5^{wt}$ generated both factor IX protein forms: normal factor IX protein (lane 2, upper band) and truncated factor IX protein (lane 2, lower band). Normal, full-length proteins were more abundant than truncated factor IX proteins, in contrast to *F9* splicing pattern, where the latter was more abundant (compare Fig.12.1B, lane 1 with Fig. 2.12.1D, lane 2). Possibly, part of the truncated protein is degraded in the cytoplasm or is inefficiently secreted due to aberrant folding providing an explanation for the discrepancy observed between WB analysis and splicing pattern results. On the other hand, transfection of pSC-FIX exon 5^{PPT-8G}, exon 5^{PPT-9G}, exon 5^{-2C}, exon 5^{-2G} and exon 5^{-2T} minigenes led to the

production of only the ~51 kDa form (Fig. 2.12.1D, lanes 3, 6, 9, 12, 15, respectively).

Strikingly, for the three pSC-FIX exon 5 mutants the presence of either U1FIXwt or U1FIX 9 completely restored the full-length factor IX protein (Fig. 2.12.1D,lanes 4-5, 7-8 and 10-11,13-14, 16.17). To quantify the secreted factor IX protein levels I carried out a sandwich ELISA assay. The antibody used cannot discriminate between the full-length active isoform and the shorter inactive form. The analysis showed an ~2-3 fold increase of factor IX protein levels in

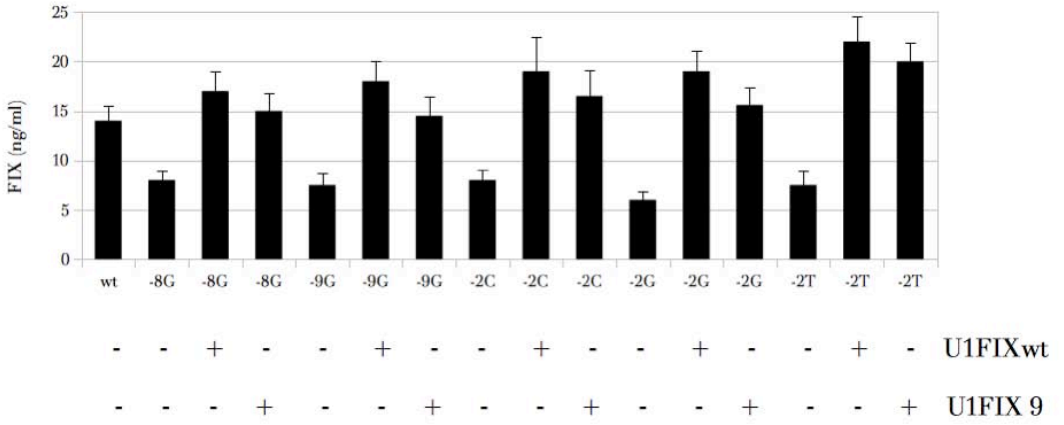

Figure 2.12.2. **Rescue of secreted FIX protein levels by modified U1FIX snRNAs in BHK cells supernatant**

ELISA assay showing concentration of secreted factor IX protein in supernatant of BHK cells transfected with pSC-FIX exon 5^{wt}, pSC-FIX exon $5^{PPT\text{-}8G}$, pSC-FIX exon $5^{PPT\text{-}9G}$, pSC-FIX exon 5^{-2G}, pSC-FIX exon 5^{-2C} or pSC-FIX exon 5^{-2T} either alone or with U1FIXwt or U1FIX 9. Presence of modified U1FIX snRNAs increased notably factor IX protein levels in *BHK* cells supernatant.

supernatants of cells transfected with pSC-FIX exon 5 mutant vectors and U1FIXwt or U1FIX 9 (Fig. 2.12.2). Finally, I tested secreted factor IX protein activity through a factor IX-specific activity assay (see Materials and Methods). As expected, no factor IX activity was detected when pSC-FIX exon $5^{PPT\text{-}8G}$, exon $5^{PPT\text{-}9G}$, exon 5^{-2C}, exon 5^{-2G} and exon 5^{-2T} were transfected alone, since truncated factor IX protein undergoes structural rearrangements that greatly impair its proper folding and secretion. Factor IX activity was recovered nearly to *wt* levels in cell supernatant where mutant expression vectors were co-transfected

with either U1FIXwt or ExSpeU1FIX 9 (Fig. 2.12.3). Taken together these results demonstrated that U1FIXwt and exon-specific U1FIX 9 are able to correct *F9* exon 5 aberrant splicing caused by both PPT and donor splice site mutations at positions -2 in a splicing competent expression vector system, re-establishing protein synthesis, secretion and activity.

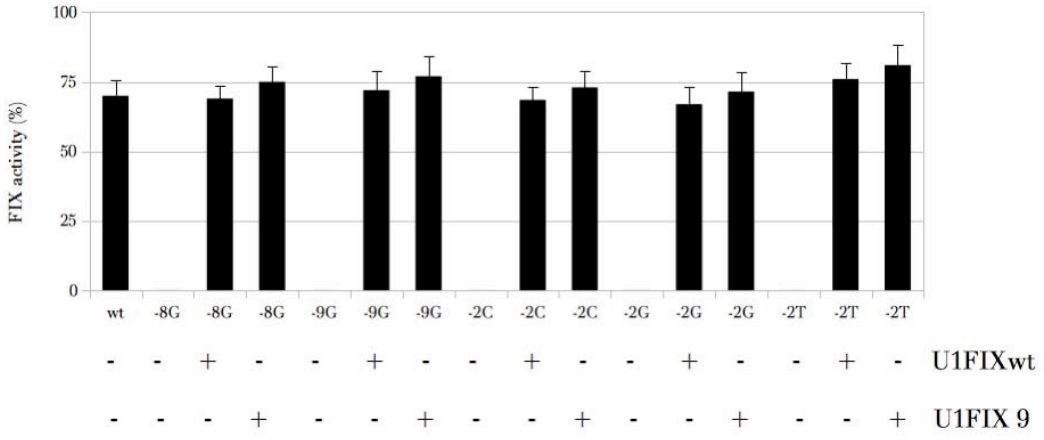

Figure 2.12.3. **Rescue of secreted FIX protein specific activity by modified U1FIX snRNAs in BHK cells supernatant**
Factor IX specific activity of secreted factor IX protein in supernatant of BHK cells transfected with pSC-FIX exon 5^{wt}, pSC-FIX exon $5^{PPT\text{-}8G}$, pSC-FIX exon $5^{PPT\text{-}9G}$, pSC-FIX exon 5^{-2G}, pSC-FIX exon 5^{-2C} or pSC-FIX exon 5^{-2T} either alone or with U1FIXwt or U1FIX 9. Presence of modified U1FIX snRNAs recovered notably factor IX protein specific activity in BHK cells supernatant.

3. Discussion

Section 3.1. **CFTR *exon 12 and* F9 *exon 5 mutations induce aberrant splicing***

In this thesis work, I have used two gene models, *CFTR* exon 12 and *F9* exon 5, to analyse the effect on pre-mRNA splicing of several disease-associated genome variants (GVs) and to evaluate the role of U1 snRNP in the resulting splicing defects. I have studied three major types of mutations: transversions within the polypyrimide tract (PPT), mutations in exonic regulatory elements and mutations that affect the consensus 5' ss.

To understand the effect of these mutations on splicing I used the minigene splicing approach. I have used as backbone for the heterologous minigene the pTB construct, which has been previously validated for the CFTR exon 12 (Pagani et al. 2000, 2002, 2003a) and for other gene systems (Baralle et al. 2003; Zhang et al. 2004; Esperante et al. 2008; Goina et al. 2008; Skoko et al. 2008; Pastor et al. 2009; Raponi et al. 2009). This backbone was also useful for the study of *F9* exon 5, as the pFIX ex5 *wt* minigene is processed *in vitro* as efficiently as it is *in vivo*. Actually, in both cases the exon 5 *wt* is alternatively spliced, with approximately 20% or the exon excluded from the mature mRNA (Fig. 2.2.2).

Minigene experiments showed that several mutations induced exon skipping. In *F9* exon 5, the two PPT transversions (-8T>G and -9T>G) and the donor splice site mutations in position -2 (A to G, C and T), -1T, +1A and +2C (Fig. 2. 2.3B, lanes 1-7) induced exon skipping. Notably, the +4G did not change the pattern of splicing (Fig. 2.2.3B, lane 8) whereas the +13G activated a cryptic donor site (Fig. 2.2.3B, lane 9). In *CFTR* exon 12, all mutations either in the exon (17A and 52T) or at the donor site induced exon skipping (Fig. 2.2.2.1B, lanes 1-9). Exclusion of *CFTR* exon 12 from the mature mRNA maintains the

reading frame but produces an aberrant protein lacking part of the NBD-1, with severe folding and secretion abnormalities (Slomski et al. 1992; Tzetis et al. 2001). In turn, skipping of *F9* exon 5 also maintains the reading frame and will lead to a formation of a protein lacking a fundamental EGF-2 domain, with consequent loss of functionality of the protein (Wilkinson et al. 2002; Schmidt et al. 2003). Similarly, the new cryptic site due to the A to G substitution at position +13 downstream the 5' ss will create a pathological transcript 13 bp longer than the *wt* one with a shift of the reading frame and the production of a truncated protein (Ketterling et al. 1999).

It is noteworthy that without an appropriate splicing assay the synonymous mutations located at position -2 of *F9* exon 5' ss might be erroneously identified as benign variants. This analysis highlights the importance of the minigene-based strategy as diagnostic tool to evaluate the effect of genomic variants on splicing.

Section 3.2. ***Bioinformatic predictions do not always correlate with functional splicing assays***

Several *in silico* tools have been designed to predict the impact of GVs occurring within donor and acceptor splice sites and establish their pathogenicity. Most of the available bioinformatic prediction tools assign a score value to the input sequence, which is regarded as a predictor of splice site strength or "quality" (i.e., consensus donor site sequence has the maximum score). Therefore, I have evaluated the score predictions for the aforementioned *CFTR* exon 12 donor site mutations and *F9* exon 5 PPT and donor site mutations using 11 web-based tools (see Table 2.1.2). For exonic GVs, some bioinformatic programs such as ESEFinder (Cartegni et al. 2003), RESCUE-ESE (Fairbrother et al. 2004) or PESX (Zhang et al. 2004) can be used to estimate whether they creation or disrupt putative exonic splicing regulatory sequences. I have subjected *CFTR* exon 12 exonic mutants 17A and 52T sequence to analysis with these softwares but no clear correlation between their impact on splicing and the predicted sites was found (data not shown), in agreement with previous studies (Raponi et al. 2007).

The very low score (0.21) of the normal *wt F9* exon 5, could itself be considered pathological (Table 2.1.2). However, despite the low score the exon 5 *wt* is predominantly included in the mature mRNA (Fig. 2.2.3, lane 1). It is likely that the natural A at position +5 of *F9* exon 5 *wt* donor site penalizes its score, since it is known that a G is conserved at that position in the ~78% of U2-dependent, human 5' GT splice sites (Sahashi et al. 2007; Roca et al. 2008). Additionally, the C at position +3 significantly deviates from consensus, with only ~3% of conservation (versus ~60% of consensus A) as well as the A at position +6 (~17% vs. ~48% of consensus T). Consequently, other mechanisms might compensate for its weak donor splice site. It is well-known that both exonic and intronic *cis*-acting regulatory sequences play a key role in exon definition

(Mardon et al. 1987; Cartegni et al. 2002). In addition, it has been proposed that the presence of a strong 3' ss modulates the impact of GVs occurring in 5' ss(Liu et al. 2011). This is supported by previous observations showing that secondary mutations that promote a better agreement with the consensus sequence on acceptor sites suppressed the effect of primary mutations on donor splice sites (Carothers et al. 1993). Furthermore, the near equal contribution of both acceptor and donor splice sites to the definition of constitutive and alternative exons has been recently proven through minigene-based analysis and a genome-wide survey (Shepard et al. 2011). Indeed, *F9* exon 5 *wt* 3' ss score is very similar to the consensus (Table 2.1.2). It is reasonable that this strong *F9* exon 5 acceptor site compensate for the weak donor splice site, although other exonic and/or intronic splicing regulatory elements might also be involved. Further studies will be required to identify additional splicing regulatory sequences involved in *F9* exon 5 definition.

Despite the low score of *F9* exon 5 *wt* 5' ss, the impact of most of the donor site mutations were consistently predicted. Score predictions for the invariant GT-destroying mutations (*CFTR* exon 12 +2C and *F9* exon 5 +1A and +2C), PPT (*F9* exon 5 -8G and -9G) and donor site mutations (*CFTR* exon 12 -1A, -1T, +3C, +5A and *F9* exon 5 -2T, +2G, +2C and -1T) and the cryptic donor splice site created by mutation +13G in *F9* exon 5 agreed with the minigene experiments. However, two interesting cases exhibited a clear discrepancy between score predictions and splicing functional assays: mutations +3G on *CFTR* exon 12 and +4G on *F9* exon 5.

Mutation +3G within *CFTR* exon 12 donor site caused complete exon skipping (Fig. 2.2.1), notwithstanding its score values were only slightly reduced for 9 out of 11 *in silico* tools (Table 2.1.2). Based on nucleotide frequencies, human 5' ss have either an A (59.5%) or a G (34.6%) (Sahashi et al. 2007) at position +3. Thus, the pathogenicity of the +3G variant on *CFTR* exon 12 cannot be explained simply by conservation at this position. As a matter of fact the

bioinformatic tools (S&S, MDD, AST, HSF; Table 2.1.2) that assigned it a score value very similar to *wt* sequence are based on position-specific weight matrices (PSWM) which take into account relative nucleotide frequencies on each donor site position. In addition, a G at position +3 should be able to base pair with the corresponding pseudouridine on the U1 snRNP 5'-tail, which explains why the scores of Hbond (hydrogen bonding) and ΔG (thermodynamical stability of 5' ss:U1 snRNP duplex) are high. Based on statistical analysis it has been proposed the A to G transitions at position +3 are associated to aberrant splicing when located in weak donor sites (Madsen et al. 2006) and are usually accompanied by nonconsensus nucleotides at positions +4 and +5 (Roca et al. 2008). A similar observation was previously anticipated (Ohno et al. 1999), stating that in presence of a +3G variant normal splicing can take place when at least two of three nucleotides at positions +4 to +6 are complementary to the U1 snRNA. Furthermore, mutant human 5' ss containing an aberrant splicing-causing +3G along with a consensus +5G are most frequently accompanied by a nonconsensus T at position +4 (82.6%) (Guédard-Méreuze et al. 2009). Indeed *CFTR* exon 12 5' ss contains a T at position +4 and a consensus G at position +5. The results of splicing functional assays with pCF exon 12 +3G fully agree with these previous observations.

On the other hand, mutation +4G on *F9* exon 5 had no effect on splicing though its score value is low and similar to other *F9* exon 5 donor site mutants that cause exon skipping. Interestingly, only SplicePort, Hbond and ΔG methods assigned it a score similar to *wt* sequence. A G in position +4 although conserved only in ~12% of human donor sites (Sahashi et al. 2007; Roca et al. 2008) is still able to establish non-Watson-Crick bonds with the corresponding pseudouridine of U1 snRNP 5'-tail, which could explain the accurate performance of Hbond and ΔG methods. It remains unclear its association with the *F9* deficiency (Haemophilia B Mutation Database). Probably the presence of a guanidine in this position is classifiable as a rare

single nucleotide polymorphism (SNP), perhaps in linkage disequilibrium with another pathological mutation in the *F9* gene.

In an effort to distinguish neutral from splicing-disrupting GVs, Ketterling and colleagues (1999) proposed the "5-6" hypothesis. They classified donor splice sites in two categories: poorly conserved, with five or six matches to the consensus sequence and highly conserved, with at least seven matches. They argued that GVs occurring in poorly conserved donor splice sites (apart from invariant GT dinucleotide) would abolish splicing, while those affecting highly conserved donor sites would exhibit a moderate effect. This hypothesis was derived from a large analysis of eighty splicing-related mutations on *F9* gene and in consequence, it might be useful in our case to predict the effect of *F9* exon 5 mutations. However, both the effect of *F9* exon 5 mutation +4G and all the *CFTR* exon 12 mutations do not fit with this hypothesis. *F9* exon 5 mutation +4G reduces the number of base pairings with U1 snRNP from six to five and contrary to what the "5-6" hypothesis predicts, exon splicing is not affected (Fig. 2.2.3B, lane 8). In addition, since *CFTR* exon 12 *wt* bears seven matches to the donor site consensus sequence it should be considered as a highly conserved 5' ss and consequently GVs should cause at most moderate reductions of exon inclusion. However, all the donor site mutations caused full exon skipping (Fig. 2.2.1B).

Another proposal to distinguish neutral from splicing-disrupting GVs has been suggested upon the evaluation of six web-based tools, including NN and MaxEnt as in this thesis work, on 39 retinoblastoma 1 gene (RB1) variants (Houdayer et al. 2008). They arbitrary chose a 20% cut-off value as predictor of positive and deleterious effect of the variant. Using this threshold value, they observed a correlation between *in silico* predictions and *in vitro* splicing assays in 84% of the RB1 variants, with no false positives. Unfortunately, using such a cut-off value for the data presented in this thesis would have not improved correlation with splicing functional assays results, because predictions listed in

Table 2.1.2 present great variability.

Based on the data presented here, only 2 out of 11 bioinformatic predictions are in concordance with the splicing functional assays for mutants at donor sites. SplicePort and Hbond correctly predicted that mutation +3G on *CFTR* exon 12 created a defective splice site and that mutation +4G in *F9* exon 5 did not modify the score in comparison to the *wt* donor site. Interestingly, SplicePort browses for splicing signals and motif sequences across a 160 bp input sequence, thus considering neighbouring splicing regulatory elements around splicing signals (Dogan et al. 2007) In turn, Hbond score is built upon calculation of hydrogen bonding formation between the U1 snRNP 5'-tail and the 5' ss, taking into account a 11 bp, longer sequence, including positions +7 and +8 of the 5' ss. Thus, the only splice site prediction tools that correctly predicted consequence of donor site variants are those which are not built solely on sequence composition analysis, confirming previous observations (Thanaraj 2000).

The poor performance of splice site prediction tools scores have been reported by several publications (Hartmann et al. 2008; Houdayer et al. 2008; Wang et al. 2009; Tuffery-Giraud et al. 2009; Raponi et al. 2009; Lu et al. 2011), and some recommendations have been proposed. The IARC Unclassified Genetic Variants Working Group suggested that in presence of a new GV at least three prediction programs should be used in order to obtain an accurate prediction on its putative effect (Spurdle et al. 2008). However it does not specify which three. Other reports (Hartmann et al. 2008; Houdayer et al. 2008; Betsalel et al. 2010) made a similar recommendation, making an appeal for the combined used of several available bioinformatic tools running different algorithms to achieve maximum sensitivity. They also advocate for the development of new reliable *in silico* diagnostic tools which ultimately will become part of the clinical diagnosis. Meanwhile the majority of the studies (Tournier et al. 2008; Spurdle et al. 2008; Raponi et al. 2009; Wang et al. 2009; Théry et al.

2011) affirmed that all data generated using current score predicting tools must be validated by the use of splicing functional assays or RT-PCR from patient's RNA samples. The results presented in this thesis agree with these recommendations, underpinning the use of hybrid minigene analysis for the assessment of pathogenicity of new GVs and validation of their bioinformatic tools predictions. Moreover these splicing functional demonstrated that aberrant splicing events, chiefly exon skipping, triggered by the *CFTR* exon 12 exonic and donor site mutations and the *F9* exon 5 PPT and donor site mutations are the molecular mechanism behind their association to cystic fibrosis or haemophilia B, respectively. For most of the studied mutations this is the first report that confirms their contribution to the disease.

Section 3.3. ***Composite nature of 3' ss makes prediction of*** **F9** ***exon 5 PPT transversions impact on splicing a difficult task***

For *F9* exon 5 PPT transversions T to G, the predicted reduction in score values for seven tools (S&S, MaxEnt, MM, WMM, NN, AST,and HSF) was lower than 17% and one does not even find the *F9* exon 5 *wt* 3' ss (SplicePort) (Table 2.1.2). Nonetheless these mutations induced full exon skipping (Fig. 2.2.3, lanes 2 and 3). Thus one can hardly state that these PPT mutations actually impair splicing processing only by calculating these differences. The composite nature of the 3' ss makes prediction of PPT GVs impact a difficult task, since its recognition relies on several signals: canonical AG, PPT tract and the branch-point. For example, some "weak" PPTs can be compensated by strong, consensus-like branchpoint sequences (and viceversa)(Roscigno et al. 1993; Mullen et al. 1991) and exon enhancer elements . Furthermore, the PPT nucleotide composition, location and length varies greatly in human acceptor sites, contrary to the donor site sequences, and this affects splicing efficiency (Roscigno et al. 1993; Coolidge et al. 1997).

