# Università degli Studi di Ferrara

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

### CICLO XXIV

COORDINATORE Prof. Bernardi Francesco

# Modified U1snRNAs as innovative therapeutic strategy for inherited coagulation factor deficiencies

Settore Scientifico Disciplinare BIO/11

**Dottorando**
Dott. Balestra Dario

*(firma)*

**Tutore**
Prof. Bernardi Francesco
Prof. Pinotti Mirko
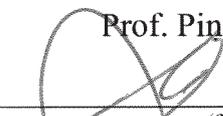
*(firma)*
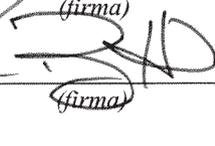
*(firma)*

Anni 2008/2011

# Abstract

A significant proportion of disease-causing mutations affects pre-mRNA splicing inducing skipping of the exon from the mature transcript. Using coagulation *F7* exon 7 and *F9* exon 5 models, we characterized a panel of mutations at either donor splice site or acceptor site. These mutations were found in patients with coagulation factor VII (FVII) or IX (FIX) deficiency with various degree of severity. By expression studies in mammalian cells, the pathological effect of the mutations has been characterized. As expected most of them, altering the sequences relevant for spliceosome assembly and thus removal of intron during the pre-mRNA maturation, altered the splicing process, causing exon skipping and/or activation of a cryptic donor splice site.

For all these mutations, we assessed the ability of engineered U1snRNA, the key component of the U1snRNP  that recognizes donor splice sites, to restore correct splicing.

In the FVII deficiency model caused by the IVS7+5g/a mutation, a modified U1 (U1+5a) was able to partially restore the splicing defects and to increase the levels of the secreted functional protein. The U1+5a was also demonstrated to rescue the expression of human FVII in the mouse model.

In the cellular model of FIX deficiency, loading of U1snRNA by complementarity to the normal or mutant donor splice site corrected exon 5 skipping caused by different types of mutations. To improve specificity and reduce off-target effects, engineered U1snRNA variants targeting non conserved intronic sequences downstream of the mutated donor splice site, called Exon Specific U1 (ExSpeU1), were evaluated. We identified a unique ExSpeU1 that was able to rescue splicing impaired by different mutations either at the polypyrimidine tract or at the donor splice site. The correction was also  demonstrated on secreted functional FIX levels.

The rescue of secreted coagulation FVII and FIX levels obtained both *in vitro* and *in vivo* could have, if translated in patients, a therapeutic impact, thus encouraging further studies aimed at proposing the modified U1snRNA, and particularly the ExSpeu1, as innovative treatment for coagulation factor disease caused by aberrant splicing. This approach could be extended to other human genetic disorders associated to aberrant splicing, a frequent cause of clinically severe disease forms.

# Contents

# 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.

| | |
|---|---|
| 3' ss | Acceptor splice site |
| 5' ss | Donor splice site |
| aa | Amino acid |
| AAV | Adeno-associated Virus |
| bp | Base pairs |
| cDNA | Complementary DNA |
| 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 |
| FVII | Coagulation factor VII protein |
| hnRNP | Heterogenous ribonuclear protein |
| IPTG | Isopropyl-β-d-thiogalactopyranoside |
| ISE | Intronic Splicing Enhancer |
| ISS | Intronic Splicing Silencer |
| IVS | iper-variable sequence |
| Kb | Kilobase |
| kDa | Kilodalton |
| NMD | Nonsense-mediated decay |
| nt | Nucleotides |
| PBS | Phosphate bufer saline |
| PCR | Polymerase Chain Reaction |
| PPT | Polypyrimidine tract |

| | |
|---|---|
| R | Purine (G or A) |
| RNA pol II | RNA polymerase II |
| RNA pol III | RNA polymerase III |
| RNA | Ribonucleic acid |
| RRM | RNA Recognition Motif |
| RS | Arginine-Serine rich motif |
| SDS | N-lauroylsarcosine sodium salt |
| snRNA | Small nuclear RNA |
| snRNP | Small nuclear ribonucleoprotein particles |
| SR | Arginine-serine rich protein |
| SRE | Splicing Regulatory Elements |
| ss | Splice site |
| UTR | untranslated region |
| wt | Wild-type |
| Y | Pyrimidine (T or C) |

# I. Introduction

When you bleed, the body launches a series of reactions that help the blood clot. This is called the coagulation cascade. The process involves special proteins called coagulation factors. When one or more of these clotting factors are missing, or defective, there is usually a higher chance of bleeding. The severity of patients' symptoms vary considerably, from epistasis to life-threatening bleeding episodes (gastro-intestinal or intra-cranial bleeding) and generally rely on the levels of the disease protein/s. In the inherited form of these deficiencies, the deficiency of these clotting factors is generally caused by mutations occurring within gene. Among all mutations found in patients, about 10% are associated to splicing, which become much more frequent in the severe forms (up to 20%). In recent years, mutations related to altered splicing are emerging as relatively common among all genes. In fact, although the frequency of splicing mutations varies considerably between individual genes, it is considered that approximately 30% of pathogenic mutations cause disease through aberrant splicing mechanisms. 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 and Baralle 2004; Buratti, Baralle et al. 2006). 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, Wei et al. 1998; Ohno, Tsujino et al. 2001; Aznarez, Chan et al. 2003).