Pathological mutations occurring within PPT are far less abundant than those falling within donor sites or exonic sequences, showing a homogeneous distribution between positions -4 and -15 (Krawczak et al. 2007). As a matter of fact the first pathological PPT mutation described in the scientific literature is a T to G transversion at position -8 in the PPT of β-globin intron 2, which has been isolated in individual suffering from β-thalassemia intermedia (Sébillon et al. 1995). The authors demonstrated that the -8G transversion decrease splicing efficiency because the spliceosomal complex A formation is impaired. The authors also suggested that the heterogenous ribonucleoprotein (hnRNP) C bounds preferentially to the mutant PPT, which could also contribute to the pathological exon skipping. Another publication demonstrated that PPT transversions reduce U2AF^{65} binding (Millevoi et al. 2006). Therefore the -8G

and -9G transversions in *F9* exon 5 likely impair proper recognition of the acceptor site (either reducing U2AF^{65} affinity or through recognition of additional *trans*-acting splicing factors) leading to aberrant exon definition. RNA binding assays with U2AF and *F9* exon 5 -8G and -9G probes could be performed to confirm this hypothesis. This mechanism may be shared by other PPT transversions which have been described to induce pathological exon skipping in several genetic diseases such as Usher syndrome type 1b (Weston et al. 1996), Wilson disease (Loudianos et al. 2002), X-linked retinitis pigmentosa (Pomares et al. 2009), Maroteaux-Lamy syndrome (Garrido et al. 2007, 2008), growth hormone insensitivity (David et al. 2010) and IFAP (Ichtyosis follicularis, alopecia and photophobia) syndrome (Oeffner et al. 2011)-..

Section 3.4. ***Compensatory U1 snRNPs correct different types of splicing defects in* CFTR *exon 12 and* F9 *exon 5***

To understand the role of U1 snRNP in the recognition of normal and mutated CFTR exon 12 and F9 exon 5, I performed several co-transfection experiments with modified U1 snRNAs. In the first set of experiments, these U1 snRNPs were loaded by complementarity directly onto the donor site. Loading of U1 snRNAs with increased complementarity to the *wt* or mutant donor sites improved splicing of several exon skipping-causing mutations. This was particularly effective for the exonic mutations (17A and 52T) in *CFTR* exon 12 and for the PPT mutations (-8G and -9G) in *F9* exon 5 (Fig. 2.3.2A, lanes 1-7 and Fig. 2.3.5C). In the case of donor site mutants, the rescue efficiency depends on the type of mutation and the context. In fact, exon skipping caused by donor site mutations -1A, -1T, +3C, +3G and +5A within *CFTR* exon 12 and mutations -2C, -2G and -2T within *F9* exon 5 were corrected (although with variable efficacy) while *F9* exon 5 -1T, +1A and +2C failed to be rescued (Fig 3.3A, lanes 4-13 and Fig. 2.3.6A, lanes 1-12).

The rescue of several 5' ss mutations by modified U1 snRNAs indicates that the disease-causing mechanism of these mutations is the reduction of the interaction between U1 snRNA *wt* and the 5' ss, as these mutations disrupt the complementarity to the U1 snRNA *wt* 5' tail. However, the fact that compensatory U1 snRNPs also corrected exon skipping associated to exonic and PPT mutations underlines the key role of this small nuclear RNA in exon definition. This process relies on a multitude of weak interactions between the factors bound to different signals within and around the exon (Robberson, Gilbert et al. 1990). These comprises the U1 snRNP complex loaded at 5' ss, U2AF bound to the PPT and SR proteins on exonic splicing regulatory elements. Recognition of the 5' ss by the U1 snRNP is one of the first events in splicing and it is fundamental for the selection of an upstream 3' ss (Grabowski

et al. 1991; Kuo et al. 1991; Hoffman et al. 1992). The results presented here show that the negative effect of several mutations on exon definition can be reversed by reinforcing U1 snRNP-mediated 5' ss recognition through the introduction of compensatory changes into its 5'-tail (U1CFwt and U1FIXwt; Fig. 2.3.1A and 3.4A). *CFTR* exon 12 exonic mutations 17A and 52T and *F9* exon 5 PPT transversions -8G and -9G weaken exon definition likely altering adequate recognition by *trans*-acting splicing factors or spliceosomal components, respectively (Guth et al. 1999; Cartegni et al. 2002; Haque et al. 2010). Consequently splicing is probably blocked at the first step when the exon has to be recognized co-transcriptionally. In these cases, modified U1 snRNAs loaded on the 5' ss probably compensate the missing interactions and facilitate the formation of the correct network of splicing factors over the exon.

Section 3.5. ***Exon-Specific U1 snRNPs correct different types of splicing defects in* CFTR *exon 12 and* F9 *exon 5***

To gain further insight into U1 snRNP role on exon recognition and to identify a possible unique U1 snRNA acting on multiple splicing defect on the same exon, I prepared a set of modified U1 snRNAs, named exon-specific U1 snRNA (ExSpeU1). These ExSpeU1s have a modified 5'-tail that allows their binding to intronic sequences in the proximity of the 5' ss (Fig. 2.4.1 and 5.1). To promote exon definition, U1 snRNP does not necessarily have to perfectly bind at the 5' ss. Some atypical 5' ss are recognized by U1 snRNA shifted by one nucleotide (Roca et al. 2009) and Cohen and colleagues (1994) have shown that U1 snRNAs base pairing at distance from the 5' ss can rescue exon inclusion induced by artificial donor site mutations. Using this strategy, I have identified the most active ExSpeU1s, CF 11 and FIX 9, for *CFTR* exon 12 and *F9* exon 5, respectively (Fig. 2.4.2A, 4.3A and 4.4A; Fig. 2.5.3A). Interestingly, these ExSpeU1s act on different types of mutations. Exonic mutations (17A and 52T) in *CFTR* exon 12 (Fig. 2.4.2A) and the PPT mutations (-8G and -9G) in *F9* exon 5 (Fig. 2.5.3A, lanes 1-4) were very efficiently rescued by the corresponding ExSpeU1s. In *F9*, ExSpeU1FIX 9 completely rescued the defective donor sites of the three synonymous mutations in position -2 (Fig. 2.5.3A, lanes 5-10) and, in *CFTR*, ExSpeU1CF 11 rescued all 5' ss mutations with different efficiency (Fig. 2.4.3A and 4.4A).

In addition, ExSpeU1s seem to be useful also to counterbalance the impact of mutations creating cryptic donor splice sites, as *F9* exon 5 +13G. ExSpeU1FIX 9 is able to stimulate usage of the correct 5' ss, despite the fact that it base pairs onto the cryptic splice site (Fig. 2.6.1 and Fig. 2.6.2D). This finding agrees with the notion that 5' ss commitment and activation are two separate events (Cohen et al. 1994). Furthermore it highlights the potential of ExSpeU1s-based molecular therapies: one single ExSpeU1 can correct exon skipping events

caused by mutations occurring within PPT, exonic elements and donor sites or redirect 5' ss usage.

Interestingly, both modified U1 snRNP and ExSpeU1s displayed a similar activity pattern, suggesting a common mechanism of splicing rescue.

Section 3.6. ***Donor splice site mutations show a different sensitivity to modified U1 snRNA and ExSpeU1s mediated correction***

The 5' ss mutations showed a different sensitivity to the splicing correction mediated by modified U1 snRNP and ExSpeU1s. The -1T +1T and +2C mutations in *F9* exon 5 did not respond to the ExSpeU1FIX 9 or to the corresponding modified U1 snRNPs; the 5' ss mutations in *CFTR* exon 12 showed a variable response to ExSpeU1CF 11 or to the corresponding modified U1 snRNPs: in this case the rescue efficiency was optimal for the +3G, intermediate for the -1A, +3C and +5A and low for the -1T (Fig. 2.3.3A lanes 4-13; Fig. 2.4.3A and 4.4A). This results suggest that loading of U1 snRNA directly on the mutant 5' ss or on downstream intronic sequences may not be sufficient (completely or in part) for rescueing exon skipping implying that additional factors must be involved.

The fact that the +1A and +2C mutants in *F9* exon 5 were not corrected by modified U1 snRNPs was expected since the GT dinucleotide represent 99.6% of all U2-type donor splice sites (Sheth et al. 2006) and that 64% of disease-causing aberrant splicing mutations affect this nucleotide pair (Krawczak et al. 2007). In addition, it is known that mutant AT and GC dinucleotides undergo correct 5' ss cleavage but either cannot sustain 3' ss cleavage and exon joining (AT) or show a drastic reduction in splicing efficiency (GC)(Aebi et al. 1987). Consequently the finding that neither fully complementary U1 snRNPs U1FIX +1A and U1FIX +2C nor ExSpeU1F9s suppressed exon skipping could be attributed to a block in splicing processing progression after the 5' ss recognition step.

The only case of a human GT-abolishing mutation that do not completely disrupts the 5' ss and is partially suppressed by a compensatory U1 snRNP loaded at the mutant donor site has been recently reported for the Fanconi anemia C (*FANCC*) gene (Hartmann et al. 2010). In this example, a donor site

mutation on *FANCC* exon 2 is found in patients afflicted by moderate Fanconi anemia and genomic sequencing revealed that canonical G at position +1 of the *FANCC* exon 2 donor splice is mutated to a T, turning GT dinucleotide to TT. This mutant TT dinucleotide is nevertheless used in 30% of the spliced transcripts and co-transfection of patient's fibroblasts with a fully compensatory U1 snRNP (improved base pairing from position -3 to position +8) rises mutant 5' ss usage to ~57%. The authors proposed that an overall high complementarity to the compensatory U1 snRNP might be a prerequisite for splicing. Additionally, positions -1 to +8 (with the obvious exception of +1T) of *FANCC* exon 2 are occupied by consensus nucleotides, allowing for extensive base pairing with normal U1 snRNP and interaction with U6 snRNP (at positions +4 to +6). This may explain why there is 30% of in-frame TT splicing in patient's fibroblasts and addition of fully complementary U1 snRNP almost doubled the percentage of TT-spliced transcripts. In the case of *F9* exon 5 +1A and +2C mutants, I did not detect in-frame AT or GC spliced transcripts.

Compensatory U1 snRNPs and ExSpeU1CFs were able to correct exon skipping caused by all the *CFTR* exon 12 5' ss mutants (Fig 3.3A , 4.3A and 4.4A), indicating that these mutations impair mainly the U1 snRNP-mediated 5' ss recognition. Nevertheless, the rescue efficiency of compensatory U1 snRNAs and ExSpeU1CFs clearly exhibited mutation-dependent variability. For instance, although they occur at the same position, correction of exon skipping caused by *CFTR* exon 12 mutant -1A was more pronounced than that of mutant -1T in presence of either corresponding compensatory U1 snRNPs (Fig. 2.3.3A,B, lanes 5 and 7) or ExSpeCF 11 (Fig. 2.4.3, lanes 5 and 11), a trend also noticed for mutant +3G and +3C (Fig. 2.3.3A,B, lanes 9 and 11).

Pairwise association studies showed the -1A+4T is more represented than -1T+4T, what could explain their differential sensitivity to modified U1 snRNPs treatment (Roca et al 2008). In fact, once the adequate base pairing of U1 snRNP at position -1 or right downstream the mutant 5' ss has been restored,

the presence of an adenine was favoured against a thymidine, likely because of the interaction of mutant -1A with the conserved U-rich loop of U5 snRNP (Wyatt et al. 1992). Nevertheless I cannot rule out the possibility that some unidentified *trans*-acting splicing factor can modulate *CFTR* exon 12 mutant -1T by direct binding to the mutant sequence or by stabilizing the 5' ss: U1 snRNP 5'-tail duplex.

CFTR exon 12 donor site mutations at position +3 (+3C and +3G) also displayed differential responsiveness to compensatory U1 snRNPs treatment, as mutant +3G exhibited higher inclusion levels than mutant +3C (Fig. 2.3.3A,B, lanes 9 and 11). The level of exon inclusion attained upon co-transfection with their compensatory U1 snRNPs can be related to the proportional bond strength formed at position +3: a G can form a non Watson-Crick bond with U1 snRNP 5'-tail while a C introduces a mismatch (Fig. 2.3.1A). This is further evidenced by the finding that *CFTR* exon 12 mutant +3G was the only donor site mutant that responded to U1CFwt treatment, since it is known that G-U bonds are stabilized by adjacent Watson-Crick bonds (Carmel et al. 2004). In addition, ExSpeU1CF 1 was active only in presence of a +3G mutation and not in mutant +3C. Indeed, previous reports have demonstrated that internal mismatches impair the activity of modified U1 snRNPs with shifted base pairing(Cohen et al. 1994).

Mutation +5A on *CFTR* exon 12 5' ss triggered exon skipping, which was partially corrected either in presence of U1CF +5A or ExSpeU1CF 11 (Fig. 2.3.3A, B, lane 13; Fig.4.4A,B, lane 17). Consensus G at position +5 in human 5' ss is highly conserved (77.6%) (Sahashi et al. 2007) and its base pairing with the 5'-tail of U1 snRNP is believed to be fundamental for initial donor splice site recognition (Ast 2004b). Furthermore position +5 is known to be bound by U6 snRNP, being important for catalytic steps of splicing reaction (Konforti et al. 1994). Therefore the occurrence of an adenine (present in only 8.9% of human 5' ss)(Sahashi et al. 2007) severely impairs U1 snRNP base pairing and

furthermore it likely reduces affinity for U6 snRNP *A*CAGAG box. In fact, this double potential negative effect of mutation +5A may explain why it was the least sensitive to both compensatory U1 snRNP and ExSpeU1CFs of all *CFTR* exon 12 donor site mutants studied in this work.

Section 3.7. ***Different rescue efficiency of the -1T donor site mutation in* CFTR *exon 12 versus* F9 *exon 5 context*.**

Interestingly, the use of compensatory U1 snRNPs or ExSpeU1s recovered exon inclusion when -1T occurs within *CFTR* exon 12 5' ss but was unsuccessful when present in *F9* exon 5 5' ss context (compare Fig. 2.3.3A lane 7 with Fig. 2.3.6A, lane 8 and Fig. 2.4.3A, lane 11 with Fig. 2.5.3A, lane 12). Thus responsiveness to modified U1 snRNPs treatment seems to depend on the sequence context in which mutation -1T occurs. *F9* exon 5 5' ss may represent an unfavourable environment for base substitutions at position -1. Position -1 of human donor sites is very conserved, being regularly occupied by a G (80.3%) while the presence of a T is underrepresented, comprising only 10% of 5' ss (Sahashi et al. 2007).

This high degree of conservation correlates well with the finding that position -1 is a frequent mutation hotspot in human donor sites (after the invariant GT) (Krawczak et al. 2007). One way to ponder the contribution of neighbouring nucleotides is the study of nucleotide pairwise associations on 5' ss. Some nucleotides combinations are enriched and others depleted in human donor sites apparently due to mechanistic constraints imposed by the spliceosome, since every 5' ss nucleotide interacts with several factors during the splicing reaction. For example, some reports have indicated that within natural human donor sites there is a clear association between positions -1 and +5 and +6 (Carmel et al. 2004; Roca et al. 2008). Normally a consensus G at position -1 allows for any base substitution at positions +5 and +6 while consensus G and T at positions +5 and +6, respectively, can accompany any nucleotide at position -1. *F9* exon 5 *wt* 5' ss contains a -1G+5A+6T combination, what suggests that the consensus G at position -1 may compensate for the two nonconsensus nucleotides at positions +5 and +6. The occurrence of a nonconsensus T at position -1 creates a -1T+5A+6T combination, which is

strongly depleted in human donor sites (Roca et al. 2008) . This unfavourable combination might not only reduce binding to U1 snRNP 5' but affect subsequent interactions with other snRNPs.

In fact, position -1 is known to interact with U5 snRNP conserved U-rich loop (Wyatt et al. 1992; Reyes et al. 1996) while positions +5 and +6 are important for U1 snRNP binding and known to be bound by U6 sRNP (Konforti et al. 1994). As demonstrated in this work, restoration of base pairing using U1FIX -1T or loading of ExSpeU1CF 11 on its vicinity did not improve exon inclusion (Fig. 2.3.6A, lane 8 and Fig. 2.5.3A, lane 12), indicating that in this donor site context this substitution likely impairs recognition by other spliceosomal components. Therefore it seems that in *F9* exon 5 donor site context position -1 plays a key role in 5' ss selection, being needed not only for proper U1 snRNP binding but also to guarantee progression of the splicing reaction. Complementation studies with additional spliceosomal factors such as U6 snRNP and U5 snRNP could bring some further insights into the mechanism of action behind this mutation.

On the other hand, the *CFTR* exon 12 *wt* 5' ss contains a -1G+5G+6T combination, which are three consensus nucleotides. The occurrence of a nonconsensus T at position -1 in this 5' ss affected chiefly U1 snRNP binding, as the use of U1CF -1T or ExSpeU1CF 11 promoted exon inclusion (Fig. 2.3.3A lane 7 and Fig. 2.4.3A, lane 11). It can be speculated that consensus +5G and +6T in this 5' ss partially buffered the impact of mutation -1T, unlike in *F9* exon 5 donor site situation. Indeed, the -1T+5G+6T combination as in *CFTR* exon 12 -1T mutant 5' ss is positively represented in human donor sites (Roca et al. 2008) which suggests that this compensation effect is mechanistically feasible.

This also agrees with a previous report on the exon 20 donor site (CAA|guaagu) of the IkappaB kinase complex-associated protein (*IKBKAP*) gene, in which a T to C substitution at position +6 is associated to familiar dysautonomia (FD)(Carmel et al. 2004). In this work the authors

demonstrated the existence of a clear compensation effect between position -1 and +6 (the donor site in this case bears a consensus G at position +5) as either the use of compensatory U1 snRNP base pairing to +1A or the introduction of a consensus G at position -1 counterbalanced the nonconsensus +6C, improving exon inclusion levels. Summarizing, the results presented here reveal that the success of modified U1 snRNPs treatment for splicing correction of -1T mutants depended on the composition of the +5 and +6 pair: consensus +5G and +6T allowed progression of splicing reaction once the base pairing with U1 snRNP was restored.

Section 3.8. ***Biosynthesis of the co-transfected modified U1 snRNAs: comparison between modified U1 snRNAs and ExSpeU1s suggest that the majority of transfected U1 snRNAs are efficiently processed***

Modification of U1 snRNAs 5'-tail, either with point mutations (Zhuang et al. 1986; Susani et al. 2004; Pinotti et al. 2008, 2009) or introduction of several tens of nucleotides (De Angelis et al. 2002; Denti et al. 2006) is effective is several cases. However some modifications may hamper adequate U1 snRNA maturation, introducing a confounding factor in the interpretation of the results (Yuo and Weiner 1989). In my analysis, I did not evaluate the correct biosynthesis of the co-transfected modified U1 snRNPs. Due to the modifications introduced in their 5'-tail some engineered U1 snRNPs might not be properly processed or exported into the nucleus. However, the finding that both compensatory and ExSpeU1 snRNPs were active on the same mutants and with similar efficiency argues against the existence of major U1 snRNA processing alteration. For example, the ExSpeU1FIX 9 is active on several mutants but not on *F9* exon 5 mutants -1T, +1A and +2C (Fig. 2.5.3A, lanes 11-16). These mutants contain the corresponding binding site (Fig. 2.2.5.1) thus their nonresponsiveness cannot be ascribed to a inefficient ExSpeU1FIX 9 biosynthesis. Equally, modified U1 snRNP and ExSpeU1s corrected several CFTR exon 12 mutant 5' ss splicing defects with similar mutation-dependent efficiency. Hence, non responsiveness or reduced effect is probably due to the mutation itself and not to defective biosynthesis of transfected U1 snRNAs.

Section 3.9. ***ExSpeU1s can correct* F9 *pre-mRNA processing and factor IX synthesis and secretion in a splicing-competent cDNA model***

One interesting result of my experiments is that the recruitment of ExSpeU1 by complementarity to intronic sequences downstream of the exon corrects exon skipping caused by multiple splicing mutations occurring either at 5' or 3' splice sites or in exonic regulatory elements. Both CFTR and factor IX are expressed in differentiated cells (epithelial cells of the airways and hepatocytes) and no cellular models with the mutations I have studied are available. Thus, to understand whether the ExSpeU1s can rescue pre-mRNA splicing and protein production I focused on *F9* exon 5 mutations using a splicing competent cDNA expression vector. This vector carries the *wt* or mutant exon 5 and flanking intronic regions (Fig. 2.12.1A). As shown in Fig. 2.12.1B splicing of the *wt* minigene was less efficient in comparison to the *in vivo* situation and this is probably due to its context. I can speculate, as previously suggested in other minigene systems (Romano et al. 2002), that the absence of upstream intronic sequences can significantly affect the downstream splicing efficiency. Even if the minigene is in a less favourable context (i.e., the exon 5 is not well recognised) U1FIXwt or ExSpeU1FIX 9 induced a nearly complete exon inclusion: the three synonymous mutations at position -2 and the two PPT transversions (-8G and -9G)(Fig. 2.12.1B, C). Restoration of correct splicing corresponded to synthesis of the full-length factor IX protein (Fig. 2.12.1D), with levels and activity similar to the *wt* situation (Fig. 2.12.2 and Fig. 2.12.3). This result if translated *in vivo* would reach a level of correction consistent with a significant therapeutic benefit.

Section 3.10. ***An intronic splicing silencer element is present downstream* CFTR *exon 12 and* F9 *exon 5 5′ ss and ExSpeU1s activity may be explained in part by interference with ISS recognition***

In both *CFTR* exon 12 and *F9* exon 5, the most active ExSpeU1s target a region at 20 bp downstream the 5' ss (spanning from position +11 on *CFTR* exon 12 and +9 on *F9* exon 5). ExSpeU1s can help to define the exon by recruiting enhancing splicing factors and/or interfering with intronic splicing regulatory elements. The fact that modified U1 snRNP complementary to the 5' ss and ExSpeU1 showed the same rescue profiles in term of efficacy and efficiency suggest a recruitment model. However, in *CFTR* exon 12 and in *F9* exon 5, the consecutive deletion mutants identified two weak intronic splicing silencers (ISS) that overlap to the corresponding binding sites of ExSpeU1s. In addition, ExSpeU1s activity is strictly dependent on base pairing to this target sequence as evaluated in co-transfection experiments with the ISS deletion mutants (Fig. 2.8.1A, B, C).

To better understand the relationship between the ISS and the 5' ss mutations, I inserted, in *CFTR* exon 12, the ISS deletion mutants in the context of the exonic (17A) and 5' ss mutations (-1A and +3G). Deletion of ISS sequence promoted increased exon inclusion in the 17A minigene (Fig. 2.7.1D, E) but had no effect on splicing pattern of donor site mutant minigenes (Fig. 2.7.2A, B). Thus, the *CFTR* exon 12 ISS seems to have a weak inhibitory activity evident only when the exon definition is compromised by exonic mutations or in the *wt* situation. To further get insight on the role of the ISS I targeted *CFTR* exon 12 and *F9* exon 5 ISSs with antisense oligonucleotides (ASO). Only a modest increase of exon inclusion was observed in the presence of ASO^{CFTR}_{25}, targeted to *CFTR* exon 12 ISS. As I tested only a limited numbers of ASO, (two for each system) I cannot completely exclude that their lack of efficacy was due to sequence constraints. A deeper analysis including a larger number of ASOs is

required to understand whether the interference of ISS recognition is sufficient of mimic ExSpeU1s activity.

Section 3.11. ***Role of hnRNPA1 and RNA secondary structure in ISS in* CFTR *exon 12 and* F9 *exon 5***

Both *CFTR* exon 12 and *F9* exon 5 ISSs exhibit positional and partial sequence similarity to SMN2 exon 7 ISS (Fig. 2.9.1A), a well-described regulatory element situated right downstream of the 5' ss which has been demonstrated to be bound by hnRNP A1 (Hua et al. 2008). For this reason, I performed a Western blot analysis upon pulldown using deletion mutant RNA oligonucleotides to assess whether hnRNP A1 also binds *CFTR* exon 12 and *F9* exon 5 ISS. The experiment results showed that *F9* exon 5 ISS is bound by hnRNP A1/A2 complexes while *CFTR* exon 12 ISS is not (Fig 9.1B). Deletion of the ISS in *F9* exon 5 disrupts the interaction with the splicing factor (Fig 9.1B). This result suggest that hnRNP A1/A2 interaction with the ISS might be involved in *F9* exon 5 splicing regulation and not in CFTR exon 12.

To further explore the mechanism behind *CFTR* exon 12 ISS activity, I performed a mutagenesis analysis on *CFTR* exon 12 ISS, introducing point mutations across the entire element sequence (Fig. 2.11.1A). In agreement with the impact on splicing of *CFTR* exon 12 ISS deletion mutants (Fig. 2.7.1A-E), some substitutions promoted exon inclusion (from positions +16 to +19)(Fig. 2.10.1B, C, lanes 5-10). However, others displayed either negative (substitutions +13C and +15C)(Fig.10.1B, C, lanes 3 and 4) or neutral effect on *CFTR* exon 12 inclusion (substitutions from position +20 to +23)(Fig. 2.10.1B, C, lanes 11-16). Analysis of the predicted RNA structure region indicate that the ISS element of *CFTR* exon 12 is part of a stem-loop structure (Fig. 2.10.2) and that some of the mutants affect this structure. (Fig.10.3 and Fig. 10.4). Predicted stability of the stem-loop structure in presence of ISS mutants exhibited a direct correlation with their impact on splicing pattern: exon skipping was associated to a more stable stem-loop structure (mutation +13C) while mutations that opened the stem structure induced higher exon inclusion rates (mutations at positions

+16 to +19). Furthermore the deletion of part of CFTR exon 12 ISS stimulated exon inclusion, in agreement with the predicted abolishment of the stem loop. Altogether these results suggest a model where the stem-loop structure predicted for CFTR exon 12 ISS influence exon splicing, probably by regulating accessibility of U1 snRNP to the 5' ss.

Interestingly, a similar stem-loop structure was previously noticed in the *CFTR* exon 12 5' ss region (Meyer et al. 2005) but its role on splicing not evaluated. The modulation of splicing outcome by RNA secondary structures involving stem-loop structures at exon/intron boundaries has been described in different gene models such as *HNRNPA1* exon 7B (Blanchette et al. 1997; Chabot et al. 1997), *SMN2* exon 7 (Singh et al. 2006), *COL2A1* exon 2 (McAlinden et al. 2005), HIV-1 RNA (Jacquenet et al. 2001) and *tau* exon 10 (Grover et al. 1999). The stem-loop of CFTR exon 12 is similar to the stem-loop described in *tau* exon 10 (Hutton et al. 1998; Spillantini et al. 1998). Alternative splicing of *tau* exon 10 produces two Tau isoforms: Tau 3R (lacking exon 10) and Tau 4R (with exon 10), displaying a ratio ~1 in adult human brain. Several mutations occurring in the vicinity of *tau* exon 10 5' ss are associated to an increase of the 3R/4R ratio and consequently with the development of tauopathies (Hutton et al. 1998; Spillantini et al. 1998; Yasuda et al. 2000). These mutations destabilize the stem-loop structure (Yasuda et al. 2000; Donahue et al. 2006) making 5' ss more accessible for U1 snRNP recognition with a consequent increase in exon 10 inclusion and 3R/4R isoform ratio (Varani et al. 1999; Jiang et al. 2000; Kalbfuss et al. 2001; Kar et al. 2011). Conversely stabilization of the stem-loop either by point mutagenesis (Donahue et al. 2006) or interaction with small molecules (Zheng et al. 2009) reduces exon 10 inclusion. Remarkably, predicted free energies for *wt* tau exon 10 and *CFTR* exon 12 stem-loops are similar: -8.2 and -6.4 kcal/mol, respectively (Donahue et al. 2006)(Fig. 2.10.2) as the estimated ΔG in mutants that induce exon inclusion and stabilize the stem loops: *CFTR* exon 12 +13C stem-loop has a predicted free energy of -11.4

kcal/mol and *tau* exon 10 +17T stem-loop -13.5 kcal/mol (Donahue et al. 2006).

Section 3.12. ***Therapeutic approaches based on ExSpeU1s and their clinical significance***

So far only a limited number of studies have explored the role and potential therapeutic effect of modified U1 snRNAs on splicing correction of donor site mutations (Hitomi et al. 1998; Baralle et al. 2003; Susani et al. 2004; Pinotti et al. 2008, 2009; Tanner et al. 2009; Hartmann et al. 2010; Schmid et al. 2011). These approaches have been proposed for the correction of splicing defects in human malignant infantile osteoporosis (arOP)(Susani et al. 2004), coagulation factor VII (FVII) deficiency , retinitis pigmentosa (Tanner et al. 2009; Glaus et al. 2011), Fanconi anemia (Hartmann et al. 2010) and Bardet-Biedl syndrome (Schmid et al. 2011). In all cases, the modified 5'-tails of U1 snRNA have few nucleotide changes in comparison to the *wt* sequences and base pair exactly to the mutant donor sites. Thus, their therapeutic effect is restricted to each mutation and the potential complementary of this modified U1 snRNA to other normal 5' ss might affect splicing of other exons, in particular those alternatively spliced. The ExSpeU1s I have developed here, binding to not conserved intronic sequences, do not interact directly with normal 5' ss and have the advantage of correcting different types of mutations associated to exon skipping and affecting different splicing regulatory elements. In particular, ExSpeU1s are active on several 5' ss mutations in *CFTR* exon 12 and *F9* exon 5, on two transversions at the PPT in *F9* exon 5 and on two exonic substitutions in *CFTR* exon 12. To our knowledge, this is the first time that an U1 snRNA variant is shown to correct PPT and exonic mutants. Another advantage of a putative ExSpeU1-based corrective splicing therapy is the short length of the ExSpeU1s cassette, of approximately 500 bp which can be easily inserted into viral vectors like adenoassociated viruses (AAV). This make them suitable for the development of gene therapy of large genes like *CFTR*, whose transcript length might represent a limiting step for carrier

vector construction. Additionally, ExSpeU1s acting on pre-mRNA, will also maintain the regulation of the gene expression in the correct chromosomal context. In the case of dominant-negative mutations where the replacement therapy is not feasible, the splicing correction mediated by ExSpeU1 will act directly reducing the amount of the mutated toxic protein. Moreover, the binding of the ExSpeU1s to nonconserved intronic sequences will significantly reduce the possibility of off-target events at complementary not conserved donor sites. Although binding of ExSpeU1s to unrelated sequences not directly involved in splicing might have an effect on other pre-mRNA processing steps (Beckley et al. 2001; Kaida et al. 2010), our result indicate that it will be possible to design ExSpeU1s with improved specificity by inserting subtle changes in sequence target and/or length. For example, ExSpeU1FIX 9, 10 and 13 (Fig. 2.5.1) bind at nearby or overlapping intronic sequences and are functionally active. Future studies in cellular and animal models will be required to address this point.

More in general, ExSpeU1s-mediated induction of exon inclusion could be used to regulate the synthesis of specific alternatively spliced isoforms for therapeutic purposes. For example, therapeutic induction of specific alternative splicing isoforms affect tumor progression (Ghigna et al. 2008)), angiogenesis (Merdzhanova et al. 2010) or aging (Fong et al 2009) and modified U1 snRNAs have been recently used to promote splicing to inhibit HIV replication (Mandal et al. 2010).

4. Future Directions

Future studies should provide new insights into *CFTR* exon 12 and *F9* exon 5 splicing regulation and ExSpeU1s mechanism of action.

The regulation of *F9* exon 5 recognition and splicing is still unclear. In this work I identified an ISS, which is to my knowledge the first *cis*-acting regulatory element described for this exon. Further studies aimed at the discovery of exonic regulatory sequences should analyse the splicing impact of the several disease-associated mutations reported for F9 exon 5. Besides synonymous mutants at position -2, two additional silent mutations, V107V and R116R have been reported for *F9* exon 5 (Haemophilia B Mutation Database). The molecular mechanism of these silent mutations is probably related to splicing, as it has been described for several disease-associated silent substitutions like *CFTR* exon 12 52T. Experimentation with splicing functional assays could uncover new exonic regulatory elements implicated in exon definition, as demonstrated in *CFTR* exon 12 (Pagani et al. 2003a).

Furthermore the possible effect of missense mutations on *F9* exon 5 splicing should not be dismissed as it has been demonstrated that the real mechanism behind some missense mutations is not the predicted aminoacid change but actually the impairment of adequate splicing processing. For example, the molecular consequence of D648V, and T665S mutations of *CFTR* gene was previously predicted to cause amino acid changes that could affect the chloride channel activity of the CFTR protein. Nevertheless, subsequent studies showed that when D648V and T665S mutations were introduced into CFTR cDNA they showed no effect on the chloride channel activity of the corresponding expressed mutant CFTR proteins (Vankeerberghen et al. 1998).Later on, mutations D648V and T665S were found to occur within an exonic splicing enhancer element of *CFTR* exon 13 (Aznarez et al. 2003).

The mechanism of action of *CFTR* exon 12 and *F9* exon 5 ISS should also be the

focus of future studies. My data indicates that these ISSs act play a role in exon definition by different ways: the former is embedded in a stem-loop structure whereas the latter is bound by hnRNP A1.

The results presented here suggest that in CFTR exon 12 the stem-loop structure at the ISS may regulate the accessibility to the 5' ss, likely competing with U1 snRNA 5'-tail for the donor site sequence. This hypothesis of the existence of a stem-loop structure at the exon/intron boundary of *CFTR* exon 12 5' ss should be evaluated through RNA structure probing or UV melting experiments as it has been done for *SMN2* exon 7 terminal stem-loop 2 (Singh et al. 2006) and *tau* exon 10 (Yasuda et al. 2000; Donahue et al. 2006), respectively.

I showed that *F9* exon 5 ISS sequence is bound by hnRNP A1 through Western blotting analysis upon RNA pulldown. Additional experiments such as RNAi and point mutagenesis analysis will be required to clearly establish the involvement of hnRNP A1 in *F9* exon 5 ISS recognition and activity, and to uncover possible role of other splicing factors.

Part of the ExSpeU1s effect is due to "antisense" mechanism, blocking the binding of trans-acting splicing factors (like hnRNP A1 for *F9* exon 5) to the ISS or interfering with the formation of a RNA secondary structure (*CFTR* exon 12). The contribution of the antisense mechanism could be explored by designing specific ASOs targeted to the ISS and/or using U7 snRNA-derived systems (Gorman et al. 1998; Brun et al. 2003).

The finding that both compensatory U1 and ExSpeU1 snRNPs restore exon definition and splicing with similar efficiency suggest that they may share a common mechanism of action. Apart from their effect on the ISS they may act recruiting additional splicing factors that help in exon definition. Specific proteins normally associated to U1 snRNA might be involved in ExSpeU1 effect. U1-70k emerges as the main candidate, since it is known to interact with SR proteins like SF2/ASF and SC35 in order to properly define the exon (Wu et al.

1993; Cho et al. 2011). The use of ExSpeU1 mutants for U1-70K binding (Gunderson et al. 1998; Ashe et al. 2000; Abad et al. 2008) in co-transfection experiments with *CFTR* exon 12 and *F9* exon 5 mutants should help to clarify this point.

5. Conclusions

In this thesis work, I have evaluated the pathogenic role of several natural mutations occurring at different splicing signals on *CFTR* exon 12 and *F9* exon 5 and the role of U1 snRNP in defective splicing processing.

Using a minigene system that reproduces the endogenous pattern of splicing, I showed that most of the substitutions affecting PPT, exonic regulatory elements and donor splice sites in *CFTR* exon 12 and *F9* exon 5 induce exon skipping, with the exception of two *F9* exon 5 mutations that either activated a cryptic site (+13G) or resulted in no pathological changes in the splicing pattern (+4G). With the exception of CFTR exon 12 17A, 52T and +3G mutants, for the disease-associated GVs studied here this is the first report to demonstrate their impact on splicing.

Prediction of splice sites score values using several bioinformatic tools showed great variability. Although in general terms a reduction in splice site score values was correlated with the hybrid minigene analysis results, the effect of some mutations (*CFTR* exon 12 +3G and *F9* exon 5 +4G) was underestimated for the bulk of the *in silico* tools. Interestingly, the three prediction tools which showed the best performances on donor splice site predictions do not rely on relative nucleotide frequencies but rather on thermodynamical constraints (Hbond and ΔG) or take into account the sequence context (SplicePort). These results agrees with previous publications that indicated the poor performance of *in silico* tools that consider only local sequence information and advocate for the use of functional splicing assays when it comes to certify pathogenicity of a new GVs.

Given the central role played by U1 snRNP as initiator of exon definition process, I have evaluated modified U1-snRNAs with a 5'-tail specifically engineered for recognizing the defective donor splice sites. These compensatory U1 snRNPs were able to suppressed the effect of PPT and 5' ss

mutants in *F9* exon 5 and exonic and 5' ss mutants in *CFTR* exon 12. This confirmed that these mutations impaired either proper exon definition (*F9* exon 5 PPT mutations and *CFTR* exon 12 exonic mutants) or reduced affinity of U1 snRNP 5'-tail: 5' ss base pairing (donor site mutations).

Suppression of mutation -1T showed a context-dependent efficiency to compensatory U1 snRNPs since this substitution was efficiently suppressed when located on *CFTR* exon 12 5' ss but not on *F9* exon 5 5' ss. Comparison of the nucleotide dependencies within the *F9* exon 5 5' ss indicates that position -1 in that context plays a important role for splicing reaction progression to the catalytic steps. Suppression of this mutation would require further complementation with additional splicing factors.

Next I explored the potential use of exon-specific U1 snRNPs as an aberrant splicing corrective approach and demonstrated that it can be useful to correct *F9* pre-mRNA processing in presence of either PPT or donor site mutations with consequent recovery of factor IX protein biosynthesis, secretion and activity. The factor IX synthesized upon transfection of ExSpeFIX 9 was correctly secreted and achieved *wt* levels of activity, which would be compatible with therapeutic improvement of the disease severity. The target sequence of the most active ExSpeU1s coincided with weak ISS elements, located just downstream of both *CFTR* exon 12 and *F9* exon 5 5' ss. The data presented here suggests that these ISS elements exert their action by different mechanisms: while hnRNP A1 seems to be involved in *F9* exon 5 ISS recognition, *CFTR* exon 12 ISS forms part of a stem-loop structure which may govern access to its 5' ss. Thus ExSpeU1s may function through recruitment of splicing factors and partial interference with ISS recognition. In addition, ExSpeFIX 9 was also able to correct 5' ss recognition in presence of a cryptic splice site created by mutation +13G. This result further expand the spectra of mutations that can be suppressed by ExSpeU1s, underlying their potential for the development of innovative therapeutic approaches.

6. Materials and Methods

6.1. Chemical reagents

General chemicals were purchased from Sigma Chemical Co., Merck, Gibco BRL, Boehringer Mannheim, Carlo Erba and Serva.

*6.2.*Standard solutions

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

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

b) PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO4, 1.8 mM KH_2PO4, pH 7.4

c) 10X TBE: 108 g/l Tris, 55 g/l Boric acid, 9.5 g/l EDTA

d) 6X DNA sample buffer: 0.25 % w/v bromophenol blue, 0.25 % w/v xylene cyanol FF, 30 % v/v glycerol in H_2O.

*6.3.*Enzymes

Restriction enzymes were from New England Biolabs, Inc. DNA modifying enzymes such as *Taq* Polymerase, DNase I RNase free, and T4 DNA ligase were obtained from Roche Diagnostic. Klenow fragment of *E. coli* DNA polymerase I and T4 polynucleotide Kinase were from New England Biolabs, Inc. RNase A was purchased from Sigma Chemicals Ltd. A 10 mg/mL solution of RNase A was prepared in sterile water and boiled for 10 minutes to destroy trace amounts of DNase activity. All enzymes were used following manufacturer's instructions.

*6.4.*Synthetic oligonucleotides

Synthetic DNA oligonucleotides for PCR and cloning were purchased from

Sigma-Genosys Ltd. RNA oligonucleotides for RNA pulldown analysis were synthesized by Integrated DNA Technologies (Leuven, Belgium). 2'-O-methoxyethyl-phosphorothiate antisense oligonucleotides (ASO) synthesized by Dr. Daniela Perrone, University of Ferrara (Ferrara, Italy).

6.5. **Bacterial culture**

The *E. coli* K12 strain DH5 was transformed with the plasmids described in this study and used for their amplification. Plasmids were maintained in the short term as single colonies on agar plates at 4 °C. For long term storage they were kept on glycerol stocks, made by adding sterile glycerol to a final 30% v/v concentration to liquid bacterial cultures. Glycerol stocks were stored at -80°C. When necessary, an overnight culture of bacteria taken from the glycerol stocks was grown in Luria-Bertani medium [LB medium: per litre: 10 g Difco Bactotryptone, 5 g Oxoid yeast extract, 10 g NaCl, (pH 7.5)]. Bacterial growth media were sterilized before use by autoclaving. When appropriate, ampicillin was added to the media at a final concentration of 200 μg /mL.

*6.6.***Cell culture**

The cell line used for transfection and co-transfection experiments of hybrid minigenes were HeLa cells, human cervical carcinoma. For transfection of splicing-competent factor IX expression vectors and rescue of factor IX protein secretion and activity BHK (Baby hamster kidney) cells were used.

*6.7.***DNA preparation**

6.7.1.Small scale preparation of plasmid DNA from bacterial cultures

Rapid purification of small amounts of recombinant plasmid DNA was performed with the method previously described by Sambrook (Sambrook,

Fritsch et al. 1989). Briefly, alkaline lysis of recombinant bacteria was performed by resuspending the bacterial pellet in 200 µL of ddH_2O; 200 µL of solution II (0.2 M NaOH, 1 % w/v SDS) were then added and the contents mixed by inversion. 300µl of solution III (3 M potassium acetate pH 5.2) were then added and the contents mixed by inversion. The bacterial lysate was then centrifuged in an Eppendorf microcentrifuge at maximum speed and the supernatant transferred to a new tube. An equal volume of 1:1 v/v phenol:chloroform solution was added to the supernatant. The tube was then vortexed and centrifuged as above. The aqueous phase containing the DNA was transferred to a new tube. An equal volume of chloroform was added to the supernatant. The tube was then vortexed and centrifuged as above. The aqueous phase containing the DNA was then recovered and the DNA pelleted by ethanol precipitation. The final pellet was resuspended in 50 µL of ddH_2O +3 µL of RNAse A (10 mg/mL) for at least 1h at 37°C and 5 µL of such preparation were routinely taken for analysis by restriction enzyme digests.

6.7.2.Large scale preparations of plasmid DNA from bacterial cultures

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

*6.8.***RNA preparation from cultured cells**

Cultured cells were washed with PBS and then RNA Trizol (Invitrogen SRL, Milano, Italia) was added. Then, chloroform extraction was performed. Supernatant was precipitated with isopropanol. The pellet was resuspended in 50 µL of ddH_2O and digested with 1U of DNase RNase free. The mix was incubated at RT for 30 minutes, and then the RNA was purified by acid phenol

extraction. The final pellet was resuspended in 33 µL of ddH_2O and frozen at -80 °C. The RNA quality was checked by electrophoresis on 0,8% agarose gels.

6.9. **Estimation of nucleic acid concentration**

An optical density of 1.0 at 260 nm is usually taken to be equivalent to a concentration of 50 µg/mL for double stranded DNA, 40 µg/mL for single stranded DNA and RNA, and approximately 20 µg/mL for single-stranded oligonucleotides samples. The ratio of values for optical densities measured at 260 nm and 280 nm is considered as 1.8 for pure sample of DNA and 2 for RNA and these are reduced by protein contaminants (Sambrook, Fritsch et al. 1989). Therefore, these values were used to determinate not only the concentration but also the purity of the samples.