Based on these recent data, the demand of new tools able to modulate the splicing mechanism is growing up. Actually, the only tools able to act at mRNA level are based on utilization of antisense oligonucleotides, as well as siRNA or antisense U7snRNA/U1snRNA (Denti, Rosa et al. 2006; Asparuhova, Marti et al. 2007; Geib and Hertel 2009; Fragall, Adams et al. 2011; Goyenvalle and Davies 2011). Except for siRNA, , through which it is possible modulate the degradation of the mutated mRNA, the antisense oligonucleotide-based approach (antisense U7/U1snRNAs act in the same manner) is directed towards the modulation of the splicing of the target mRNA masking crucial sequences involved in the splicing process. Nowadays no tools able to modulate the splicing process has been developed, and they could represent an alternative and innovative approach in the treatment of splicing related-diseases.

The laboratory in which I conducted my PhD thesis has an established experience in the molecular bases of coagulation factor deficiencies. We used FVII and FIX deficiencies as models to study the relationship between mutations and patients' phenotype and to develop an innovative therapeutic approach both *in vitro* and *in vivo*, thanks to a productive collaborations with Dott. Franco Pagani at ICGEB, and Dott. Valder Arruda at Children's Hospital of Philadelphia, who has a longstanding

experience in basic research in premRNA splicing and a big experience in the gene therapy of haemophilia B with AAV vectors, respectively.

## I.1 The blood-clotting cascade

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

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

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

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

of two peptide bonds, thereby releasing two small amino-terminal peptides, fibrinopeptide A and fibrinopeptide B, and giving rise to fibrin monomers that polymerise spontaneously. The insoluble fibrin network is subsequently stabilized by the transglutaminase action of FXIIIa, activated by thrombin in the presence of $Ca^{2+}$ ions (Furie and Furie 1988).

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

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

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

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

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

Vascular damage is always accompanied by inflammation and repair reactions. Thrombin plays a key role in these processes by chemotactically drawing leukocytes to the site of injury and by stimulating tissue remodelling and mitogenesis. P-selectin expressed on the platelet membrane in the haemostatic plug acts as a receptor for monocytes and neutrophils which, in addition to providing ideal membrane surfaces for blood coagulation, sustain the inflammatory response. During wound healing, the fibrin clot in the vessel is degraded by plasmin, a serine protease generated form the plasma zymogen plasminogen, in a process known as fibrinolysis. The activation of plasminogen is catalysed by several plasminogen activators, including two serine proteases, tissue plasminogen activator (t-PA) and urokinase (UK). Their activity is in turn regulated by two plasma protease inhibitors, plasminogen activator inhibitor and $\alpha_2$-antiplasmin (Furie and Furie 1988).
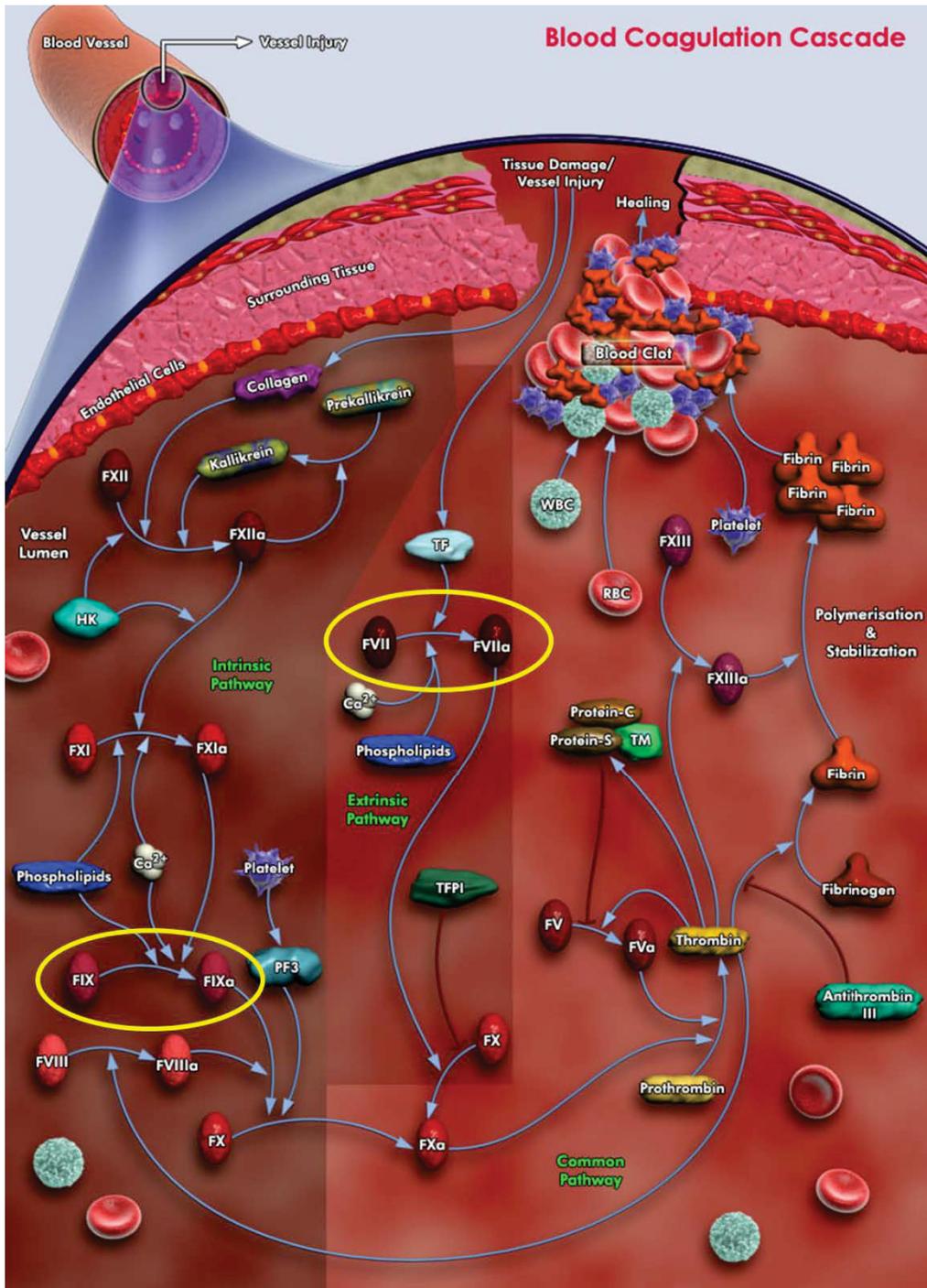
**Figure 1:Scheme showing the intrinsic and extrinsic pathways of coagulation cascade leading to clot formation.**
*A deficiency of coagulation factor VII or factor IX ( yellow ovals) compromises the coagulation cascade activation , leading to inefficient clot formation and resulting in prolonged bleeding time (FVII deficiency and haemophilia B). The image has been taken by the QIAGEN site.*