6.10. **Enzymatic modification of DNA**

6.10.1. Restriction enzymes

Restriction endonucleases were used in the construction and analysis of recombinant plasmids. Each restriction enzyme functions optimally in a buffer of specific ionic strength. All buffers were supplied by the same company that supplied the enzymes and were used according to the manufacturer's instructions.

For analytical digests 100-500 ng of DNA were digested in a volume of 20 µL containing 5 U of the appropriate restriction enzyme. The reaction was incubated for 2-3 hours at 37 °C. Preparative digestion was made of 5-10 µg DNA using the above conditions and 5 U of restriction enzyme for µg of DNA in 200 µL reaction volume.

6.10.2. Large fragment of E. coli Polymerase I and T4 Polynucleotide Kinase

These enzymes were used to treat PCR products for blunt-end ligation during construction of recombinant plasmids. The large fragment of DNA Polymerase I (Klenow) is a proteolytic product of *E. coli* DNA Polymerase I. It retains polymerization and 3'->5' exonuclease activity, but has lost 5'->3' exonuclease activity. This was useful for digesting a specific residues added by *Taq* DNA polymerase at the 3' terminus to create compatible ends for ligation. T4 Polynucleotide Kinase catalyses the transfer of phosphate from ATP to the 5' hydroxyl terminus of DNA. It was useful for the addition of 5'-phosphate to PCR products to allow subsequent ligation. Klenow fragment (2.5 U) was added to 23 µL of PCR product in 5 mM $MgCl_2$ buffer. The mixture was incubated at room temperature for 10 minutes. EDTA to a final concentration of 0.2 mM, ATP to a final concentration of 1 mM, 10 U of T4 Polynucleotide Kinase and the proper quantity of Kinase buffer were added to the above mixture and incubated at 37 °C for 30 min. The enzymes were inactivated by incubation at 80 °C for 20 min.

6.10.3. T4 DNA ligase

T4 DNA ligase catalyses the formation of a phosphodiester bond between adjacent 3' hydroxyl and 5' phosphoryl termini in DNA, requiring ATP as a cofactor in this reaction. This enzyme was used to join double stranded DNA fragments with compatible sticky or blunt ends, during generation of recombinant plasmid DNAs. 20 ng of linearized vector were ligated with a 5-10 fold molar excess of insert in a total volume of 20 µL containing 1X ligase buffer and 1U of T4 DNA ligase. Reaction was carried out at room temperature for 2-4 hours at RT for sticky end ligations and ON at 16 °C for blunt end ligations.

In some reactions synthetic oligonucleotide were included in the reaction. In

these cases, the amounts added to each reaction to obtain inclusion of oligonucleotides in the resulting plasmid were about 100 fold molar excess over the DNA vector.

6.11. **Agarose gel electrophoresis of DNA**

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

6.12. **Elution and purification of DNA fragments from agarose gels**

This protocol was used to purify small amounts (less than 1 µg) of DNA for sub-cloning. The DNA samples were electrophoresed onto an agarose gel as described previously. The DNA was visualized with UV light and the required DNA fragment band was excised from the gel. This slab was cut into pieces, and the JETquick Spin Column Technique (Genomed, Inc., Florida, USA) was used according to the manufacturer's instructions. Briefly, 300 µL of gel solubilisation solution L1 ($NaClO_4$, Na acetate and TBE) were added for each 100 mg of the gel slice pieces and incubated at 55 °C for 15 min vortexing every 5 min. The mixture was loaded into a prepared JETquick column and it was centrifuged at maximum speed for 1 min. The flowthrough was discarded. 700 µL of washing and reconstituted solution L2 (ethanol, NaCl, EDTA and Tris-

HCl) were added into the spin column and after 2 min, the column was centrifuged in the same conditions twice. The flowthrough was again discarded both times. To elute the bound DNA, 30-50 µL of pre-warmed sterile water were added onto the centre of the silica matrix of the spin column and the system was centrifuged for 2 min. The amount of DNA recovered was approximately calculated by UV fluorescence of intercalated ethidium bromide in an agarose gel electrophoresis.

6.13. **Preparation of bacterial competent cells**

Bacterial competent cells were prepared following the method described by Chung and Niemela (Chung, Niemela et al. 1989). *E. coli* strains were grown overnight in 10 mL of LB at 37°C. The following day, 140 mL of fresh LB were added and the cells were grown in the shaker at room temperature for 30-45 min until the OD_{600} was 0.3-0.38 at l=600nm. The cells were transferred into a Falcon 50ml tube and then placed in ice and centrifuged at 4 °C and 1000 rpm for 10 min. The pellet was resuspended in 1/10 mL of the initial volume of cold TSS solution (10% w/v PEG molecular weight 4000, 5% v/v DMSO, 35mM Mg Cl_2, pH 6.5 in LB medium). The cells were aliquoted, rapidly freeze in liquid nitrogen and stored at –80°C. Competence was determined by transformation with 0.1 ng of pUC19 and was deemed satisfactory if this procedure resulted in more than 100 colonies.

6.14. **Transformation of bacteria**

Transformations of ligation reactions were performed using 1/2 of the reaction volume. Transformation of clones was carried out using 1 ng of the plasmid DNA. The DNA was incubated with 60 µL of competent cells for 20 min on ice and at 42°C for 2 minutes. After the step of the heat shock, 60 µL of LB were added and the bacteria allowed to recover for 10 min at 37 °C. The cells were

then spread onto agarose plates containing the appropriate antibiotic. The plates were then incubated for 12-15 hours at 37 °C. When DNA inserts were cloned into β-galactosidase-based virgin plasmids, 30 µL of IPTG 100 mM and 20 µL of X-Gal (4 % w/v in dimethylformamide) were spread onto the surface of the agarose before plating to facilitate screening of positive clones (white colonies) through identification of β-galactosidase activity (blue colonies) which indicates the negative clones.

6.15. **Amplification of selected DNA fragments**

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

6.16. **Sequence analysis for cloning purpose**

Sequence analysis of plasmid DNA was performed using the CEQ 2000 sequencer (Beckman Coulter). The plasmid DNA of interest (approximately 100 ng) was purified through a MicroSpin S-400 HR Column (Amersham Pharmacia Biotech). The DNA was then amplified using fluorescent labeled dideoxynucleotide terminators according to the manufacturer's instructions. The samples were analyzed by loading them into the automatic sequencer.

6.17. **Construction of minigene system**

A minigene, as its name indicates, is a simplified version of a gene. In this thesis the minigene used is a hybrid construct, the pTB minigene, containing exons from α-globin and fibronectin, under the control of the α-globin promoter and SV40 enhancer (Fig. 6.1) (Pagani et al. 2003a). The intronic region between the two fibronectin exons contains a unique *Nde*I site which facilitates subcloning of an exon with flanking intronic regions.This system allows to insert *wt* or mutant exons, and study their impact on splicing outcome. In the presence of a *wt* exon, the splicing pattern of the minigene should be equivalent to that of the endogenous exon in the specific tissue or organ. On the other hand, the presence of mutations may affect pre-mRNA processing, causing aberrant splicing pattern due to exon skipping, intron retention, nonsense-mediated decay (NMD) or activation of cryptic sites. If in presence of a mutation the splicing pattern is not altered, it can be considered as neutral.

Oligonucleotides for PCR and cloning	
FIXex5 dir	5'-catatggttatacattaataaatag-3'
FIXex5 rev	5'-catatgcagaaatcacacaaattaattgc-3'
Alfa2-3	5'-caacttcaactcctaagccactgc-3'
Bra2	5'-gtcaccaggaagttggttaaatca-3'
pCI	5'-gactcactataggctagcctcg-3'
E8-75	5'-aagtacttacctgtaacgcttcacattccagatctgtc-3'
pBsK-FIXdir	5'-ggtaaattggaagagtttgt-3'
pBsK-FIXrev	5'-attaacgatagagcctccac-3'
U7CF11-dir	5'-acagaggcctttccgcaaataagtaaggtattcaaatttttgg-3'
SP6	5'-atttaggtgacactatag-3'
pTB2160-dir	5'-tacttcagatattatgtctagg-3'
pTB2270-rev	5'-cccatgtgagatatctagcg-3'

Table 6.1. List of oligonucleotides used for PCR amplification of spliced transcripts and cloning.

6.17.1. Generation of pCF ex12 and pFIX ex5 minigenes

For this thesis work I have used several human *CFTR* exon 12 and *F9* exon 5 minigene constructs. Human pCF ex12 *wt*, 17A and 52T minigenes have been previously described (Pagani et al. 2003a, 2005). *CFTR* exon 12 donor site mutations S589N (c.1766G>A), S589I (c.1766G>T), c.1766+2T>C, c.1766+3A>C, c.1766+3>G and c.1766+5G>A were introduced into pCF ex12 *wt* by by PCR-mediated site directed mutagenesis into the *Acc*I-*BamH*I cassette. They were called pCF exon 12 -1A, -1T, +2C, +3C, +3G and +5A, respectively. A DNA fragment of 156 bp containing the human *F9* exon 5 along with intronic flanking regions was amplified with PstI-dir and XbaI-rev oligonucleotides (Table 6. 1). Subsequently this amplicon was cloned in the *Pst*I/*Xba*I cassette of pTB minigene (Figure 6.1) and the resulting construct was called pFIX ex5 *wt*. Donor site (c.17896 (A>G, A>C and A>T; c.17897G>T, c.17897+1G>A, c.17897+2T>C, c.17897+4A>G and c.17897+13A>G) and polypyrimidine tract (c.17660T>G and c.17661T>G) mutations in exon 5 were introduced into pFIX ex5 *wt* by PCR mediated site-directed mutagenesis. They were called pFIX ex5 -2G, pFIX ex5 -2C, pFIX ex5 -2T, pFIX ex5 -1T, pFIX ex5 -1A, pFIX ex5 +1A, pFIX ex5 +2C, pFIX ex5 +4G and pFIX ex5 +13G, respectively. The *CFTR* exon 12 and *F9* exon 5 intronic deletion mutant minigenes and the *CFTR* exon 12 ISS mutants were created by overlapping PCR using suitable primers on pCF ex12 or pFIX ex5 *wt* as templates, accordingly. The first round of PCRs was performed using two sets of primers, combining an external primer with an internal one. The external oligonucleotides were pTB2160-dir and pTB2270-rev (Table 6.1). The internal oligonucleotides carried either the deletions (Table 6.3) or the point mutations (Table 6.4) and annealed downstream the 5' ss. The second round of PCR was carried out using the external primers and the two amplified products as template, obtained from the previous two separated PCRs. Finally the mutant amplicons were cloned into the *Acc*I/*BamH*I or

*Pst*I/*Xba*I cassettes of pCF exon 12 or pFIX exon 5, accordingly. Deletion mutants Δ12-15, Δ17-23, Δ24-34, Δ34-44, Δ45-55 and Δ55-65 were subcloned into *CFTR* exon 12 *wt* and 17A minigenes; deletion mutants Δ12-15 and Δ17-23 contexts into *CFTR* exon 12 -1A and +3G minigenes. *F9* exon 5 intronic deletion mutants Δ8-10, Δ11-13, Δ15-19 and Δ19-24 were subcloned into pFIX ex5 *wt*. pCI-SMN2 exon 7 was obtained from Dr. Adrian R. Krainer (CSHL, NY, USA). The identity of all minigene constructs was ultimately confirmed through sequencing analysis.

6.17.2. Generation of splicing-competent factor IX exon 5 expression vector

Splicing competent expression vector pSC-FIX was synthesized by GeneScript Inc. (Piscataway, NJ, USA). It consists of a fragment of 6139 bp cloned into a pcDNA 3.1+ backbone using *Bgl*II and *Apa*I restriction sites. It carries a simian virus 40 promoter, factor IX cDNA sequence from exons 1 to 4, first 544 bp of intron 4, *Nde*I restriction site, last 312 bp of intron4, exon 5 (129 bp), first 287 bp of intron 5, *Nde*I restriction site, last 908 bp of intron 5, cDNA sequence of *F9* from exons 6 to 8 and simian virus 40 polyadenilation site. *F9* exon 5 *wt* and mutant *Nde*I-*Nde*I cassettes were taken from pFIX exon 5 minigenes and subcloned into pSC-FIX create *wt* and mutant factor IX expression vectors. Thus five pSC-FIX constructs were created: pSC-FIX exon 5^{wt}, exon $5^{PPT\text{-}8G}$, exon $5^{PPT\text{-}9G}$, exon 5^{-2C}, exon 5^{-2G} and exon 5^{-2T}.

6.18. **Expression vector for generation of U1 snRNAs**

The vector used for the production of U1 snRNAs was the pGEM3, a standard cloning vector. The *wt* U1 snRNA gene was cloned in *BamH*I restriction site (Fig. 6.2) and the region between *Bgl*II and *Bcl*I sites encoding for the 5' tail complementary to 5' ss was replaced with specific annealed oligos. The sequences of each oligo used for creation of modified U1 snRNAs are reported in Table 6.2. The identity of all modified U1 snRNAs expression vectors was

ultimately confirmed through sequencing analysis.

6.19. **Maintenance and analysis of cells in culture**

HeLa and BHK cell lines were grown in Dulbecco's modified Eagle's medium with Glutamax I (Gibco) (DMEM with glutamine, sodium pyruvate, pyridoxine and 4.5 g/l glucose) supplemented with 10% foetal calf serum (Euro Clone) and Antibiotic Antimycotic (Sigma) according to the manufacturer's instructions.

Plates containing a confluent monolayer of cells were treated with 0.1% w/v trypsin as follows. Cells washed with PBS solution, were incubated at 37° C with 1-2 mL of PBS/EDTA/trypsin solution (PBS containing 0.04% w/v EDTA and 0.1% w/v trypsin) for 2 minutes or until cells were dislodged. After adding 10 mL of media, cells were pelleted by centrifugation and resuspended in 5 mL pre-warmed medium. 1-2 mL of this cell suspension was added to 10 mL medium in a fresh plate and was gently mixed before incubation.

6.20. **Transfections**

The DNA used for transfections was prepared with JetStar purification kit (Genomed) as previously described. HeLa cells grown on six-well plates ($3x10^5$/well) were transfected performed using Effectene reagent (Qiagen) according to manufacturer's indications. Briefly 0.5 µg of each *CFTR* exon 12, *F9* exon 5 or *SMN2* exon 7 minigenes were mixed with 4 µL of Enhancer for each transfection and the mixture was incubated at room temperature for 5 minutes to allow the condensation of the DNA. Then, 5 µL of Effectene were added to the mixture and incubated for 10 minutes. After the addition of 500 µL of complete culture medium the mixture was added to the cells in 1,5 mL of the same medium and incubated at 37°C. After 24 h, cells were harvested and RNA isolation using TRIReagent (Ambion Inc) following manufacturer's protocol. Co-transfections of pCF ex12 or pFIX ex5 minigenes with ExSpeU1

snRNPs were carried out using a 1:1 ratio (i.e., 0.5 µg pCF ex12 *wt* and 0.5 µg ExSpeU1CF).

Co-transfection of antisense oligonucleotides with pCF ex12 or pFIX ex5 constructs was performed using Lipofectamine 2000 (Invitrogen SRL, Milano, Italia) following manufacturer's instructions.

Transfection of BHK cells with the splicing-competent pSC-FIX expression vectors was carried out done as described previously (Pinotti et al. 2009).

6.21. cDNA preparation and Polymerase Chain Reaction (PCR)

In order to synthesize cDNA, the 3 µg of total RNA extracted from cells were mixed with random primers (Finnzymes Oy., Vantaa, Finland) in a final volume of 20 µL. After denaturation at 65°C for 5 min the RNA and the primer were incubated for 1 hour at 37 °C in the following solution: 1X First Strand Buffer (Gibco), 0,1 mM DTT, 10 mM dNTPs, RNase inhibitor 20 U (Ambion) and Moloney murine leukemia virus reverse transcriptase 100 U (Gibco) and

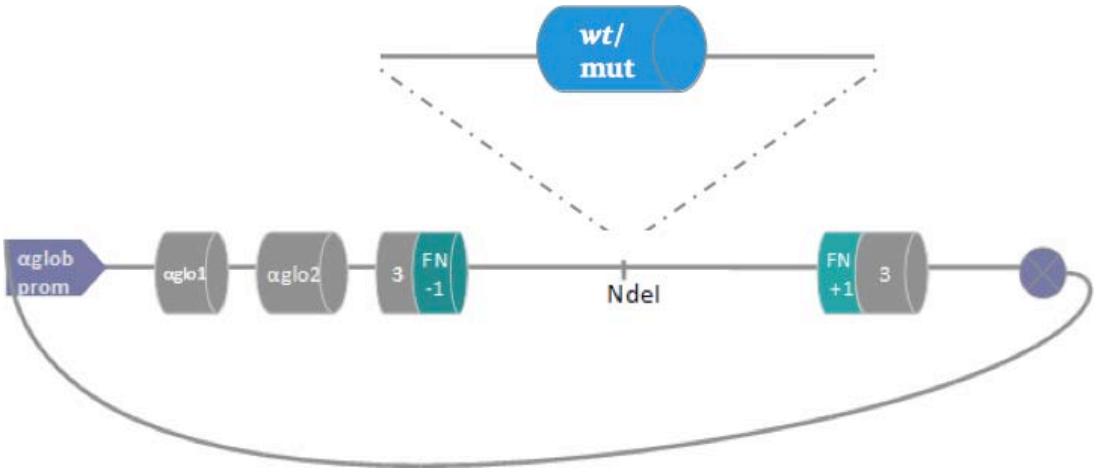

Figure 6.1. **Schematic representation of pTB NdeI- minigene.**

It contains at the 5' end a α-globin gene promoter and SV40 enhancer sequences to allow polymerase II transcription in the transfected cell lines. The reporter gene is composed by α-globin (grey) and fibronectin exons (green boxes) while at the 3' end a functionally competent polyadenylation site, derived from the α-globin gene, is present. The *Nde*I is used for subcloning of *wt* or mutant exon and flanking intronic regions.

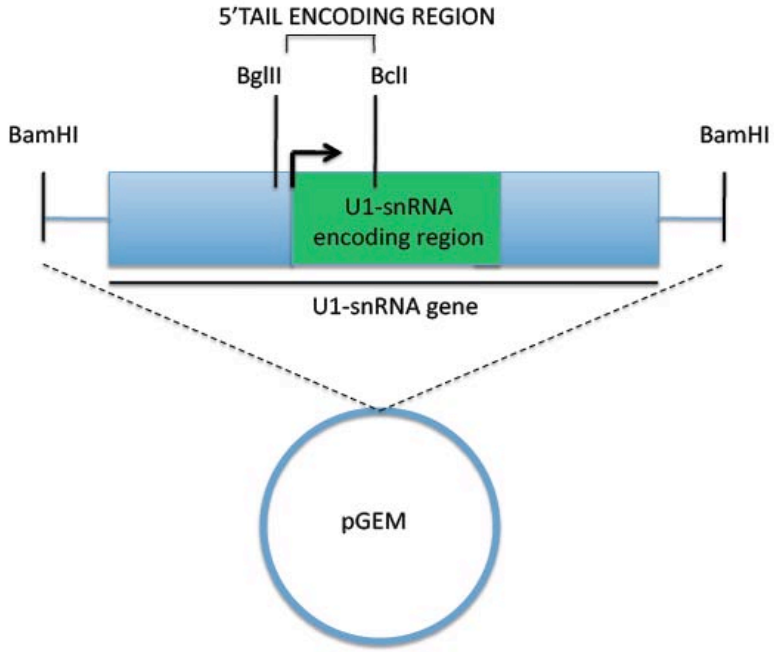

Figure 6.2. **Schematic representation of cloning strategy for creation of modified-U1-snRNAs.** The wild-type U1 snRNA gene was cloned in *BamH*I restriction site located in the multiple cloning site of pGEM3 vector. The region between *Bgl*II and *Bcl*I sites encoding for the 5'-tail of the U1 was replaced by specific annealed oligos for create engineered U1 snRNAs able to recognize specific target sequences with their modified 5'-tails.

retrotranscribed with poly(dT) primer. 3 µL of the cDNA reaction mix was used for the PCR analysis. To amplify only the spliced transcripts derived from pCF ex12 or pFIX ex5 *wt* and mutant, RT-PCR were carried out using as sense primer, alpha2,3 and an antisense primer, bra2, which recognize specific regions of these minigenes. To amplify spliced transcripts from pCI-SMN2 construct, pCI and E8-75 oligonucleotides were used as described previously (Hua et al. 2008). Splicing products from pSC-FIX exon 5 expression vectors were obtained using pSC-FIXdir and pSC-FIXrev oligonucleotides. Liver RNA (First Choice Human Total RNA Survey Panel, Ambion, Inc.) was retrotranscribed in standard conditions and amplified with FIX140 and FIX279 primers. The conditions used for the PCRs were the following: 94°C for 5 min for the initial denaturation, 94°C for 45 sec, 56°C for 45 sec, 72°C for 45 sec for 35 cycles, and 72°C for 10 min for the final extension. The PCRs were optimized

to be in the exponential phase of amplification. The results of all the transfections are the representative of at least three independent experiments. PCR products were resolved by 2% agarose gel electrophoresis. Quantification of exon inclusion was performed using ImageJ 1.38 software (http://rsb.info.nih.gov/ij/). The variability among different experiments was always <20%.