## I.2 Blood clotting proteins

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

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

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

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

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

Fibrinogen is composed of two identical monomers, each comprising three polypeptide chains: Aα, Bβ and γ; its structure is unique because each of the three subunits is encoded by a separate gene.

During fibrinogen activation, thrombin removes fibrinopeptide A from the Aα chain and fibrinopepitde B from the Bβ chain, unmasking sites for the spontaneous polymerization of the fibrin monomers.

## I.3 F7 gene

The gene encoding human Factor VII (*F7*), localized through ibridization *in situ*, is gene located in a single copy on the long arm of chromosome 13, region q34-qter, 2.8 kb downstream FX gene (Pfeiffer, Ott et al. 1982).

The complete nucleotide sequence of the human F7 was reported by O'Hara et al. in 1987, one year after the isolation and characterization of cDNA clones coding for FVII by Hagen et al. in 1986 (Hagen, Gray et al. 1986; O'Hara, Grant et al. 1987). Clones were obtained from two different cDNA libraries prepared from poly(A) RNA from human liver and human hepatoma HepG2 cells.
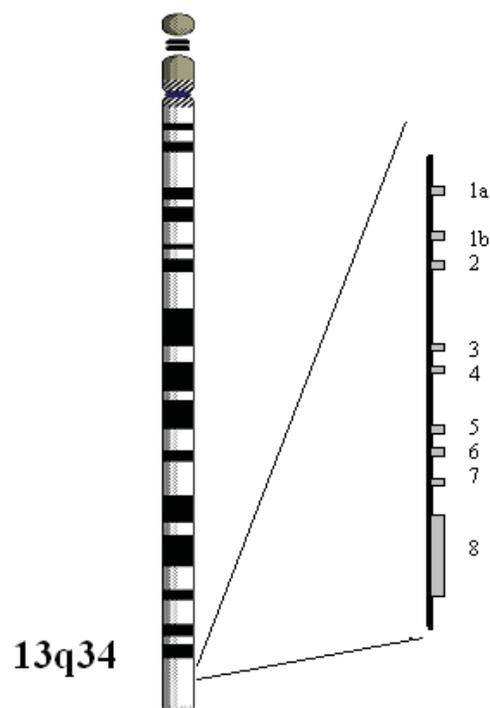


*Figure 2: Representation of the human chromosome 13 and organization of F7 gene*

The F7 gene spans about 12.8 kb of genomic DNA and contained nine exons and eight introns (see table below). The introns ranged in size from 68 nucleotides to nearly 2.6kb. The exons also vary considerably in size, ranging from 25 nucleotides to 1.6kb. Exons 1a, 1b and part of exon 2 encode a prepro leader sequence that is removed during processing. The remainder of exon 2 and exons 3-8 encode the 406 amino acids present in the mature protein circulating in blood. Two different cDNAs coding for FVII were isolated by Hagen et al.: one clone coded for a prepro leader sequence

15

of 38 amino acids and the second coded for a 60 amino acids leader sequence. The genomic sequence data revealed that the second clone contained an optional exon (exon 1b) that was absent in the first clone. cDNA clones representing the two mRNa in which exon 1b is either present or absent both give rise to functional transcripts which code for a biologically active FVII. In normal liver, however, mRNA lacking exon 1b is far more abundant than mRNA possessing this exon (Berkner, Busby et al. 1986).

The −17 to −1 region of the vitamin K-dependent proteins functions as a γ-carboxylation signal domain. It is partially conserved in FIX, FX, PC, PS and PT and is present in the 38 or 60 amino acid leader of FVII. The 38 amino acid sequence resembles the other vitamin K proteins leader in sequence, size, hydrophobicity pattern, exon-intron structure, and predicted signal peptide cleavage more closely than the 60 amino acid leader. The predicted cleavage for the 60 amino acid leader would leave 16 amino acids encoded by exon 1b attached to the putative γ-carboxylase signal, but this is evidently tolerable.

*Table 1 – F7 gene organization in relation to exon/intron size*

| EXON | INTRON | CODING REGION | DIMENSION (bp) |
|---|---|---|---|
| 1a | | pre-proleader sequence | 100 |
| | 1a | | 1068 |
| 1b | | pre-proleader sequence | 66 |
| | 1b | | 2574 |
| 2 | | pre-proleader, Gla domain | 161 |
| | 2 | | 1928 |
| 3 | | Gla domain | 25 |
| | 3 | | 70 |
| 4 | | EGF domain | 139 |
| | 4 | | 1716 |
| 5 | | EGF domain | 141 |
| | 5 | | 971 |
| 6 | | activation site | 110 |
| | 6 | | 595 |
| 7 | | catalytic domain | 124 |
| | 7 | | 817 |
| 8 | | catalytic domain | 1622 |

As noted before for other vitamin K protein genes, exons in F7 gene encode discrete domains of the protein: prepro leader, γ-carboxylase region, growth factor domains, activation region and the serine protease domain. The conservation of domains, intron position, and intron phase (the conservation of phase preserves the reading frame) among the members of the vitamin K protein family supports the theory of differentiation by exon shuffling. The least conserved region is the activation domain. Unlike FIX, FX, PC and prothrombin, FVII does not release an activation peptide and the other proteins vary considerably in size and sequence in this region.

The striking degree of similarity between the exons of genes coding for vitamin K proteins contrasted with the lack of resemblance in the sequences or sizes of the introns of these genes (O'Hara, Grant et al. 1987). For instance, introns in the human F9 and PC genes contain Alu repeats, whereas the F7 gene lacks such sequences. On the other hand, F7 gene contains five minisatellite imperfect tandem repeats (sequences that are repeated directly adjacent to each other) with monomer lengths ranging from 14 to 37 bp and copy numbers ranging from 6 to 52, while F9 and PC lack minisatellite DNA (O'Hara and Grant 1988). These tandem repeats in F7 gene are often responsible for polymorphism due to allelic variation in the repeat copy number. Tandem repeats may evolve because of random crossover in DNA whose sequence is not maintained by selection. This suggests that much of the sequence information present in the introns and the 3' untraslated portion of FVII messenger may be dispensable.

The F7 gene contains copies of sequences that are typically associated with the regulation of transcription and translation. The sequence surrounding the ATG initiation codon is TCATC<u>ATGG</u> in which 7 of 9 nucleotides match with the translation initiation site consensus sequence of Kozak CCACC<u>ATGG</u>. The sequence from nucleotide −366 to −260 exhibits limited homology with regions upstream of the putative transcription initiation sites of F9 and PC genes.

Promoter and silencer elements of the immediate 5' flanking region of the human F7 gene were identified and characterized (Erdmann and Heim 1995; Greenberg, Miao et al. 1995; Pollak, Hung et al. 1996). Differently from the promoter regions of most eukaryotic genes, the 5' flanking sequence of F7 lack the typical TATA-box; the same feature has been found in promoters of other coagulation proteins like FIX, FX, FXII and prothrombin. Although a CAAT-box is critical in the promoters of Factor IX and X, no such element is apparent in the Factor VII promoter. Sequence alignment of F7 and F10 promoters revealed similarity of 86% for a small 37-bp element, but the functional relevance, if any, of this sequence in unclear as its deletion did not alter significantly reporter gene expression. The major transcription start site has been identified within a strong initiator element (-57 CCCGTCAGTCCC -46), at position −51 upstream from the start site of translation (+1). Multiple other start sites have been recognized in the region surrounding the major

transcription star site, a typical feature of constitutively expressed genes which lack TATA and CAAT sequences.

The major start site of FVII is only 8 bases downstream from a consensus sequence for the transcription factor hepatocyte nuclear factor-4 (HNF-4), which binds at nucleotide –63 to –58 (ACTTTG). This close proximity is similar to that of the FIX major start site, which is 18 bp from the HNF-4 sequence. HNF-4 binds F7 promoter with a lower affinity as compared to F9 and F10 promoters and this may explain the large differences in steady-state mRNA levels and plasma concentrations of these coagulation proteins. HNF-4 has been found to play an important role in the transcription initiation of a number of genes expressed in the liver but it is not the only limiting factor in non-hepatic cells, so additional liver-specific factors are probably required to fully activate F7 promoter. To date, the proximal 185 base pairs upstream the ATG translation initiation codon were sufficient to confer liver-specific expression and maximal promoter activity in HepG2 cells.

The sequence from –101 to –94 (CCCCTCCC) was shown to be a binding site for the ubiquitous transcription factor Sp1 and a sequence with homology to a hormone responsive element has been detected at –227 to –21.

Functional studies of promoter deletions in the HepG2 cell line showed that deletion of sequences from -1601 down to -1212 increases expression about 2-fold, which suggests that a negative element is present upstream of the -1212 (Figure 3).
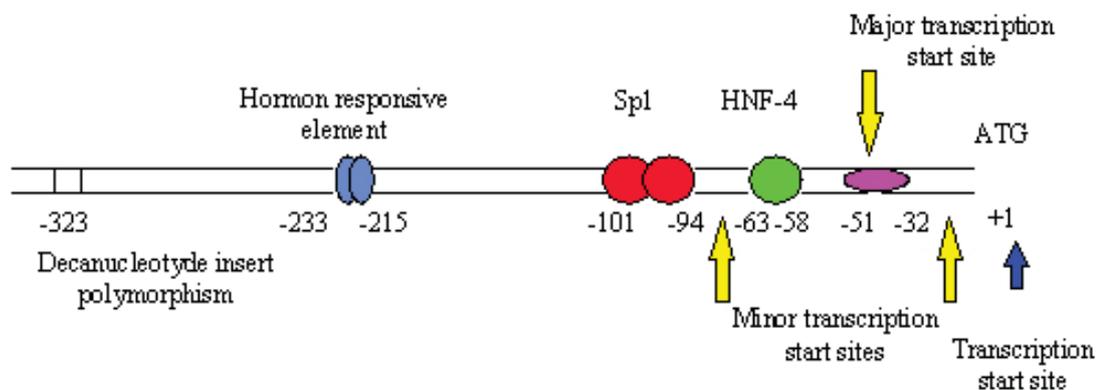