6.22. **Preparation of samples for capillary electrophoresis**

RNA extracted from cells was amplified using alpha2,3 and a labelled fluorescent bra2 primer (FAM). The PCR products were diluted 1:100 and then 1 µL of DNA was dehydrated at the temperature of 60°C for 20 min.

6.23. ***In silico*** **predictions**

RNA secondary structure predictions were performed (under default parameters) using the freely available mfold server. This web server uses mfold (version 3.5) by M. Zuker (2003): http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form.

Donor splice site scores were calculated using the following prediction tools:

*Shapiro & Senapathy consensus matrix (*S&S*)*

The Shapiro and Senapathy (S&S) consensus matrix is a position-specific weight matrix (PSWM) which reflects the degree of conservation at each position of the consensus 5' ss motif in an alignment of 1446 5' ss (Shapiro et al. 1987). The mammalian 5' ss consensus sequence is MAG|GURAGU (M = A or C; R = purine), and spans from position -3 (the third nucleotide from the 3' end of the upstream exon) to +6 (the sixth nucleotide in the intron). The mammalian 3' ss consensus sequence is CNNUUYUUYUYYYYUNCAG|G (Y= pyrymidine;

N= any base), spanning from position -14 to +3 (with respect to 3' ss). The S&S is one of the most commonly used 5' ss scoring methods, and it assumes independence between individual positions of the 9 nucleotide motif.

Maximal Entropy model (MaxEnt)

It can monitor the importance of dependencies between different positions by using a maximum-entropy distribution consistent with low-order marginal constraints estimated from available data (real and decoy splice sites). It considers local adjacent and nonadjacent dependencies (Yeo et al. 2004). This tool, together with MDD, MM and WMM, is available at http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html.

Maximum dependence decomposition model (MDD)

This tool relies on the assumption that there is interdependence between splice site positions and it considers conditional probabilites between adjacent and nonadjacent positions. It is a iterative decision-tree approach that emphasizes the strongest dependencies in the early branches of the tree (Burge et al. 1997).

First-order Markov model (MM)

It takes into account for nearest-neighbour dependencies, ignoring dependencies between nonadjacent positions.

Weight matrix models (WMM)

It assumes independence between nucleotide positions, quantifying the relative likelihood of candidate splice site sequence with respect to the background nucleotide distribution from a training set of signals.

Neural Network method (NN)

The NN algorithm is a machine-learning approach that recognizes sequence

patterns once it is trained with a set of DNA sequences encompassing authentic 5' ss. It is available at http://www.fruitfly.org/ seq_tools/ splice.html .

Analyzer Splice Tool (AST)

It is a PSWM, based on S&S calculation, using instead an updated splice site database (Carmel et al. 2004) The tool can be consulted at http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm.

Free energy calculation (ΔG)

It predicts the variation of free energy that comes from the base pairing of the 5' ss with the U1 snRNP 5'-tail (Carmel et al. 2004). This calculation can be obtained at http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm.

SplicePort

It is based on a feature generation algorithm for sequence classification, which searches on a large space of sequence-based features (nucleotide composition and position) to identify the predictive features (Dogan et al. 2007). A support vector machine classifier uses the identified features to produce splice site predictions. The software can be used at http://spliceport.cs.umd.edu/.

H-bond Model

It is a calculation of the hydrogen bond pattern between the 5' ss and all 11 nucleotides of the free 5'-tail of U1 snRNA at individual positions and also considers nonadjacent nucleotides dependencies. This tool can be consulted at http://www.uni-duesseldorf.de/rna/html/hbond_score.php.

Human Splice Finder (HSF)

It is developed on modified matrices derived from S&S original calculation,

using algorithms derived from the Universal Mutation Database (UMD) for evaluation of 5' and 3' ss and branchpoint scores (Desmet et al. 2009). It is available at http://www.umd.be/HSF/.

The default parameters of all these prediction tools were always used for estimation of score values, except for NN in which basal score was set to 0.01.

6.24. **Affinity purification of RNA binding proteins and Western Blot analysis**

Synthetic RNA oligonucleotides used as targets for pulldown assays were synthesized by Integrated DNA Technologies (Leuven, Belgium). Sequences of these RNA oligonucleotides are listed in table 6. 5.

Briefly 200 µL of agarose resin (for each sample) were washed with 10 mL of 0.1 M NaAcetate pH 5. After shaking, it was centrifuged (3000 rpm) at room temperature for 5 minutes. Washing was repeated 4 times (5 min/each) anf inally 0.1 M NaAcetate was added to reach a final volume of 300 µL. The agarose beads (300 µL)were then incubated with 150 µL of RNA oligo an let for overnight incubation on a rotator in the dark at 4°C. The next day beads were pelleted and washed 3 times with 1 mL NaCl. The pellet was washed (3 times) with 1 mL of 1X solution D [20 mM Hepes pH 7.5; 100 mM Kcl; 0.2 mM EDTA; 0.5 mM DTT; glycerol 6% and water for a final volume of 50 mL]. The nuclear extract mix was prepared for each sample [50 µL 10X solution D; 50 µL 1 M Kcl; 100 µL HeLa NE+; 100 µL H_2O; heparin 2.5 mg/mL final concentration] and 300 µL of the NE mix were added to the resin (final volume 500 µL) and incubated for 30 min on a rotator at room temperature. Afterwards, samples were centrifugated at room temperature, 4000 rpm for 5 min., and the pellet was washed 6 times using 1 mL 1X solution D (last 3 times using the rotator). 100 µL of SDS-PAGE sample buffer were added and cracked for 4 min. at 94°C for final loading onto a 12% polyacrylamide gel. After electrophoresis the gel was electroblotted on a PVDF membrane according to standard protocols

(Amersham Biosciences) and the membrane was blocked with 1X PBS-5% skimmed milk. Proteins were identified using different antibodies and Western blots signals were detected with a chemiluminescence kit (ECL, Pierce Biotechnology). For hnRNP A1/A2, a rabbit polyclonal anti-hnRNP A1 serum was used (generously provided by R. Klima (ICGEB, Trieste, Italy)).

6.25. **Factor IX activity and protein assays.**

Factor IX coagulant activity was assessed by aPTT assays (Bernardi et al 2009). Factor IX antigen levels in conditioned medium were evaluated by ELISA (Factor IX antigen F.IX; Affinity Biologicals, Ancaster, Canada). For Western blot analysis, 26 µl of conditioned medium were incubated 5 min at 95°C and run on 4-12% SDS-PAGE (NuPAGE Bis-Tris gel, Invitrogen®; Carlsbad, CA). Proteins were transferred onto a 0.2 µm nitrocellulose membrane (Whatman®, Dassel, Germany), which was blocked overnight with PBS buffer supplemented with 0.1% Tween-20 (PBS-T) and 5% low fat dry milk (Bio-Rad, Hercules, CA). Membranes were then incubated for 3 hours at room temperature with an anti-Human F.IX peroxidase conjugated (GAFIX-APHRP; Affinity Biologicals, Ancaster, Canada). The Supersignal® West Femto reagent (Thermo Scientific, Rockford, IL) was exploited for detection. Plasma derived factor IX or rFIX-wt were used to optimize the assay.

Oligonucleotide		Sequence
Modified U1 snRNAs for FIX exon 5		
3C	oligo rev	5'- gatcatggtatctcccctgccaggtcagttatga-3'
	oligo dir	5'- gatctcataactgacctggcagggagataccat-3'
5T6A	oligo rev	5'- gatcatggtatctcccctgccaggtaatatatga-3'
	oligo dir	5'- gatctcatatattacctggcaggggagataccat-3'
ex5 wt	oligo rev	5'- gatcatggtatctcccctgccaggtcatatatga-3'
	oligo dir	5'- gatctcatatatgacctggcaggggagataccat-3'
ex5 -2C	oligo rev	5'- gatcatggtatctcccctgcccggtcatatatga-3'
	oligo dir	5'- gatctcatatatgaccgggcaggggagataccat-3'
ex5 -2G	oligo rev	5'- gatcatggtatctcccctgccgggtcatatatga-3'
	oligo dir	5'- gatctcatatatgaccgcgcaggggagataccat-3'
ex5 -2T	oligo rev	5'- gatcatggtatctcccctgcctggtcatatatga-3'
	oligo dir	5'- gatctcatatatgaccgagcaggggagataccat-3'
ex5 -1T	oligo rev	5'- gatcatggtatctcccctgccatgtcatatatga-3'
	oligo dir	5'- gatctcatatatgacatggcaggggagataccat-3'
ex5 +1A	oligo rev	5'- gatcatggtatctcccctgccagatcatatatga-3'
	oligo dir	5'- gatctcatatatgatctggcaggggagataccat-3'
ex5 +2C	oligo rev	5'- gatcatggtatctcccctgccaggccatatatga-3'
	oligo dir	5'- gatctcatatatggcctggcaggggagataccat-3'
Exon-specific U1 snRNAs for FIX exon 5		
fix -7	oligo rev	5'-gatcatggtatctcccctgcccagcaggtcatatga-3'
	oligo dir	5'-gatctcatatgacctgctgggcaggggagataccat-3'
fix 1	oligo rev	5'-gatcatggtatctcccctgcgtcataatctatga-3'
	oligo dir	5'-gatctcatagattatgacgcaggggagataccat-3'
fix 7	oligo rev	5'- gatcatggtatctcccctgcatctgaataagatga-3'
	oligo dir	5'- gatctcatcttattcagatgcaggggagataccat-3'
fix 9	oligo rev	5'- gatcatggtatctcccctgcctgaataagaatga-3'
	oligo dir	5'- gatctcattcttattcaggcaggggagataccat-3'
fix 10	oligo rev	5'- gatcatggtatctcccctgctgaataagatatga-3'
	oligo dir	5'- gatctcatatcttattcagcaggggagataccat-3'
fix 13	oligo rev	5'-gatcatggtatctcccctgcataagattttatga-3'
	oligo dir	5'-gatctcataaaatcttatgcaggggagataccat-3'
fix 16	oligo rev	5'-gatcatggtatctcccctgcagatttttttatatga-3'
	oligo dir	5'-gatctcatataaaaaatctgcaggggagataccat-3'
fix 22	oligo rev	5'-gatcatggtatctcccctgctttaaagaaatatga-3'
	oligo dir	5'-gatctcatatttctttaaagcaggggagataccat-3'

fix 33	oligo rev	5'- gatcatggtatctcccctgctctgtatctgaatga-3'
	oligo dir	5'- gatctcattcagatacagagcaggggagataccat-3'
fix 38	oligo rev	5'- gatcatggtatctcccctgcatctgaaactatga-3'
	oligo dir	5'- gatctcatagtttcagatgcaggggagataccat-3'
fix 63	oligo rev	5'- gatcatggtatctcccctgcacctacataaatga-3'
	oligo dir	5'- gatctcatttatgtaggtgcaggggagataccat-3'
Modified U1 snRNAs for *CFTR* exon 12		
ex12 wt	oligo rev	5'- gatcatggtatctcccctgccaggtatgtatga-3'
	oligo dir	5'- gatctcatacatacctggcaggggagataccat-3'
ex12 -1A	oligo rev	5'- gatcatggtatctcccctgccaagtatgtatga-3'
	oligo dir	5'-gatctcatacatacttggcaggggagataccat-3'
ex12 -1T	oligo rev	5'-gatcatggtatctcccctgccatgtatgtatga-3'
	oligo dir	5'-gatctcatacatacatggcaggggagataccat-3'
ex12 +3C	oligo rev	5'-gatcatggtatctcccctgccaggtctgtatga-3'
	oligo dir	5'-gatctcatacagacctggcaggggagataccat-3'
ex12 +3G	oligo rev	5'-gatcatggtatctcccctgccaggtgtgtatga-3'
	oligo dir	5'-gatctcatacacacctggcaggggagataccat-3'
ex12+5A	oligo rev	5'-gatcatggtatctcccctgccaggtatatatga-3'
	oligo dir	5'-gatctcatatatacctggcaggggagataccat-3'
Exon-specific U1 snRNAs for *CFTR* exon 12		
cf 1	oligo rev	5'-gatcatggtatctcccctgcgtatgttctttgaga-3'
	oligo dir	5'-gatctctcaaagaacatacgcaggggagataccat-3'
cf 9	oligo rev	5'-gatcatggtatctcccctgctttgaatacctatga-3'
	oligo dir	5'-gatctcataggtattcaaagcaggggagataccat-3'
cf 11	oligo rev	5'-gatcatggtatctcccctgctgaataccttacttatga-3'
	oligo dir	5'-gatctcataagtaaggtattcagcaggggagataccat-3'
cf 15	oligo rev	5'-gatcatggtatctcccctgctaccttacttataaga-3'
	oligo dir	5'-gatctcttataagtaaggtagcaggggagataccat-3'
cf 33	oligo rev	5'-gatctctattttagcatgagcaggggagataccat-3'
	oligo dir	5'-gatcatggtatctcccctgctcatgctaaaataga-3'

Table 6.2. List of oligonucleotides used for construction of modified and exon-specific U1 snRNAs.

CFTR exon 12 intronic deletion mutants		
Δ12-15	oligo rev	5'-tataagtaaggtaaagaacatacctttcaaata-3'
	oligo dir	5'-aggtatgttctttaccttacttataatgctcatgc-3'
Δ17-23	oligo rev	5'-tgagcattatatattcaaagaacatacctttc-3'
	oligo dir	5'-ttctttgaatatataatgctcatgctaaaataa-3'
Δ24-34	oligo rev	5'-tattttagcaagtaaggtattcaaagaacatac-3'
	oligo dir	5'-taccttacttgctaaaataaaagaaagacagac-3'
Δ34-44	oligo rev	5'-tctttctttgagcattataagtaaggtattc-3'
	oligo dir	5'-cttataatgctcaaagaaagacagactgtcc-3'
Δ45-55	oligo rev	5'-atggatcctatgatgggacagtctattttagcatgagcattataa-3'
	oligo dir	5'-ttataatgctcatgctaaaatagactgtcccatcataggatccat-3'
Δ55-65	oligo rev	5'-atggatcctatgagtctttcttttattttagcatgagc-3'
	oligo dir	5'-gctcatgctaaaataaaagaaagactcataggatccat-3'
FIX exon 5 intronic deletion mutants		
Δ8-10	oligo rev	5'-cagattttctttaaaaaatcttattcttatgacct-3'
	oligo dir	5'-ctgtgaaccagcaggtcataagaataagatt-3'
Δ11-13	oligo rev	5'-atacagattttctttaaaaaatcttaagattatgacctg-3'
	oligo dir	5'-accagcaggtcataatcttaagattttt-3'
Δ15-19	oligo rev	5'-gatacagattttctttaaaaaattcagattatgacct-3'
	oligo dir	5'-ccagcaggtcataatctgaatttttaaagaaaat-3'
Δ19-24	oligo rev	5'-gtttcagatacagattttctttttcttattcagattatga-3'
	oligo dir	5'-agcaggtcataatctgaataagaaaagaaaatctgta-3'

Table 6.3. Oligonucleotides for preparation of CFTR exon 12 and FIX exon 5 intronic deletion mutants.

CFTR exon 12 ISS mutagenesis		
+12A	oligo rev	5'-gtaaggtatttaaagaacatacctttc-3'
	oligo dir	5'-gaaaggtatgttctttaaataccttac-3'
+13C	oligo rev	5'-gtaaggtatgcaaagaacatacctttc-3'
	oligo dir	5'-gaaaggtatgttctttgcataccttac-3'
+15C	oligo rev	5'-gtaaggtgttcaaagaacatacctttc-3'
	oligo dir	5'-gaaaggtatgttctttgaacaccttac-3'
+16T	oligo rev	5'-gcattataagtaaggaattcaaagaac-3'
	oligo dir	5'-gttctttgaattccttacttataatgc-3'
+16G	oligo rev	5'-gatctctattttagcatgagcaggggagataccat-3'
	oligo dir	5'-gatcatggtatctcccctgctcatgctaaaataga-3'
+17A	oligo rev	5'-gcattataagtaagttattcaaagaac-3'
	oligo dir	5'-gttctttgaataacttacttataatgc-3'
+18G/A	oligo rev	5'-gcattataagtaagytattcaaagaac-3'
	oligo dir	5'-gttctttgaatarcttacttataatgc-3'
+19G	oligo rev	5'-gcattataagtaacgtattcaaagaac-3'
	oligo dir	5'-gttctttgaatacgttacttataatgc-3'
+20G/A	oligo rev	5'-gcattataagayggtattcaaagaac-3'
	oligo dir	5'-gttctttgaataccrtacttataatgc-3'
+21G	oligo rev	5'-gcattataagaacgtattcaaagaac-3'
	oligo dir	5'-gttctttgaataccttgcttataatgc-3'
+22A/T	oligo rev	5'-gcattataagwaaagtattcaaagaac-3'
	oligo dir	5'-gttctttgaatacctttwttataatgc-3'
+23G	oligo rev	5'-gcattatacgtaaggtattcaaagaac-3'
	oligo dir	5'-gttctttgaataccttacgtataatgc-3'

Table 6.4. Oligonucleotides used for CFTR exon 12 ISS mutagenesis by overlapping PCR.

RNA oligonucleotides for pulldown assay	
CFex12 *wt*	5'-aagguauguucuuugaauaccuuacuuaua-3'
CFex12 Δ12-15	5'-aagguauguucuuuaccuuacuuaua-3'
FIX ex5 *wt*	5'-caggucauaaucugaauaagauuuuuua-3'
FIX ex5 Δ15-19	5'-caggucauaaucugaauuuuuua-3'
Antisense oligonucleotides (ASO)	
ASO^{CFTR}_{20}	5'-tgaataccttacttataatg-3'
ASO^{CFTR}_{25}	5'-tgaataccttacttataatgctcat-3'
ASO^{FIX}_{20}	5'-tgaataagatttttaaaga-3'
ASO^{FIX}_{25}	5'-tgaataagatttttaaagaaaatc-3'

Table 6.5. RNA oligonucleotides for pulldown assays and antisense oligonucleotides targeting CFTR exon 12 and F9 exon 5 ISS

7.References

Abad, X., Vera, M., Jung, S. P., Oswald, E., Romero, I., Amin, V., Fortes, P., and Gunderson, Samuel I (2008). Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. Nucleic Acids Research 36, 2338-52.

Adams, D. J., van der Weyden, L, Mayeda, A, Stamm, S, Morris, B. J., and Rasko, J. E. (2001). ZNF265--a novel spliceosomal protein able to induce alternative splicing. The Journal of cell biology 154, 25-32.

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

Amaral, M. D. et al. (2004). Quantitative methods for the analysis of CFTR transcripts/splicing variants. Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society 3 Suppl 2, 17-23.

Angelis, F. G. De, Sthandier, O., Berarducci, B., Toso, S., Galluzzi, G., Ricci, E., Cossu, G., and Bozzoni, I. (2002). Chimeric snRNA molecules carrying antisense sequences against the splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in del48-50 DMD cells. Proceedings of the National Academy of Sciences 99, 1-6.

Anson, D. S., Choo, K. H., Rees, D. J. G., Giannelli, F., Gould, K., Huddleston, J. A., and Brownlee, G. G. (1984). The gene structure of human anti-haemophilic factor IX. The EMBO journal 3, 1053-1060.

Ashe, M. P., Furger, a, and Proudfoot, N. J. (2000). Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. RNA (New York, N.Y.) 6, 170-7.

Asparuhova, M. B., Marti, G., Liu, S., Serhan, F., Trono, D., and Schümperli, D (2007). Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing Tat and Rev exon skipping. The journal of gene medicine 9, 323-34.

Ast, G. (2004). How did alternative splicing evolve? Nature Reviews Genetics 5, 773-782.

Aznarez, I., Chan, E. M., Zielenski, J, Blencowe, B. J., and Tsui, L.(2003). Characterization of disease-associated mutations affecting an exonic splicing enhancer and two cryptic splice sites in exon 13 of the cystic fibrosis transmembrane conductance regulator gene. Human Molecular Genetics 12, 2031-2040.

Babich, A., Nevins, J. R., and Darnell, J. E. (1980). Early capping of transcripts from the adenovirus major late transcription unit. Nature 287, 246-8.

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

Bakkour, N. et al. (2007). Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. PLoS pathogens 3, 1530-9.

Baralle, D., Lucassen, A., and Buratti, E (2009). Missed threads. The impact of pre-mRNA splicing defects on clinical practice. EMBO reports 10, 810-6.

Baralle, D., and Baralle, M. (2005). Splicing in action: assessing disease causing sequence changes. Journal of medical genetics 42, 737-48.

Baralle, M., Baralle, D., Conti, L. De, Mattocks, C., Whittaker, J., Knezevich, A., Ffrench-constant, C., and Baralle, F E (2003). Identification of a mutation that perturbs NF1 gene splicing using genomic DNA samples and a

minigene assay. Journal of Medical Genetics 40, 220-222.

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

Bass, B. L., and Weintraub, H. (1988). An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55, 1089-98.

Batsché, E., Yaniv, M., and Muchardt, C. (2006). The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nature structural & molecular biology 13, 22-9.

Beckley, S. A., Liu, P., Stover, M. L., Gunderson, S I, Lichtler, A. C., and Rowe, D. W. (2001). Reduction of target gene expression by a modified U1 snRNA. Molecular and cellular biology 21, 2815-2825.

Berasain, C. (2010). Impairment of pre-mRNA splicing in liver disease: Mechanisms and consequences. World Journal of Gastroenterology 16, 3091.

Berget, S M (1995). Exon recognition in vertebrate splicing. The Journal of biological chemistry 270, 2411-4.