***Figure 3:****Schematic representation of regulatory sequences*
*in factor 7 gene 5'-flanking region*

A comparison of different cDNA clones for Factor VII showed that alternative sites for polyadenylylation occurring downstream from the poly(A) signal of AATAAA were present in the 3'UTR. These multiple copies in F7 gene may direct polyadenylylation at more than one site (O'Hara, Grant et al. 1987).

## I.4 Biosynthesis and post-translational modifications

FVII is synthesized by the hepatocytes and secreted as a serine protease precursor in plasma, where it reaches the concentration of about 10 nM (500ng/ml) (Wion, Kelly et al. 1985).

The domain structures of the vitamin K-dependent coagulation factors FVII, FIX, FX, prothrombin, PC and PS, deduced from their cDNA sequences, demonstrate that they contain common structural features (Kaufman 1998). All contain a signal peptide, that is required for translocation into the lumen of the endoplasmic reticulum (ER), which is followed by a propeptide that directs vitamin K-dependent γ-carboxylation of the mature polypeptide. Upon transit through the trans-Golgi apparatus the propeptide is cleaved away. The amino terminus of the mature protein contains a γ-carboxy glutamic acid rich region (Gla) that includes a short α-helical stack of aromatic amino acids. Then there are two epidermal growth factor (EGF) like domains. The next region is the activation peptide that is glycosylated on an asparagine residue and presents the site of proteolytic cleavage. The remainder of the vitamin K-dependent protein contains the serine protease catalytic triad (Figure 4).
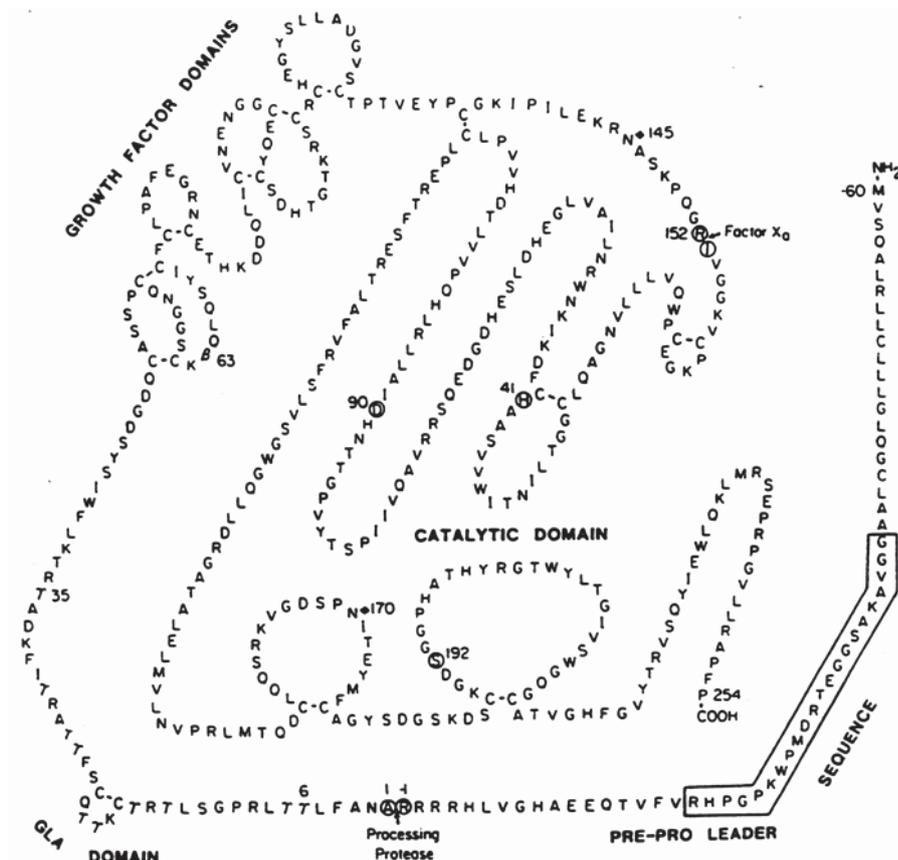


*Figure 4:* amino acidic sequence-based FVII structure

*Signal peptide cleavage.* The precursor form of FVII that is translocated through the ER contains a hydrophobic signal sequence of 38 or 60 amino acids. The signal peptide mediates association of the nascent polypeptide with the cytosolic face of the ER. It is composed of three regions: 1) an amino terminal segment with a net positive charge, 2) a central hydrophobic core of 6-15 residues, and 3) a C-terminal region that often has a helix breaking residue (i.e Gly, Pro, Ser). The cleavage site for signal peptidase is marked by small amino acids (Ala or Gly) in the -3 and -1 positions relative to the cleavage site. Cleavage of the signal sequence by the signal peptidase then releases the mature amino-terminus into the lumen of the ER and is required for the translocation to the secretory pathway (Kaufman 1998).