Berglund, J. A., Chua, K., Abovich, N., Reed, R, and Rosbash, M (1997). The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell 89, 781-7.

Bernstein, L. B., Manser, T., and Weiner, A M (1985). Human U1 small nuclear RNA genes: extensive conservation of flanking sequences suggests cycles of gene amplification and transposition. Molecular and cellular biology 5, 2159-71.

Betsalel, O. T. et al. (2010). Characterization of novel SLC6A8 variants with the use of splice-site analysis tools and implementation of a newly developed LOVD database. European Journal of Human Genetics 19, 56-63.

Bindereif, A., and Green, M R (1987). An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. The EMBO journal 6, 2415-24.

Black, D L, Chabot, B, and Steitz, J A (1985). U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell 42, 737-50.

Black, D L (2003). Mechanisms of alternative pre-messenger RNA splicing. Annual review of biochemistry 72, 291-336.

Blanchette, M., and Chabot, B (1997). A highly stable duplex structure sequesters the 5' splice site region of hnRNP A1 alternative exon 7B. RNA (New York, N.Y.) 3, 405-19.

Blaustein, M. et al. (2005). Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT. Nature structural & molecular biology 12, 1037-44.

Bompadre, S. G., Ai, T., Cho, J. H., Wang, Xiaohui, Sohma, Y., Li, M., and Hwang, T.-C. (2005). CFTR gating I: Characterization of the ATP-dependent gating of a phosphorylation-independent CFTR channel (DeltaR-CFTR). The Journal of general physiology 125, 361-75.

Bompadre, S. G., Cho, J. H., Wang, Xiaohui, Zou, X., Sohma, Y., Li, M., and Hwang, T.-C. (2005). CFTR gating II: Effects of nucleotide binding on the stability of open states. The Journal of general physiology 125, 377-94.

Bourgeois, Cyril F, Lejeune, F., and Stévenin, J. (2004). Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. Progress in nucleic acid research and molecular biology 78, 37-88.

Bowen, D. J. (2002). Haemophilia A and haemophilia B: molecular insights. Molecular pathology : MP 55, 127-44.

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

Bruce, S. R., and Peterson, M. L. (2001). Multiple features contribute to efficient constitutive splicing of an unusually large exon. Nucleic acids research 29, 2292-302.

Brun, C., Suter, D., Pauli, C., Dunant, P., Lochmuller, H., Burgunder, J., Schumperli, D., and Weis, J. (2003). U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping. Cellular and Molecular Life Sciences 60, 557-566.

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

Buratti, E, and Baralle, F E (2004). Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. Molecular and cellular biology 24, 10505-10514.

Burge, C B., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. Journal of molecular biology 268, 78-94.

Burset, M., Seledtsov, I. a, and Solovyev, V. V. (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic acids research 28, 4364-75.

Calarco, J. A, Xing, Y., Cáceres, M., Calarco, J. P., Xiao, X., Pan, Q., Lee, C., Preuss, T. M., and Blencowe, B. J. (2007). Global analysis of alternative splicing differences between humans and chimpanzees. Genes & development 21, 2963-75.

Cao, W., and Garcia-Blanco, M. A (1998). A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding.

The Journal of biological chemistry 273, 20629-35.

Carbon, P., Murgo, S., Ebel, J. P., Krol, A., Tebb, G., and Mattaj, L. W. (1987). 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 51, 71-9.

Carmel, I., Tal, S., and Ast, G. (2004). Comparative analysis detects dependencies among the 5' splice-site positions. Rna 10, 828-840.

Carothers, A. M., Urlaub, G., Grunberger, D., and Chasin, L A (1993). Splicing mutants and their second-site suppressors at the dihydrofolate reductase locus in Chinese hamster ovary cells. Molecular and cellular biology 13, 5085-5098.

Cartegni, L, Maconi, M., Morandi, E., Cobianchi, F., Riva, S., and Biamonti, G (1996). hnRNP A1 selectively interacts through its Gly-rich domain with different RNA-binding proteins. Journal of molecular biology 259, 337-48.

Cartegni, L, Chew, Shern L, and Krainer, A R (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nature reviews. Genetics 3, 285-98.

Cartegni, L, Wang, J, Zhu, Z., Zhang, M Q, and Krainer, A R (2003). ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic acids research 31, 3568-71.

Cartegni, L, and Krainer, A R (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nature genetics 30, 377-84.

Castle, J. C., Zhang, C., Shah, J. K., Kulkarni, A. V., Kalsotra, A., Cooper, T. A and Johnson, J. M. (2008). Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. Nature genetics 40, 1416-25.

Chabot, B, Blanchette, M., Lapierre, I., and Branche, H. L A (1997). An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. Molecular and cellular biology 17, 1776-86.

Chanfreau, G., and Jacquier, A (1993). Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron. The EMBO journal 12, 5173-80.

Chang, W.-C., Chen, H.-H., and Tarn, W.-Y. (2008). Alternative splicing of U12-type introns. Frontiers in bioscience : a journal and virtual library 13, 1681-90.

Chaudhury, A., Chander, P., and Howe, P. H. (2010). Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. RNA (New York, N.Y.) 16, 1449-62.

Cho, S., Hoang, A., Sinha, R., Zhong, X.-Y., Fu, X.-D., Krainer, A R, and Ghosh, G. (2011). Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. Proceedings of the National Academy of Sciences of the United States of America, 1-6.

Choe, Y. J. et al. (2010). Novel CFTR mutations in a Korean infant with cystic fibrosis and pancreatic insufficiency. Journal of Korean medical science 25, 163-5.

Cohen, J B, Snow, J E, Spencer, S D, and Levinson, A D. (1994). Suppression of mammalian 5' splice-site defects by U1 small nuclear RNAs from a distance. Proceedings of the National Academy of Sciences of the United States of America 91, 10470-4.

Coolidge, C. J., Seely, R. J., and Patton, J G (1997). Functional analysis of the polypyrimidine tract in pre-mRNA splicing. Nucleic acids research 25, 888-96.

Cooper, D N, and Tuddenham, E. G. (1994). Molecular genetics of familial venous thrombosis. British medical bulletin 50, 833-50.

Cooper, T. A (2005). Use of minigene systems to dissect alternative splicing elements. Methods (San Diego, Calif.) 37, 331-40.

Cooper, T. A, Wan, L., and Dreyfuss, G (2009). RNA and disease. Cell 136, 777-93.

Coppola, J. A., Field, A. S., and Luse, D. S. (1983). Promoter-Proximal Pausing by RNA Polymerase II in vitro: Transcripts Shorter Than 20 Nucleotides are not Capped. Proceedings of the National Academy of Sciences 80, 1251-1255.

Coulter, L. R., Landree, M. A., and Cooper, T A (1997). Identification of a new class of exonic splicing enhancers by in vivo selection. Molecular and cellular biology 17, 2143-50.

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

Cramer, P., Cáceres, J F, Cazalla, D., Kadener, S., Muro, A. F., Baralle, F E, and Kornblihtt, A. R. (1999). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Molecular cell 4, 251-8.

Cremonesi, L. et al. (1992). Four new mutations of the CFTR gene (541delC, R347H, R352Q, E585X) detected by DGGE analysis in Italian CF patients, associated with different clinical phenotypes. Human mutation 1, 314-9.

Das, R, and Reed, R (1999). Resolution of the mammalian E complex and the

ATP-dependent spliceosomal complexes on native agarose mini-gels . Rna 5, 1504-1508.

Das, R, Dufu, K., Romney, B., Feldt, M., Elenko, M., and Reed, R (2006). Functional coupling of RNAP II transcription to spliceosome assembly. Genes & development 20, 1100-9.

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

David, A, Miraki-Moud, F., Shaw, N. J., Savage, M. O., Clark, A J. L., and Metherell, L. A (2010). Identification and characterisation of a novel GHR defect disrupting the polypyrimidine tract and resulting in GH insensitivity. European journal of endocrinology / European Federation of Endocrine Societies 162, 37-42.

David, C. J., and Manley, J. L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. Genes & Development 24, 2343-2364.

David, C J, and Manley, J L (2008). The search for alternative splicing regulators: new approaches offer a path to a splicing code. Genes & development 22, 279-85.

Denti, M. A. et al. (2006). Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model. Proceedings of the National Academy of Sciences of the United States of America 103, 3758-63.

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, Mireille, and Béroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic acids research 37, e67.

Dietz, H. C., Valle, D., Francomano, C. A., Kendzior, R. J., Pyeritz, R. E., and

Cutting, G R (1993). The skipping of constitutive exons in vivo induced by nonsense mutations. Science (New York, N.Y.) 259, 680-3.

Dirksen, W. P., Mohamed, S. A., and Fischer, S. A. (2003). Splicing of a myosin phosphatase targeting subunit 1 alternative exon is regulated by intronic cis-elements and a novel bipartite exonic enhancer/silencer element. J. Biol. Chem. 278, 9722-9732.

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

Dogan, R. I., Getoor, L., Wilbur, W. J., and Mount, S M (2007). SplicePort--an interactive splice-site analysis tool. Nucleic acids research 35, W285-91.

Donahue, C. P., Muratore, C., Wu, Jane Y, Kosik, K. S., and Wolfe, M. S. (2006). Stabilization of the Tau Exon 10 Stem Loop Alters pre-mRNA splicing. The Journal of Biological Chemistry 281, 23302-23306.

Dreumont, N., Hardy, S., Behm-Ansmant, I., Kister, L., Branlant, C, Stévenin, J., and Bourgeois, C F (2010). Antagonistic factors control the unproductive splicing of SC35 terminal intron. Nucleic acids research 38, 1353-66.

Dreyfuss, G, Matunis, M. J., Piñol-Roma, S., and Burd, C. G. (1993). hnRNP proteins and the biogenesis of mRNA. Annual review of biochemistry 62, 289-321.

Dreyfuss, G, Kim, V. N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nature reviews. Molecular cell biology 3, 195-205.

Drögemüller, C. et al. (2011). An unusual splice defect in the mitofusin 2 gene

(MFN2) is associated with degenerative axonopathy in Tyrolean Grey cattle. PloS one 6, e18931.

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

Egloff, S., O'Reilly, D., Chapman, R. D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. (2007). Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. Science (New York, N.Y.) 318, 1777-9.

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

Esperante, S. a, Rivolta, C. M., Caputo, M., González-Sarmiento, R., and Targovnik, H. M. (2008). Identification and characterization of new variants of three associated SNPs and a microsatellite in the TSH receptor gene which are useful for genetic studies. Molecular and cellular probes 22, 281-6.

Evsyukova, I., Somarelli, J. A., Gregory, S. G., and Garcia-Blanco, M. A. (2010). Alternative splicing in multiple sclerosis and other autoimmune diseases. RNA Biology 7, 462-473.

Fagnani, M. et al. (2007). Functional coordination of alternative splicing in the mammalian central nervous system. Genome biology 8, R108.

Fairbrother, W G, and Chasin, L A (2000). Human genomic sequences that inhibit splicing. Molecular and cellular biology 20, 6816-25.

Fairbrother, W G, Yeh, R.-F., Sharp, P A, and Burge, C B (2002). Predictive identification of exonic splicing enhancers in human genes. Science (New York, N.Y.) 297, 1007-13.

Fairbrother, W GF, Yeo, G. W., Yeh, R., Goldstein, P., Mawson, M., Sharp, P A, and Burge, C B (2004). RESCUE-ESE identifies candidate exonic splicing

enhancers in vertebrate exons. Nucleic acids research 32, W187-90.

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

Fedorova, O., and Zingler, N. (2007). Group II introns: structure, folding and splicing mechanism. Biological chemistry 388, 665-78.

Feeney, R. J., and Zieve, G. W. (1990). Nuclear exchange of the U1 and U2 snRNP-specific proteins. The Journal of cell biology 110, 871-81.

Fierro-Monti, I., and Mathews, M. B. (2000). Proteins binding to duplexed RNA: one motif, multiple functions. Trends in biochemical sciences 25, 241-6.

Fisette, J.-F., Toutant, J., Dugré-Brisson, S., Desgroseillers, L., and Chabot, B (2010). hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection. RNA (New York, N.Y.) 16, 228-38.

Ford, E., Strubin, M., and Hernandez, N. (1998). 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 & development 12, 3528-40.

Frendewey, D., and Keller, W (1985). Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell 42, 355-67.

Freund, M. et al. (2003). A novel approach to describe a U1 snRNA binding site. Nucleic acids research 31, 6963-75.

Fukao, T., Horikawa, R., Naiki, Y., Tanaka, T., Takayanagi, M., Yamaguchi, S., and Kondo, N. (2010). A novel mutation (c.951C>T) in an exonic splicing enhancer results in exon 10 skipping in the human mitochondrial acetoacetyl-CoA thiolase gene. Molecular genetics and metabolism 100, 339-44.

Fukuhara, T. et al. (2006). Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. Proceedings of the National Academy of Sciences of the United States of America 103, 11329-33.

Ganot, P., Jády, B. E., Bortolin, M. L., Darzacq, X., and Kiss, T (1999). Nucleolar factors direct the 2'-O-ribose methylation and pseudouridylation of U6 spliceosomal RNA. Molecular and cellular biology 19, 6906-17.

Garcia-Blanco, M. A., Baraniak, A. P., and Lasda, E. L. (2004). Alternative splicing in disease and therapy. Nature Biotechnology 22, 535-546.

Garrido, E., Chabás, A., Coll, M. J., Blanco, M., Domínguez, C., Grinberg, D., Vilageliu, L., and Cormand, B. (2007). Identification of the molecular defects in Spanish and Argentinian mucopolysaccharidosis VI (Maroteaux-Lamy syndrome) patients, including 9 novel mutations. Molecular genetics and metabolism 92, 122-30.

Garrido, E., Cormand, B., Hopwood, J. J., Chabás, A., Grinberg, D., and Vilageliu, L. (2008). Maroteaux – Lamy syndrome : Functional characterization of pathogenic mutations and polymorphisms in the arylsulfatase B gene. Molecular Genetics and Metabolism 94, 305-312.

Gatto, F. Del, Plet, A., Gesnel, M., Fort, C., and Breathnach, R. (1997). Multiple interdependent sequence elements control splicing of a fibroblast growth factor receptor 2 alternative exon. Mol. Cell. Biol. 17, 5106-5116.

Geib, T., and Hertel, K J (2009). Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs. PloS one 4, e8204.

Gersappe, A, and Pintel, D. J. (1999). CA- and purine-rich elements form a novel bipartite exon enhancer which governs inclusion of the minute virus of mice NS2-specific exon in both singly and doubly spliced mRNAs. Molecular and cellular biology 19, 364-75.

Ghigna, C., Valacca, C., and Biamonti, G (2008). Alternative splicing and tumor progression. Current genomics 9, 556-70.

Giannelli, F., and Green, P. M. (1996). The molecular basis of haemophilia A and B. Baillière's clinical haematology 9, 211-28.

Glaus, E., Schmid, F., Costa, R. D., Berger, W., and Neidhardt, J. (2011). Gene Therapeutic Approach Using Mutation-adapted U1 snRNA to Correct a RPGR Splice Defect in Patient-derived Cells. Molecular Therapy, 1-6.

Goina, E., Skoko, N., and Pagani, F. (2008). Binding of DAZAP1 and hnRNPA1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant. Molecular and cellular biology 28, 3850-60.

Goldman, B. S., Sherman, D. A., and Kranz, R. G. (1997). Comparison of the bacterial HelA protein to the F508 region of the cystic fibrosis transmembrane regulator. Journal of bacteriology 179, 7869-71.

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

Gorman, L., Suter, D., Emerick, V., Schümperli, D, and Kole, R. (1998). Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. Proceedings of the National Academy of Sciences of the United States of America 95, 4929-34.

Goss, C. H., and Rosenfeld, M. (2004). Update on cystic fibrosis epidemiology. Current opinion in pulmonary medicine 10, 510-4.

Goyenvalle, A., Vulin, A., Fougerousse, F., Leturcq, F., Kaplan, J.-C., Garcia, L., and Danos, O. (2004). Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. Science (New York, N.Y.) 306, 1796-9.

Gozani, O., Feld, R., and Reed, R (1996). Evidence that sequence-independent

binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes & development 10, 233-43.

Grabowski, P J, Nasim, F H, Kuo, H.-C. C., and Burch, R. (1991). Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. Mol. Cell. Biol. 11, 5919-5928.

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

Graveley, B R, Hertel, K J, and Maniatis, T (1998). A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. The EMBO journal 17, 6747-56.

Graveley, B R, and Maniatis, T (1998). Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. Molecular cell 1, 765-71.

Gregory, R. J. et al. (1990). Expression and characterization of the cystic fibrosis transmembrane conductance regulator. Nature 347, 382-6.

Gridasova, A. A., and Henry, R W(2005). The p53 Tumor Suppressor Protein Represses Human snRNA Gene Transcription by RNA Polymerases II and III Independently of Sequence-Specific DNA Binding. Molecular and cellular biology 25, 3247-3260.

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

Gromoll, J., Lahrmann, L., Godmann, M., Müller, T., Michel, C., Stamm, S, and Simoni, M. (2007). Genomic checkpoints for exon 10 usage in the luteinizing hormone receptor type 1 and type 2. Molecular endocrinology (Baltimore, Md.) 21, 1984-96.

Grover, a, Houlden, H., Baker, M., Adamson, J., Lewis, J., Prihar, G., Pickering-Brown, S., Duff, K., and Hutton, M. (1999). 5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. The Journal of biological chemistry 274, 15134-43.

Guil, S., Gattoni, R., Carrascal, M., Abián, J., Stévenin, J., and Bach-Elias, M. (2003). Roles of hnRNP A1, SR proteins, and p68 helicase in c-H-ras alternative splicing regulation. Molecular and cellular biology 23, 2927-41.

Gunderson, S I, Polycarpou-Schwarz, M., and Mattaj, I. W. (1998). U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. Molecular cell 1, 255-64.

Guo, W., and Helfman, D. M. (1993). cis-elements involved in alternative splicing in the rat beta-tropomyosin gene: the 3'-splice site of the skeletal muscle exon 7 is the major site of blockage in nonmuscle cells. Nucleic acids research 21, 4762-8.

Guth, S., Martínez, C., Gaur, R. K., and Valcárcel, J (1999). Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). Molecular and cellular biology 19, 8263-71.

Guédard-Méreuze, S. L., Vaché, C., Molinari, N., Vaudaine, J., Claustres, Mireille, Roux, A.-F., and Tuffery-Giraud, S. (2009). Sequence contexts that determine the pathogenicity of base substitutions at position +3 of donor splice-sites. Human mutation 30, 1329-39.

Hamm, J., Dathan, N. A., Scherly, D., and Mattaj, I. W. (1990). Multiple domains of U1 snRNA, including U1 specific protein binding sites, are required for splicing. The EMBO journal 9, 1237-44.

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

Haque, A., Buratti, E, and Baralle, F E (2010). Functional properties and evolutionary splicing constraints on a composite exonic regulatory element of splicing in CFTR exon 12. Nucleic acids research 38, 647-659.

Hartmann, L. et al. (2010). 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. American journal of human genetics 87, 480-93.

Hartmann, L., Theiss, S., Niederacher, D., and Schaal, H. (2008). Diagnositcs of pathogenic splicing mutations: does bioinformatics cover all bases? Frontiers in Bioscience 13, 3252-3272.

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

Hernandez, N. (2001). Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. The Journal of biological chemistry 276, 26733-6.

Heyd, F., and Lynch, K. W. (2010). Phosphorylation-Dependent Regulation of PSF by GSK3 Controls CD45 Alternative Splicing. Molecular cell 40, 126-37.

Hiller, M., Zhang, Zhaiyi, Backofen, R., and Stamm, S (2007). Pre-mRNA Secondary Structures Influence Exon Recognition. PLoS Genetics 3, e204.

Hirsch, H. A., Gu, L., and Henry, R W (2000). The retinoblastoma tumor suppressor protein targets distinct general transcription factors to regulate RNA polymerase III gene expression. Molecular and cellular biology 20, 9182-91.

Hitomi, Y., Sugiyama, K., and Esumi, H. (1998). Suppression of the 5' splice site mutation in the Nagase analbuminemic rat with mutated U1snRNA.

Biochemical and biophysical research communications 251, 11-6.

Hocine, S., Singer, R. H., and Grünwald, D. (2010). RNA processing and export. Cold Spring Harbor perspectives in biology 2, a000752.

Hoffman, B. E., and Grabowski, P J (1992). U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. Genes & development 6, 2554-68.

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

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

Honoré, B. et al. (1995). Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromosomes. The Journal of biological chemistry 270, 28780-9.

Horabin, J. I., and Schedl, P. (1993). Sex-lethal autoregulation requires multiple cis-acting elements upstream and downstream of the male exon and appears to depend largely on controlling the use of the male exon 5' splice site. Molecular and cellular biology 13, 7734-46.

Houdayer, C. et al. (2008). Evaluation of in silico splice tools for decision-making in molecular diagnosis. Hum. Mutat 29, 975-982.

Hovhannisyan, R. H., and Carstens, R. P. (2005). 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. Molecular and cellular biology 25, 250-63.

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

Hua, Y., Vickers, T. A., Okunola, H. L., Bennett, C. F., and Krainer, A R (2008). Antisense Masking of an hnRNP A1 / A2 Intronic Splicing Silencer Corrects SMN2 Splicing in Transgenic Mice. American Journal of Human Genetics 82, 834-848.

Huang, Q., Jacobson, M. R., and Pederson, T. (1997). 3' processing of human pre-U2 small nuclear RNA: a base-pairing interaction between the 3' extension of the precursor and an internal region. Molecular and cellular biology 17, 7178-85.

Huang, Q., and Pederson, T. (1999). A human U2 RNA mutant stalled in 3' end processing is impaired in nuclear import. Nucleic acids research 27, 1025-31.

Huang, Y., and Steitz, JA (2005). SRprises along a messenger's journey. Molecular cell 17, 613-5.

Hutton, M. et al. (1998). Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. Nature 393, 702-5.

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

Jacobs, E. Y., Ogiwara, I., and Weiner, A M (2004). Role of the C-Terminal Domain of RNA Polymerase II in U2 snRNA Transcription and 3' Processing. Molecular and cellular biology 24, 846-855.

Jacquenet, S., Ropers, D., Bilodeau, P. S., Damier, L., Mougin, A, Stoltzfus, C M,

and Branlant, C (2001). Conserved stem-loop structures in the HIV-1 RNA region containing the A3 3' splice site and its cis-regulatory element: possible involvement in RNA splicing. Nucleic acids research 29, 464-78.

Jamison, S. F., Pasman, Z., Wang, J, Will, Cindy L, Luhrmann, R., Manley, J L, and Garcia-Blanco, M. A. (1995). U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. Nucleic Acids Res 23, 3260-3267.

Jamison, S. F., and Garcia-Blanco, M. A (1992). An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. Proceedings of the National Academy of Sciences of the United States of America 89, 5482-6.

Jantsch, M F, and Gall, J. G. (1992). Assembly and localization of the U1-specific snRNP C protein in the amphibian oocyte. The Journal of cell biology 119, 1037-46.

Jiang, Z., Cote, J., Kwon, J. M., and Goate, A. M. (2000). Aberrant Splicing of tau Pre-mRNA Caused by Intronic Mutations Associated with the Inherited Dementia Frontotemporal Dementia with Parkinsonism Linked to Chromosome 17. Molecular and cellular biology 20, 4036-4048.

Kadener, S., Fededa, J. P., Rosbash, M, and Kornblihtt, A. R. (2002). Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. Proceedings of the National Academy of Sciences of the United States of America 99, 8185-90.

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

Kaida, D. et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nature chemical biology 3, 576-83.

Kalbfuss, B., Mabon, S. A, and Misteli, T (2001). Correction of alternative splicing of tau in frontotemporal dementia and parkinsonism linked to chromosome 17. The Journal of biological chemistry 276, 42986-93.

Kambach, C., and Mattaj, I. W. (1994). Nuclear transport of the U2 snRNP-specific U2B'' protein is mediated by both direct and indirect signalling mechanisms. Journal of cell science 107 (Pt 7, 1807-16.

Kamma, H., Portman, D. S., and Dreyfuss, G (1995). Cell type-specific expression of hnRNP proteins. Experimental cell research 221, 187-96.

Kandels-Lewis, S., and Séraphin, B (1993). Role of U6 snRNA in 5 ' Splice Site Selection. Science 262, 2035-2039.

Kar, A., Fushimi, K., Zhou, X., Ray, P., Shi, C., Chen, X., Liu, Z., Chen, S., and Wu, Jane Y (2011). RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site. Molecular and cellular biology 31, 1812-21.

Kashima, T., Rao, N., David, C J, and Manley, J L (2007). hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. Human molecular genetics 16, 3149-59.

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

Kerem, E., and Kerem, B (1996). Genotype-phenotype correlations in cystic fibrosis. Pediatric pulmonology 22, 387-95.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nature Reviews Genetics 11, 345-55.

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

protein inhibitors. PloS one 4, e4533.

Ketterling, R. P., Drost, J. B., Scaringe, W. A, Liao, D.-zhou, Liu, J, Kasper, C. K., and Sommer, S S (1999). Reported in vivo splice-site mutations in the factor IX gene: severity of splicing defects and a hypothesis for predicting deleterious splice donor mutations. Human mutation 13, 221-31.

Kim, U., Wang, Y., Sanford, T., Zeng, Y., and Nishikura, K. (1994). Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. Proceedings of the National Academy of Sciences of the United States of America 91, 11457-61.

Kinnunen, S., Bonache, S., Casals, T., Monto, S., Savilahti, E., Kere, J., and Järvelä, I. (2005). Spectrum of mutations in CFTR in Finland: 18 years follow-up study and identification of two novel mutations. Journal of cystic fibrosis 4, 233-7.

Kiss, T (2004). Biogenesis of small nuclear RNPs. Journal of cell science 117, 5949-51.

Koch, C., and Høiby, N. (1993). Pathogenesis of cystic fibrosis. Lancet 341, 1065-9.

Koeberl, D. D., Bottema, C. D. K., Ketterling, R. P., Bridge, P. J., Lillicrap, D. P., and Sommer, S S (1990). Mutations causing hemophilia B: direct estimate of the underlying rates of spontaneous germ-line transitions, transversions, and deletions in a human gene. American journal of human genetics 47, 202-17.

Konarska, M M, Grabowski, P J, Padgett, R. A., and Sharp, P A (1985). Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature 313, 552-7.

Konforti, B. B., and Konarska, M M (1994). U4/U5/U6 snRNP recognizes the 5' splice site in the absence of U2 snRNP. Genes & Development 8, 1962-

1973.

Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nature chemical biology 3, 570-5.

Krawczak, M., Reiss, J., and Cooper, D N (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Human genetics 90, 41-54.

Krawczak, M., Thomas, N. S. T., Hundrieser, B., Mort, M., Wittig, M., Hampe, J., and Cooper, D N (2007). Single Base-Pair Substitutions in Exon – Intron Junctions of Human Genes : Nature , Distribution , and Consequences for mRNA Splicing. Human Mutation 28, 150-158.

Kuhlman, T. C., Cho, H., Reinberg, D., and Hernandez, N. (1999). The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. Molecular and cellular biology 19, 2130-41.

Kuhn, A. N., Santen, M. a van, Schwienhorst, A., Urlaub, H., and Lührmann, R(2009). Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. RNA (New York, N.Y.) 15, 153-75.

Kuo, H, Nasim, F H, and Grasowski, P. J. (1991). Splicing by the Control of Alternative Binding of U1 Small Nuclear Differential Particle Ribonucleoprotein. Science 251, 1045-1050.

Köhler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nature reviews. Molecular cell biology 8, 761-73.

Lamichhane, R., Daubner, G. M., Thomas-Crusells, J., Auweter, S. D., Manatschal, C., Austin, K. S., Valniuk, O., Allain, F. H.-T., and Rueda, D.

(2010). RNA looping by PTB: Evidence using FRET and NMR spectroscopy for a role in splicing repression. Proceedings of the National Academy of Sciences of the United States of America 107, 4105-10.

Lamond, A I (1993). A glimpse into the spliceosome. Current biology : CB 3, 62-4.

Lamond, A I, Konarska, M M, and Sharp, P A (1987). A mutational analysis of spliceosome assembly : evidence for splice site collaboration during spliceosome formation. Genes & Development 1, 532-543.

Lander, E. S. et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Lavigueur, A, Branche, H., Kornblihtt, a R., and Chabot, B (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. Genes & Development 7, 2405-2417.

Lesser, C. F., and Guthrie, C (1993). Mutations in U6 snRNA That Mutations Splice Site Specificity : Implications for the Active site. Science 262, 1982-1988.

Lewis, L. G., and Cohen, M. B. (1995). A selective advantage for cystic fibrosis carriers. Journal of pediatric gastroenterology and nutrition 21, 117-8.

Li, J, Hawkins, I. C., Harvey, C. D., Jennings, J. L., Link, A. J., and Patton, J G (2003). Regulation of alternative splicing by SRrp86 and its interacting proteins. Molecular and cellular biology 23, 7437-47.

Liao, X., Kretzner, L, Seraphin, B., and Rosbash, M (1990). Universally conserved and yeast-specific U1 snRNA sequences are important but not essential for U1 snRNP function. Genes & Development 4, 1766-1774.

Lin, C. H., and Patton, J G (1995). Regulation of alternative 3' splice site selection by constitutive splicing factors. RNA (New York, N.Y.) 1, 234-45.

Liu, H. X., Cartegni, L, Zhang, M Q, and Krainer, A R (2001). A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nature genetics 27, 55-8.

Liu, H. X., Chew, S L, Cartegni, L, Zhang, M Q, and Krainer, A R (2000). Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. Molecular and cellular biology 20, 1063-71.

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

Liu, Y., Jiang, Y., Si, Y., Kim, J.-Y., Chen, Z.-F., and Rao, Y. (2011). Molecular regulation of sexual preference revealed by genetic studies of 5-HT in the brains of male mice. Nature 472, 95-99.

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

Lorson, C. L., Hahnen, E., Androphy, E J, and Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proceedings of the National Academy of Sciences of the United States of America 96, 6307-11.

Loudianos, G., Lovicu, M., Dessi, V., Tzetis, M, Kanavakis, E, Zancan, L., Zelante, L, Galvèz-Galvèz, C., and Cao, A (2002). Abnormal mRNA splicing resulting from consensus sequence splicing mutations of ATP7B. Human mutation 20, 260-6.

Lu, Z.-X., Jiang, P., Cai, J. J., and Xing, Y. (2011). Context-dependent robustness to 5' splice site polymorphisms in human populations. Human molecular genetics 20, 1084-1096.

Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., and

Misteli, T (2010). Regulation of alternative splicing by histone modifications. Science (New York, N.Y.) 327, 996-1000.

Luco, R. F., and Misteli, T (2011). More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. Current opinion in genetics & development.

Lund, E., and Dahlberg, J. E. (1984). True genes for human U1 small nuclear RNA. Copy number, polymorphism, and methylation. The Journal of biological chemistry 259, 2013-21.

Madocsai, C., Lim, S. R., Geib, T., Lam, B. J., and Hertel, K J (2005). Correction of SMN2 Pre-mRNA splicing by antisense U7 small nuclear RNAs. Molecular Therapy 12, 1013-22.

Madsen, P. P. et al. (2006). Short/branched-chain acyl-CoA dehydrogenase deficiency due to an IVS3+3A>G mutation that causes exon skipping. Human genetics 118, 680-90.

Mandal, D., Feng, Z., and Stoltzfus, C M (2010). Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. Journal of virology 84, 12790-800.

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

Marcucci, R., Baralle, F E, and Romano, M (2007). Complex splicing control of the human Thrombopoietin gene by intronic G runs. Nucleic acids research 35, 132-42.

Mardon, H. J., Sebastio, G, and Baralle, F E (1987). A role for exon sequences in alternative splicing of the human fibronectin gene. Nucleic acids research 15, 7725-33.

Markham, N. R., and Zuker, M. (2008). UNAFold: software for nucleic acid folding and hybridization. Methods in molecular biology (Clifton, N.J.)

453, 3-31.

Markovtsov, V, Nikolic, J. M., Goldman, J. A., Turck, C. W., Chou, M. Y., and Black, D L (2000). Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. Molecular and cellular biology 20, 7463-79.

Martinez-Contreras, R., Fisette, J.-F., Nasim, Faiz-ul Hassan, Madden, R., Cordeau, M., and Chabot, B (2006). Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. PLoS biology 4, e21.

Matlin, A. J., Southby, J., Gooding, C., and Smith, C. W. J. (2007). Repression of alpha-actinin SM exon splicing by assisted binding of PTB to the polypyrimidine tract. RNA (New York, N.Y.) 13, 1214-23.

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

Mayeda, A, Screaton, G. R., Chandler, S. D., Fu, X. D., and Krainer, A R (1999). Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. Molecular and cellular biology 19, 1853-63.

McAlinden, A., Havlioglu, N., Liang, L., Davies, S. R., and Sandell, L. J. (2005). Alternative splicing of type II procollagen exon 2 is regulated by the combination of a weak 5' splice site and an adjacent intronic stem-loop cis element. The Journal of biological chemistry 280, 32700-11.

McCullough, A. J., and Berget, S M (1997). G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. Molecular and cellular biology 17, 4562-71.

Medlin, J. E., Uguen, P., Taylor, A., Bentley, D. L., and Murphy, S. (2003). The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3'

processing of U2 snRNA. The EMBO journal 22, 925-34.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Higuchi, M., and Seeburg, P. H. (1996). RED2, a brain-specific member of the RNA-specific adenosine deaminase family. The Journal of biological chemistry 271, 31795-8.

Merdzhanova, G., Gout, S., Keramidas, M., Edmond, V., Coll, J.-L., Brambilla, C., Brambilla, E., Gazzeri, S., and Eymin, B. (2010). The transcription factor E2F1 and the SR protein SC35 control the ratio of pro-angiogenic versus antiangiogenic isoforms of vascular endothelial growth factor-A to inhibit neovascularization in vivo. Oncogene 29, 5392-403.

Mermoud, J. E., Cohen, P. T., and Lamond, A I (1994). Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. The EMBO journal 13, 5679-88.

Meyer, I. M., and Miklos, I. (2005). Statistical evidence for conserved, local secondary structure in the coding regions of eukaryotic mRNAs and pre-mRNAs. Nucleic Acid Research 33.

Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S. Z., Keller, Walter, Antoniou, M., and Vagner, S. (2006). An interaction between U2AF 65 and CF I(m) links the splicing and 3' end processing machineries. The EMBO journal 25, 4854-64.

Minovitsky, S., Gee, S. L., Schokrpur, S., Dubchak, I., and Conboy, J. G. (2005). The splicing regulatory element , UGCAUG , is phylogenetically and spatially conserved in introns that flank tissue-specific alternative exons. Nucleic Acid Research 33, 714-724.

Misteli, T, and Spector, D. L. (1997). Protein phosphorylation and the nuclear organization of pre-mRNA splicing. Trends in cell biology 7, 135-8.

Mittal, V., Cleary, M. A., Herr, W., and Hernandez, N. (1996). The Oct-1 POU-specific domain can stimulate small nuclear RNA gene transcription by

stabilizing the basal transcription complex SNAPc. Molecular and cellular biology 16, 1955-65.

Monani, U. R., Lorson, C. L., Parsons, D. W., Prior, T. W., Androphy, E J, Burghes, A. H., and McPherson, J. D. (1999). A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. Human molecular genetics 8, 1177-83.

Montejo, J. M., Magallón, M., Tizzano, E., and Solera, J. (1999). Identification of twenty-one new mutations in the factor IX gene by SSCP analysis. Human mutation 13, 160-5.

Moore, M. J., and Sharp, P A (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature 365, 364-8.

Mouaikel, J., Bertrand, E., Tazi, J, and Fischer, U (2002). Hypermethylation of the Cap Structure of Both Yeast snRNAs and snoRNAs Requires a Conserved Methyltransferase that Is Localized to the Nucleolus. Molecular cell 9, 891-901.

Mount, S M, Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J A (1983). The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell 33, 509-18.

Mount, S M (1982). A catalogue of splice site junction sequences. Nucleic Acids Research 10, 459-472.

Mullen, M. P., Smith, C. W., Patton, J G, and Nadal-Ginard, B. (1991). Alpha-tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines default exon choice. Genes & development 5, 642-55.

Muraki, M. et al. (2004). Manipulation of alternative splicing by a newly developed inhibitor of Clks. The Journal of biological chemistry 279,

24246-54.

Mérelle, M. E., Scheffer, H., Jong, D. De, and Dankert-Roelse, J. E. (2006). Extended gene analysis can increase specificity of neonatal screening for cystic fibrosis. Acta paediatrica (Oslo, Norway : 1992) 95, 1424-8.

Müller, B., and Schümperli, D (1997). The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism. Seminars in cell & developmental biology 8, 567-76.

Nelissen, R. L., Will, C L, Venrooij, W J van, and Lührmann, R (1994). The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. The EMBO journal 13, 4113-25.

Nilsen, Timothy W, and Graveley, Brenton R (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457-63.

Oberstrass, F. C. et al. (2005). Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science (New York, N.Y.) 309, 2054-7.

Oeffner, F. et al. (2011). Intronic mutations affecting splicing of MBTPS2 cause ichthyosis follicularis, alopecia and photophobia (IFAP) syndrome. Experimental dermatology 20, 447-9.

Oers, C. C. van, Adema, G. J., Zandberg, H., Moen, T. C., and Baas, P. D. (1994). Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of the human calcitonin/calcitonin gene-related peptide I pre-mRNA. Molecular and cellular biology 14, 951-60.

Ohno, K, Brengman, J. M., Felice, K. J., Cornblath, D. R., and Engel, a G. (1999). Congenital end-plate acetylcholinesterase deficiency caused by a nonsense mutation and an A-->G splice-donor-site mutation at position +3 of the collagenlike-tail-subunit gene (COLQ): how does G at position

+3 result in aberrant splicing? American journal of human genetics 65, 635-44.

Ohno, K et al. (2001). Choline acetyltransferase mutations cause myasthenic syndrome associated with episodic apnea in humans. Proceedings of the National Academy of Sciences of the United States of America 98, 2017-22.

Okunola, H. L., and Krainer, A R (2009). Cooperative-binding and splicing-repressive properties of hnRNP A1. Molecular and cellular biology 29, 5620-31.

Pagani, F., Buratti, E, Stuani, C, Romano, M, Zuccato, E, Niksic, M., Giglio, L., Faraguna, D., and Baralle, F E (2000). Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element. The Journal of biological chemistry 275, 21041-7.

Pagani, F., Buratti, E, Stuani, Cristiana, Bendix, R., Dörk, T., and Baralle, F E (2002). A new type of mutation causes a splicing defect in ATM. Nature genetics 30, 426-9.

Pagani, F., Raponi, M., and Baralle, F E (2005). Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. Proc. Natl Acad. Sci. USA 102, 6368-6372.

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

Pagani, F., Stuani, C, Zuccato, El, Kornblihtt, A. R., and Baralle, F E (2003). Promoter architecture modulates CFTR exon 9 skipping. The Journal of biological chemistry 278, 1511-7.

Pagani, F., and Baralle, F E (2004). Genomic variants in exons and introns:

identifying the splicing spoilers. Nature Reviews Genetics 5, 389-396.

Pajares, M. J., Ezponda, T., Catena, R., Calvo, A., Pio, R., and Montuenga, L. M. (2007). Alternative splicing: an emerging topic in molecular and clinical oncology. The lancet oncology 8, 349-57.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics 40, 1413-5.

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

Pastor, T., Talotti, G., Lewandowska, M. A., and Pagani, F. (2009). An Alu-derived intronic splicing enhancer facilitates intronic processing and modulates aberrant splicing in ATM. Nucleic acids research 37, 7258-67.

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

Patterson, J. B., and Samuel, C. E. (1995). Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. Molecular and cellular biology 15, 5376-88.

Paule, M. R., and White, R. J. (2000). Transcription by RNA polymerases I and III. Nucleic Acids Research 28, 1283-1298.

Petkovic, V. et al. (2007). Exon splice enhancer mutation (GH-E32A) causes autosomal dominant growth hormone deficiency. The Journal of clinical endocrinology and metabolism 92, 4427-35.

Pinotti, M., Balestra, D., Rizzotto, L., Maestri, I., Pagani, F., and Bernardi, F. (2009). Rescue of coagulation factor VII function by the U1+5A snRNA.

Blood 113, 6461-6464.

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

Piñol-Roma, S., and Dreyfuss, G (1993). hnRNP proteins: localization and transport between the nucleus and the cytoplasm. Trends in cell biology 3, 151-5.

Pleiss, J. a, Whitworth, G. B., Bergkessel, M., and Guthrie, C (2007). Rapid, transcript-specific changes in splicing in response to environmental stress. Molecular cell 27, 928-37.

Pohlenz, J., Dumitrescu, A., Aumann, U., Koch, G., Melchior, R., Prawitt, D., and Refetoff, S. (2002). Congenital Secondary Hypothyroidism Caused by Exon Skipping due to a Homozygous Donor Splice Site Mutation in the TSH beta-Subunit Gene. Journal of Clinical Endocrinology & Metabolism 87, 336 -339.

Pomares, E., Riera, M., Castro Navarrio, J., Andrés Gutierrez, Á., Gonzalez Duarte, R., and Marfany, G. (2009). Identification of an Intronic Single-Point Mutation in RP2 as the Cause of Semidominant X-linked Retinitis Pigmentosa. Investigative Ophthalmology & Visual Science 50, 5107-5114.

Pomeranz Krummel, D. A, Oubridge, C., Leung, A. K. W., Li, Jade, and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature 458, 475-80.

Proudfoot, N. J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends in biochemical sciences 14, 105-10.

Raker, V A, Hartmuth, K., Kastner, B, and Lührmann, R (1999). Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA

nonanucleotide in a specific and thermodynamically stable manner. Molecular and cellular biology 19, 6554-65.

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

Raponi, M., Baralle, F E, and Pagani, F. (2007). Reduced splicing efficiency induced by synonymous substitutions may generate a substrate for natural selection of new splicing isoforms: the case of CFTR exon 12. Nucleic acids research 35, 606-13.

Raponi, M., Buratti, E, Dassie, E., Upadhyaya, M., and Baralle, D. (2009). Low U1 snRNP dependence at the NF1 exon 29 donor splice site. The FEBS journal 276, 2060-73.

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

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

Reese, M. G., Eeckman, F. H., Kulp, D., and Haussler, D. (1997). Improved splice site detection in Genie. Journal of computational biology : a journal of computational molecular cell biology 4, 311-23.

Reyes, J. L., Kois, P., Konforti, B. B., and Konarska, M M (1996). The canonical GU dinucleotide at the 5' splice site is recognized by p220 of the U5 snRNP within the spliceosome. RNA (New York, N.Y.) 2, 213-25.