*Disulfide bond formation.* The vitamin K-dependent coagulation factors have some conserved disulfide bonds. Generally, three disulfide bonds occur within each EGF domain, and several disulfide bonds occur within the serine protease catalytic domain. In addition, in FVII, FIX, FX and PC a disulfide bond connects the amino terminal half with the carboxyl terminal half of the protein to prevent the dissociation of the two portions of the molecule after activation.

Disulfide bond formation occurs between cysteins belonging to the two different chains of activated Factor VII ($Cys^{135}$-$Cys^{262}$) and between cysteins of the same chains ($Cys^{50}$-$Cys^{61}$, $Cys^{55}$-$Cys^{70}$, $Cys^{72}$-$Cys^{80}$, $Cys^{91}$-$Cys^{102}$, $Cys^{98}$-$Cys^{112}$, $Cys^{114}$-$Cys^{127}$, $Cys^{159}$-$Cys^{164}$, $Cys^{178}$-$Cys^{193}$, $Cys^{310}$-$Cys^{329}$, $Cys^{340}$-$Cys^{368}$). Disulfide bond formation occurs in the oxidizing environment of the ER and it is possible that protein chaperones such as protein disulfide isomerase (PDI) are important to ensure proper disulfide bond (Kaufman 1998).

*Asparagine- and Serine-linked glycosylation.* FVII presents two sites of N-glycosylation ($Asn^{145}$ and $Asn^{322}$) and two sites of O-glycosylation ($Ser^{52}$ and $Ser^{60}$). Addition of N-linked oligosaccharides is an obligatory event for the folding and assembly of newly synthesized polypeptides (Helenius 1994). The presence of oligosaccharides is often required for the efficient transport of glycoproteins through the secretory pathway (Dorner, Bole et al. 1987). In addition, N-linked glycosylation frequently affects the plasma half-life and biological activity of glycoproteins. The ER lumenal enzyme oligosaccaryltransferase catalyzes the transfer of a preassembled high mannose containing an oligosaccaride core structure from a dolichol pyrophospahte precursor on to asparagine acceptor sites within the consensus sequence Asn-X-Ser/Thr, where X can be any amino acid except proline. Transit out of the ER is the rate-limiting step in secretion for the majority of proteins and may vary from 15 min to days, depending upon the rate by which a polypeptide attains a properly folded conformation. Upon transit through the Golgi apparatus a series of additional carbohydrate, modifications occur that are separated spatially and temporarily. These reactions

occur by specific glycosyltransferase that modify the high mannose carbohydrate to complex forms. Also within the Golgi apparatus, O-linked oligosaccharides are attached to the hydroxyl of serine residues through an O-glycosidic bond to N-acetylgalactosamine. O-glycosilation occurs in the Golgi complex concomitant with processing of complex N-linked oligosaccharides. Although the functional significance of these O-linked residues is not known, this unusual structure may have some importance since the S52A FVII mutant possesses only 60% coagulant activity of wild type (Bjoern, Foster et al. 1991).

*γ-Carboxylation of glutamic acid residues.* The Gla residues are essential for vitamin K-dependent coagulation proteins to attain a calcium-dependent conformation and for their ability to bind phospholipid surfaces, an essential interaction for their function. The precursor of FVII contains a propeptide that directs γ-carboxylation of 10 glutamic acid residues at the amino-terminus of the mature protein (residues 6, 7, 14, 16, 19, 20, 25, 26, 29 and 35). The propeptide (residues −17 to −1) of FVII shares amino acid similarity with other vitamin K-dependent proteins, by conservation of the γ-carboxylase recognition site and the cleavage site of the propeptide. NMR structural analysis of prothrombin identified that the propeptide is an amphipatic α-helix with the carboxylase recognition site N-terminal to the helix (Sanford, Kanagy et al. 1991).

The function of Gla residues within the coagulation factors was studied by isolation of proteins from animals treated with inhibitors of γ-carboxylation, such as dicoumarol, by proteolytic removal of the Gla domain, and by site-directed mutagenesis of specific Gla residues. For example, des-γ-carboxy prothrombin binds $Ca^{2+}$ much more weakly and is defective in procoagulant activity (Nelsestuen and Suttie 1972). Analysis of partially carboxylated prothrombins demonstrated that their activation rates in coagulation assays are proportional to the number of Gla residues present. The cleavage of the Gla domain from vitamin K dependent factors is coincident loss of low affinity $Ca^{2+}$ sites and with a greatly reduced biological activity.

Several moderate- to low-affinity calcium binding sites exist in the Gla domains of factors VII, IX, X, and protein C that are necessary for a conformational change requisite to phospholipids binding and as coordination sites for phospholipid binding.

In the absence of calcium ions, the Gla domain is disordered, whereas in the presence of calcium ions a unique structure is obtained.

The vitamin K-dependent γ-glutamyl carboxylase enzyme converts glutamate residues to Gla residues. In the presence of $CO_2$, $O_2$ and vitamin K hydroquinone ($KH_2$) the enzyme is able to carboxylate a peptide yielding Gla residues, vitamin K epoxide and $H_2O$. The vitamin K epoxide

formed is subsequently reduced by either a thiol or the enzyme vitamin K epoxide reductase to regenerate $KH_2$.

High expression levels of the vitamin K-dependent plasma proteins in transfected mammalian cells is limited by the ability of the host cells to efficiently perform γ-carboxylation of the glutamic acid residues as well as efficient cleavage of the propeptide.