Rinke, J., Appel, B., Blöcker, H., Frank, R., and Lührmann, R (1984). The 5'-terminal sequence of U1 RNA complementary to the consensus 5' splice site of hnRNA is single-stranded in intact U1 snRNP particles. Nucleic acids research 12, 4111-26.

Riordan, J. R. et al. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science (New York, N.Y.) 245, 1066-73.

Robberson, B L, Cote, G J, and Berget, S M (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. Molecular and cellular biology 10, 84-94.

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

Roca, X., and Krainer, A R (2009). Recognition of atypical 5' splice sites by shifted base-pairing to u1 snRNA. Nature structural & molecular biology 16, 176-182.

Romano, M, Buratti, E, Chivers, M., Kra, J., Baralle, M., Krainer, A R, and Vorechovský, I. (2007). Aberrant 5' splice sites in human disease genes : mutation pattern , nucleotide structure and comparison of computational tools that predict their utilization. Nucleic acids research 35, 4250-4263.

Romano, M, Marcucci, R., Buratti, E, Ayala, Y. M., Sebastio, G, and Baralle, F E (2002). Regulation of 3' splice site selection in the 844ins68 polymorphism of the cystathionine Beta -synthase gene. The Journal of biological chemistry 277, 43821-9.

Rooke, N., Markovtsov, V, Cagavi, E., and Black, D L (2003). Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1. Molecular and cellular biology 23, 1874-84.

Roscigno, R. F., Weiner, M., and Garcia-Blanco, M. A (1993). A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. The Journal of biological chemistry 268, 11222-9.

Roscigno, R. F., and Garcia-Blanco, M. A. (1995). SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA (New York, N.Y.) 1, 692-706.

Rossi, F, Forné, T., Antoine, E., Tazi, J, Brunel, C, and Cathala, G (1996). Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5' splice site-U1 snRNP interaction. The Journal of biological chemistry 271, 23985-91.

Ruskin, B, Zamore, P. D., and Green, M R (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell 52, 207-19.

Ruskin, B, Krainer, A R, Maniatis, T, and Green, M R (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38, 317-31.

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

Sadowski, C. L., Henry, R W, Lobo, S. M., and Hernandez, N. (1993). Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. Genes & development 7, 1535-48.

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

Saltzman, A. L., Pan, Q., and Blencowe, B. J. (2011). Regulation of alternative splicing by the core spliceosomal machinery. Genes & development 25, 373-84.

Sanford, J R, Ellis, J., and Cáceres, J F (2005). Multiple roles of arginine/serine-rich splicing factors in RNA processing. Biochemical Society transactions 33, 443-6.

Sanford, Jeremy R, Ellis, J. D., Cazalla, D., and Cáceres, Javier F (2005). Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. Proceedings of the National Academy of Sciences of the United States of America 102, 15042-7.

Schaal, T. D., and Maniatis, T (1999). Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. Molecular and cellular biology 19, 1705-19.

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

Schmid, F., Glaus, E., Barthelmes, D., Fliegauf, M., Gaspar, H., Nurnberg, G., et al. (2011). U1 snRNA-Mediated Gene Therapeutic Correction of Splice Defects Caused by an Exceptionally Mild BBS Mutation Human Mutation. Human Mutation 32, 815-821.

Schmidt, A. E., and Bajaj, S. P. (2003). Structure – Function Relationships in Factor IX and Factor IXa. Trends in Cardiovascular Medicine 13, 39-45.

Schoft, V. K., Schopoff, S., and Jantsch, M F (2007). Regulation of glutamate receptor B pre-mRNA splicing by RNA editing. Nucleic acids research 35, 3723-32.

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

Sharma, S., Kohlstaedt, L. A., Damianov, A., Rio, D. C., and Black, D L (2008).

Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. Nat. Struct. Mol. Biol. 15, 183-191.

Sharp, P A (1994). Split genes and RNA splicing. Cell 77, 805-15.

Shatkin, A. J. (1976). Capping of eucaryotic mRNAs. Cell 9, 645-53.

Shepard, P. J., Choi, E.-A, Busch, A., and Hertel, K J (2011). Efficient internal exon recognition depends on near equal contributions from the 3' and 5' splice sites. Nucleic acids research, 1-10.

Shepard, P. J., and Hertel, K J (2009). The SR protein family. Genome Biology 10, 242.

Sheth, N., Roca, X., Hastings, M. L., Roeder, T., Krainer, A R, and Sachidanandam, R. (2006). Comprehensive splice-site analysis using comparative genomics. Nucleic acids research 34, 3955-67.

Shi, Y., and Manley, J L (2007). A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock. Molecular cell 28, 79-90.

Siliciano, P. G., and Guthrie, C (1988). 5' splice site selection in yeast: genetic alterations in base-pairing with Ul reveal additional requirements. Genes & Development 2, 1258-1267.

Singh, N K, Singh, N N, Androphy, E J, and Singh, R N (2006). Splicing of a critical exon of human survival motor neuron is regulated by a unique silencer element located in the last intron. Mol. Cell. Biol. 26, 1333-1346.

Singh, Natalia N, Androphy, Elliot J, and Singh, Ravindra N (2004a). An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy. Biochemical and biophysical research communications 315, 381-8.

Singh, Natalia N, Androphy, Elliot J, and Singh, Ravindra N (2004b). In vivo

selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes. RNA (New York, N.Y.) 10, 1291-305.

Sironi, M., Menozzi, G., Riva, L., Cagliani, R., Comi, G. P., Bresolin, N., Giorda, R., and Pozzoli, U. (2004). Silencer elements as possible inhibitors of pseudoexon splicing. Nucleic acids research 32, 1783-91.

Skoko, N., Baralle, M., Buratti, E, and Baralle, F E (2008). The pathological splicing mutation c.6792C>G in NF1 exon 37 causes a change of tenancy between antagonistic splicing factors. FEBS letters 582, 2231-6.

Slomski, R., Schloesser, M., Berg, L. P., Wagner, M., Kakkar, V. V., Cooper, D N, and Reiss, J. (1992). Omission of exon 12 in cystic fibrosis transmembrane conductance regulator (CFTR) gene transcripts. Human genetics 89, 615-9.

Smith, C. W., and Valcárcel, J (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. Trends in biochemical sciences 25, 381-8.

Soret, J. et al. (2005). Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. Proceedings of the National Academy of Sciences of the United States of America 102, 8764-9.

Spillantini, M G, Murrell, J. R., Goedert, M, Farlow, M. R., Klug, a, and Ghetti, B. (1998). Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. Proceedings of the National Academy of Sciences of the United States of America 95, 7737-41.

Spurdle, A. B., Couch, F. J., Hogervorst, F. B. L., Radice, P., and Sinilnikova, O. M. (2008). Prediction and assessment of splicing alterations: implications for clinical testing. Human mutation 29, 1304-13.

Staffa, A., Acheson, N. H., and Cochrane, A. (1997). Novel exonic elements that modulate splicing of the human fibronectin EDA exon. The Journal of

biological chemistry 272, 33394-401.

Stamm, S, Zhang, M Q, Marr, T. G., and Helfman, D. M. (1994). A sequence compilation and comparison of exons that are alternatively spliced in neurons. Nucleic acids research 22, 1515-26.

Stamm, S (2008). Regulation of alternative splicing by reversible protein phosphorylation. The Journal of biological chemistry 283, 1223-7.

Stanek, D., and Neugebauer, K. M. (2006). The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. Chromosoma 115, 343-54.

Stanziale, P., Savino, M., Bonis, P. De, Granatiero, M., Zelante, L, and Bisceglia, L. (2005). Indirect CFTR mutation identification by PCR/OLA anomalous electropherograms. Genetic testing 9, 285-91.

Stark, H., Dube, P., Lührmann, Reinhard, and Kastner, Berthold (2001). Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. Nature 409, 539-42.

Stefanovic, B., Hackl, W., Lührmann, R, and Schümperli, D (1995). Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into Xenopus oocytes. Nucleic acids research 23, 3141-51.

Sun, H., and Chasin, L A (2000). Multiple splicing defects in an intronic false exon. Molecular and cellular biology 20, 6414-25.

Sun, Q., Mayeda, A, Hampson, R. K., Krainer, A R, and Rottman, F. M. (1993). General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes & development 7, 2598-608.

Surowy, C. S., Santen, V. L. van, Scheib-Wixted, S. M., and Spritz, R. A (1989). Direct, sequence-specific binding of the human U1-70K ribonucleoprotein antigen protein to loop I of U1 small nuclear RNA. Molecular and cellular biology 9, 4179-86.

Susani, L. et al. (2004). TCIRG1-dependent recessive osteopetrosis: mutation

analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA. Human mutation 24, 225-35.

Sébillon, P., Beldjord, C., Kaplan, J. C., Brody, E., and Marie, J. (1995). A T to G mutation in the polypyrimidine tract of the second intron of the human beta-globin gene reduces in vitro splicing efficiency: evidence for an increased hnRNP C interaction. Nucleic acids research 23, 3419-25.

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

Tanaka, K., Watakabe, A, and Shimura, Y. (1994). Polypurine sequences within a downstream exon function as a splicing enhancer. Molecular and cellular biology 14, 1347-54.

Tanner, G., Glaus, E., Barthelmes, D., Ader, M., Fleischhauer, J., Pagani, F., Berger, W., and Neidhardt, J. (2009). Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. Human mutation 30, 255-63.

Tardiff, D. F., and Rosbash, M (2006). Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. RNA (New York, N.Y.) 12, 968-79.

Tarn, W. Y., and Steitz, J. A (1994). SR proteins can compensate for the loss of U1 snRNP functions in vitro. Genes & Development 8, 2704-2717.

Tavassoli, K., Eigel, A., Wilke, K., Pollmann, H., and Horst, J. (1998). Molecular diagnostics of 15 hemophilia A patients: characterization of eight novel mutations in the factor VIII gene, two of which result in exon skipping. Human mutation 12, 301-3.

Tazi, Jamal, Bakkour, N., and Stamm, S (2009). Alternative splicing and disease. Biochimica et biophysica acta 1792, 14-26.

Tazi, J, Kornstadt, U., Rossi, Ferdinand, Jeanteur, P., Cathala, Guy, Brunel, C, and Luhrmann, R. (1993). thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. Nature 363, 283-286.

Thanaraj, T. A (2000). Positional characterisation of false positives from computational prediction of human splice sites. Nucleic acids research 28, 744-54.

Théry, J. C. et al. (2011). Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes. European journal of human genetics : EJHG.

Tian, M., and Maniatis, T (1994). A splicing enhancer exhibits both constitutive and regulated activities. Genes & Development 8, 1703-1712.

Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J, and Guigó, R. (2009). Nucleosome positioning as a determinant of exon recognition. Nature structural & molecular biology 16, 996-1001.

Tournier, I. et al. (2008). A large fraction of unclassified variants of the mismatch repair genes MLHI and MSH2 is associated with splicing defects. Human Mutation 29, 1412-1424.

Tripathi, V. et al. (2010). The Nuclear-Retained Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR Splicing Factor Phosphorylation. Molecular Cell 39, 925-938.

Tuffery-Giraud, S. et al. (2009). Genotype-phenotype analysis in 2,405 patients with a dystrophinopathy using the UMD-DMD database: a model of nationwide knowledgebase. Human mutation 30, 934-45.

Turner, I. A., Norman, C. M., Churcher, M. J., and Newman, A. J. (2004). Roles of the U5 snRNP in spliceosome dynamics and catalysis. Biochemical Society transactions 32, 928-31.

Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J A (1998). Modification

of U6 spliceosomal RNA is guided by other small RNAs. Molecular cell 2, 629-38.

Tzetis, M, Efthymiadou, a, Doudounakis, S, and Kanavakis, E (2001). Qualitative and quantitative analysis of mRNA associated with four putative splicing mutations (621+3A-->G, 2751+2T-->A, 296+1G-->C, 1717-9T-->C-D565G) and one nonsense mutation (E822X) in the CFTR gene. Human genetics 109, 592-601.

Uchikawa, H., Fujii, K., Kohno, Y., Katsumata, N., Nagao, K., Yamada, M., and Miyashita, T. (2007). U7 snRNA-mediated correction of aberrant splicing caused by activation of cryptic splice sites. Journal of human genetics 52, 891-7.

Uguen, P., and Murphy, S. (2003). The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. The EMBO journal 22, 4544-54.

Valadkhan, S., Mohammadi, A., Jaladat, Y., and Geisler, S. (2009). Protein-free small nuclear RNAs catalyze a two-step splicing reaction. Proceedings of the National Academy of Sciences of the United States of America 106, 11901-6.

Vankeerberghen, a, Wei, L., Jaspers, M., Cassiman, J. J., Nilius, B., and Cuppens, H. (1998). Characterization of 19 disease-associated missense mutations in the regulatory domain of the cystic fibrosis transmembrane conductance regulator. Human molecular genetics 7, 1761-9.

Varani, L, Hasegawa, M, Spillantini, M G, Smith, M. J., Murrell, J. R., Ghetti, B., Klug, a, Goedert, M, and Varani, G (1999). Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. Proceedings of the National Academy of Sciences of the United States of America 96, 8229-34.

Vibe-pedersen, K., Kornblihtt, A. R., and Baralle, F E (1984). Expression of a human a-globin/fibronectin gene hybrid generates two mRNAs by alternative splicing. The EMBO journal 3, 2511 -2516.

Vidal, C., Cachia, A., and Xuereb-Anastasi, A. (2009). Effects of a synonymous variant in exon 9 of the CD44 gene on pre-mRNA splicing in a family with osteoporosis. Bone 45, 736-42.

Wahl, M. C., Will, Cindy L, and Lührmann, R (2009). The spliceosome: design principles of a dynamic RNP machine. Cell 136, 701-18.

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

Wang, G.-shin, and Cooper, T A (2007). Splicing in disease : disruption of the splicing code and the decoding machinery. Nature Reviews Genetics 8, 749-762.

Wang, H. Y., Lin, W., Dyck, J. A., Yeakley, J. M., Songyang, Z., Cantley, L. C., and Fu, X. D. (1998). SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. The Journal of cell biology 140, 737-50.

Wang, K., Zhao, X., Chan, S., Cil, O., He, N., Song, X., Paterson, A. D., and Pei, Y. (2009). Evidence for pathogenicity of atypical splice mutations in autosomal dominant polycystic kidney disease. Clinical journal of the American Society of Nephrology : CJASN 4, 442-9.

Wang, Z., Rolish, M. E., Yeo, G., Tung, V., Mawson, M., and Burge, C B (2004). Systematic identification and analysis of exonic splicing silencers. Cell 119, 831-45.

Ward, A. J., and Cooper, T A (2010). The pathobiology of splicing. Journal of

Pathology 220, 152-163.

Wessagowit, V., Nalla, V. K., Rogan, P. K., and McGrath, J. a (2005). Normal and abnormal mechanisms of gene splicing and relevance to inherited skin diseases. Journal of dermatological science 40, 73-84.

Weston, M. D. et al. (1996). Myosin VIIA Mutation Screening Type 1 Patients 189 Usher Syndrome. American journal of human genetics 59, 1074-1083.

Whiley, P. J., Pettigrew, C. A., Brewster, B. L., Walker, L. C., Spurdle, A. B., and Brown, M. A. (2010). Effect of BRCA2 sequence variants predicted to disrupt exonic splice enhancers on BRCA2 transcripts. BMC medical genetics 11, 80.

Wilkinson, F. H., Ahmad, S. S., and Walsh, P. N. (2002). The factor IXa second epidermal growth factor (EGF2) domain mediates platelet binding and assembly of the factor X activating complex. The Journal of biological chemistry 277, 5734-41.

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

Will, C L, Rümpler, S., Klein Gunnewiek, J. M., Venrooij, W J van, and Lührmann, R (1996). In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. Nucleic acids research 24, 4614-23.

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

Wu, J Y, and Maniatis, T (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell 75, 1061-70.

Wu, Q., and Krainer, A R (1998). Purine-rich enhancers function in the AT-AC pre-mRNA splicing pathway and do so independently of intact U1 snRNP. Rna 4, 1664-1673.

Wu, S., Romfo, C. M., Nilsen, T W, and Green, M R (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature 402, 832-5.

Wyatt, J. R., Sontheimer, E. J., and Steitz, J. A (1992). Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. Genes & Development 6, 2542-2553.

Xiao, S, and Manley, J L. (1997). Phosphorylation of the ASF / SF2 RS domain aifects both protein-protein and protein-RNA interactions and is necessary for splicing. Genes & Development 11, 334-344.

Xiao, X., Wang, Z., Jang, M., and Burge, C B (2007). Coevolutionary networks of splicing cis-regulatory elements. Proceedings of the National Academy of Sciences of the United States of America 104, 18583-8.

Xing, Y., and Lee, C. (2006). Alternative splicing and RNA selection pressure--evolutionary consequences for eukaryotic genomes. Nature reviews. Genetics 7, 499-509.

Yasuda, M, Takamatsu, J, et al. (2000). A novel mutation at position +12 in the intron following exon 10 of the tau gene in familial frontotemporal dementia (FTD-Kumamoto). Annals of neurology 47, 422-9.

Yasuda, M, Takamatsu, J, et al. (2000). A Novel Mutation at Position +12 in the Intron following Exon 10 of the Tau Gene in Familial Frontotemporal Dementia (FTD-Kumamoto). Annals of Neurology 47, 422-429.

Yeo, G., and Burge, C B (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. Journal of computational biology : a journal of computational molecular cell biology 11, 377-94.

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

Yuo, C Y, and Weiner, A M (1989). A U1 small nuclear ribonucleoprotein particle with altered specificity induces alternative splicing of an adenovirus E1A mRNA precursor. Molecular and cellular biology 9, 3429-37.

Zdziarska, J., Chojnowski, K., Klukowska, A., Łetowska, M., Mital, A., Podolak-Dawidziak, M., Windyga, J., and Zawilska, K. (2009). Therapeutic properties and safety of recombinant factor VIII and factor IX. Polskie Archiwum Medycyny Wewnętrznej 119, 403-9.

Zhang, Li, Vincent, G. M., et al. (2004). An intronic mutation causes long QT syndrome. Journal of the American College of Cardiology 44, 1283-91.

Zhang, X. H., and Chasin, L A (2004). Computational definition of sequence motifs governing constitutive exon splicing. Genes & Development 18, 1241-1250.

Zhang, X. H.-F., Kangsamaksin, T., Chao, M. S. P., Banerjee, J. K., and Chasin, L A (2005). Exon inclusion is dependent on predictable exonic splicing enhancers. Molecular and cellular biology 25, 7323-32.

Zhang, X. H.-F., and Chasin, L A (2004). Computational definition of sequence motifs governing constitutive exon splicing. Genes & development 18, 1241-50.

Zheng, S., Chen, Y., Donahue, C. P., Wolfe, M. S., and Varani, G(2009). Article Structural Basis for Stabilization of the Tau Pre-mRNA Splicing Regulatory Element by Novantrone (Mitoxantrone). Chemistry & Biology 16, 557-566.

Zheng, Z. M., He, P. J., and Baker, C. C. (1999). Function of a bovine

papillomavirus type 1 exonic splicing suppressor requires a suboptimal upstream 3' splice site. Journal of virology 73, 29-36.

Zheng, Z. M., Huynen, M., and Baker, C. C. (1998). A pyrimidine-rich exonic splicing suppressor binds multiple RNA splicing factors and inhibits spliceosome assembly. Proceedings of the National Academy of Sciences of the United States of America 95, 14088-93.

Zheng, Z. M., and Baker, C. C. (2000). Parameters that affect in vitro splicing of bovine papillomavirus type 1 late pre-mRNAs. Journal of virological methods 85, 203-14.

Zhou, H.-L., and Lou, H. (2008). Repression of prespliceosome complex formation at two distinct steps by Fox-1/Fox-2 proteins. Molecular and cellular biology 28, 5507-16.

Zhu, J., Mayeda, A, and Krainer, A R (2001). Exon Identity Established through Differential Antagonism between Exonic Splicing Silencer-Bound hnRNP A1 and Enhancer-Bound SR Proteins. Molecular Cell 8, 1351-1361.

Zhuang, Y., and Weiner, A M (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell 46, 827-35.

Zielenski, J, and Tsui, L. C. (1995). Cystic fibrosis: genotypic and phenotypic variations. Annual review of genetics 29, 777-807.

Zielenski, J (2000). Genotype and phenotype in cystic fibrosis. Respiration 67, 117-33.

Zielenski, J, Makiewicz, D., Lin, S.-P., Huang, F.-Y., Yang-Feng, T., and Tsui L-C (1995). Skipping of exon 12 as a consequences of a point mutation (1898+5G->T) in the cystic fibrosis transmembrane conductance regulator gene found in a consanguineous Chinese family. Clinical Genetics 47, 125-132.

Zuo, P., and Maniatis, T (1996). The splicing factor U2AF35 mediates critical

protein-protein interactions in constitutive and enhancer-dependent splicing. Genes & development 10, 1356-68.

8.Acknowledgements

This work was carried out in the Molecular Pathology / Human Molecular Genetics group at the International Centre for Genetic Engineering and Biotechnology (ICGEB) in Trieste, under the supervision of Dr. Franco Pagani.
I would like to thank my boss Franco Pagani for allowing me to work in his group, for his constant support in research work during these years and for his professionalism and patience.
I want to thank the marvelous Human Molecular Genetics people: Erica, Gabriele, Elisa, Tibor, Ify, Andrea, Gosia, Chiara, Giulia for their great friendship and the splendid atmosphere they have created in the lab.
Finally, I would like to thank our collaborators at the University of Ferrara: Dr. Mirko Pinotti and his team and Dr Daniela Pe