*β-hydroxylation.* The unusual amino acid erythro-β-hydroxyaspartic acid, formed by post-translational hydroxylation of an aspartic acid residue, has been found in the EGF domain of FVII at position 63. Its function is unknown, as β-hydroxylation is unnecessary for high affinity calcium binding to the first EGF domain and inhibition of β-hydroxylation of factor IX expressed in mammalian cells did not reduce functional activity in factor IX (Derian, VanDusen et al. 1989; Sunnerhagen, Persson et al. 1993).

*Proteolytic processing.* Propeptide cleavage occurs in the trans-Golgi compartment just prior to secretion from the cell. The localization of propeptide processing to this compartment ensures that the propeptide is associated with the mature polypeptide as proteins transit the secretory compartment. Characterization of the amino acid requirements at the propeptide cleavage site has identified that arginines at position 1 and 4 are important for processing. The enzymes candidates for this process are the subtilisin-like serine protease furin/PACE and PACE4, ubiquitously expressed but to a greater extent in the hepatocytes (Kaufman 1998).

## I.5 Activation

After purification of bovine FVII in 1974 Jesty and Nemerson asserted that FVII "apparently exists in plasma not as a zymogen, but in a partially active form" (Jesty and Nemerson 1974). One year later Radcliffe and Nemerson better defined this characteristic of FVII (Radcliffe and Nemerson 1975). In the presence of $Ca^{2+}$ and phospholipids, single chain FVII is rapidly hydrolyzed by FXa and by thrombin to a two-chain form joined by disulphide bridges. This proteolysis is accompanied by an increase of at least 85-fold in the specific FVII coagulation activity with respect to the single chain species. In this report, the term "activated FVII" was used for the first time to depict the two-chain FVII form. The activation of Factor VII involves the cleavage of the single peptide bond located at $Arg^{152}$-$Ile^{153}$ in the sequence Arg-Ile-Val-Gly-Gly (Radcliffe and Nemerson 1976).

In 1977 Kisiel *et al*. proved that FVII can be converted to FVIIa also by FXIIa (Kisiel, Fujikawa et al. 1977), while in 1979 Seligsohn *et al*. demonstrated the important role of FIXa in FVII activation

(Seligsohn, Osterud et al. 1979). Several years later Nakagaki reported an autocatalytic mechanism of FVII activation following complex formation of FVIIa with TF, which may play a key role in the initiation of extrinsic coagulation in normal hemostasis (Nakagaki, Foster et al. 1991).

A detailed kinetic estimate of FVII activation was performed in 1996, by Butenas and Mann, who studied the catalytic efficiency of several plasmatic enzymes and complexes towards FVII. A very sensitive fluorogenic substrate permitted the evaluation of FVIIa activity at nanomolar and subnanomolar concentration of this enzyme (Butenas and Mann 1996). The FVIIa-TF complex was able to generate detectable levels of FVIIa only when high concentrations of the enzyme complex (0.1 nM) and of anionic phospholipids (Phosphatidylserine 25%/ phosphatistlcoline 75% vesicles, PCPS: 200 μM) were used. At physiologic concentrations, FXa was found to be a more effective activator (at least 15-fold better) of FVII than the FVIIa-TF complex.

FVIIa, in the absence of TF, failed to activate FVII at detectable rates even if very high enzyme concentrations and long incubation times were tested. Tissue factor increased the ability of FVIIa to activate FVII to approximately 2-3% of that observed for FXa. No detectable activation of FVII was observed when thrombin, FIXa or FXIa were used as activators. FVIIIa in the presence of PCPS had no effect on the ability of FIXa to activate FVII; FVa progressively decreased the FVII activation rate by FXa. These data suggested that the predominant physiological FVII activator is, most likely, membrane-bound FXa.

Basal *in vivo* levels of FVIIa are thought to be primarily generated by FIXa (Bauer 1997). This statement is based on data showing that in patients with severe FIX deficiency the mean FVIIa level is markedly suppressed and the administration of full replacement doses of FIX led to a normalization (Eichinger, Mannucci et al. 1995).

## I.6 FVIIa structure

Cleavage of the single peptide bond located at $Arg^{152}$-$Ile^{153}$ generates the mature N-terminal residue $Ile^{153}$, and enables conformational changes that create the active enzyme. The structural modifications arise in a contiguous collection of four peptide segments collectively termed the "activation domain". Among these fragments the most important is the new N-terminus, which becomes buried with its non polar side chain in a hydrophobic environment and its charged α-amino nitrogen atom forms a salt bridge with the carboxyl group of $Asp^{343}$ side chain. The three associated segments that undergo changes, creating the substrate binding cleft, are termed, based on chymotrypsin numbering, Loop 140s (142-152, in FVIIa 285-294), Loop 1 (186-194 in FVIIa 334-343) and Loop 2 (216-223 in FVIIa 365-372).

Blood coagulation FVIIa is a trypsin-like plasma serine protease and in its catalytic domains has strong primary sequence identity and tertiary structure similarity with trypsin and chymotrypsin. As in chymotrypsin, FVIIa catalytic domain folds into two domains of the antiparallel β-barrel type, each containing six β strands . The active site is situated in a fissure between the two domains. The enzyme provides a general base, a His residue, which can accept the proton from the hydroxyl group of the reactive Ser thus facilitating formation of the covalent tetrahedral transition state. The His residue is part of a catalytic triad consisting of three side chains from Asp, His and Ser. One domain contributes two of the residues in the catalytic triad, His[193] and Asp[242], whereas the reactive Ser[344] is part of the second domain. Tight binding and stabilization of the transition state intermediate is facilitated by formation of hydrogen bonds between the enzyme and the substrate. These groups are in a pocket of the enzyme called oxyanion hole, while substrate specificity is dictated by the perfect fitting of the preferred side chains into pockets of the enzyme called specificity pockets. The substrate specificity pocket accommodates the side chain of the residue preceding the scissile bond.

A fine characterization of the crystal structure of FVIIa has centred the interest on the loops between the β strands that, for their variable length and composition, usually confer specificity to proteinases and are thought to participate in the mechanism of the enzymatic activity (Kemball-Cook, Johnson et al. 1999; Pike, Brzozowski et al. 1999)(Figure 5). In fact, flexible loops seem involved in the direct recognition of substrates and in the transmission of the cofactor-induced effect from the interface to the catalytic domain (Kumar and Fair 1993; Jin, Perera et al. 2001) and they would also be privileged point of inactivation by degradation (Kemball-Cook, Johnson et al. 1999).
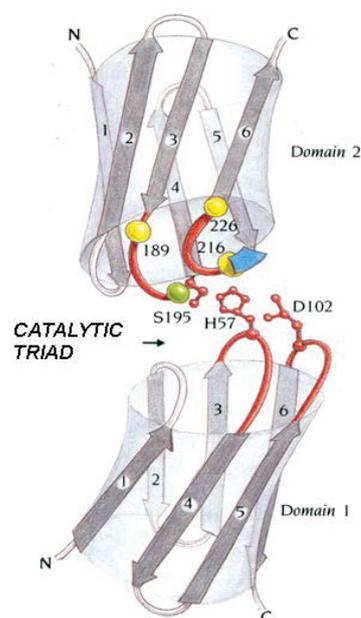


*Figure 5:Representation of the catalytic domain of chymotrypsin*

## I.7 FVII structure

In 2001, trying to isolate crystals containing a shortened FVII construct (EGF2 plus protease domains) and the potent FVIIa/TF inhibitor peptide A-183, Eigenbrot *et al.* obtained crystals containing the zymogen instead of the enzyme (Eigenbrot, Kirchhofer et al. 2001). Compared to the soluble TF-FVIIa complex structure, the key feature of this zymogen structure is a unique registration of the β strand B2 that permits $Glu^{296}$ H bonds with residues near the scissile $Arg^{152}$-$Ile^{153}$ peptide bond and precludes TF binding. Because the energetic cost of the transition between the two β strand B2 frames seemed small, it was suggested that there might be also a significant minority of FVII molecules that have a re-registered B2, resulting in loss of the $Glu^{296}$ H bonds with residues 158 and 159 and a competent TF binding region. If TF binds to FVII, it would select for this species. When TF-FVII complex undergoes the activating cleavage reaction, the $Ile^{153}$-$Asp^{343}$ salt bridge can be formed immediately, and a fully competent enzyme results. Alternatively, when FVII is cleaved before association with TF, the equilibrium mixture would include a population in which H bonds between $Glu^{296}$ and residues 158-159 prevent the formation of the $Ile^{153}$-$Asp^{343}$ salt bridge. The catalytic activity of this equilibrium mixture would then be low due to the predominance of this form.

It has long been known that coagulation factor VII/VIIa is present in circulating blood but is largely inactive and even the nominally activated FVIIa retains zymogen character. Upon exposure to the extravascular environment, FVIIa forms a complex with the cell-surface bound TF and this combination activates downstream clotting factors. TF helps localize and orient FVIIa and guides substrates to productive interactions with the FVIIa active site.

## I.8 FVII activity

FVIIa catalyses the hydrolysis of peptide bonds within a polypeptide chain to produce two new smaller peptides. It is a trypsin-like enzyme, recognizing and cleaving peptide bonds after Arg or Lys side chains and the reaction proceeds in two steps. The first step produces a covalent bond between C1 of the substrate and the hydroxyl group of the reactive Ser of the enzyme. Production of this intermediate proceeds through a negatively charged transition state intermediate. During this step the peptide bond is cleaved, one peptide product is attached to the enzyme in the intermediate and the other product rapidly diffuses away. In the second step, the intermediate is hydrolysed by a water molecule, releasing the second peptide with a complete carboxy terminus and restoring the Ser hydroxyl of the enzyme. This step proceeds through a negatively charged tetrahedral transition

state intermediate. During the formation of the intermediate, the His of the catalytic triad accept a proton, first from the Ser and then from the water molecule and its positive charge is buffered by the Asp of the triad.

In 1964, FVIIa proteolytic activity was observed first against FX, but the methods used didn't reveal whether TF first reacted with FVII to form an intermediate that then activated FX or whether TF directly activated FX. A few years later, it was demonstrated that TF interacted with FVII and the formed intermediate was catalytically active (Nemerson 1966). The reaction product of TF and FVII is a potent activator also of FIX (Osterud and Rapaport 1977)(Figure 6).

The ability of FVIIa to autoactivate FVII molecules was proved by Pedersen in 1989 (Pedersen, Lund-Hansen et al. 1989). thus leading to the concept that trace amounts of FVIIa could be responsible for the initiation of the coagulation cascade upon TF exposition in the vascular lumen, through the formation of trace concentrations of FXa that could then process FVII bound to TF (Rao and Rapaport 1988) (Figure 6).
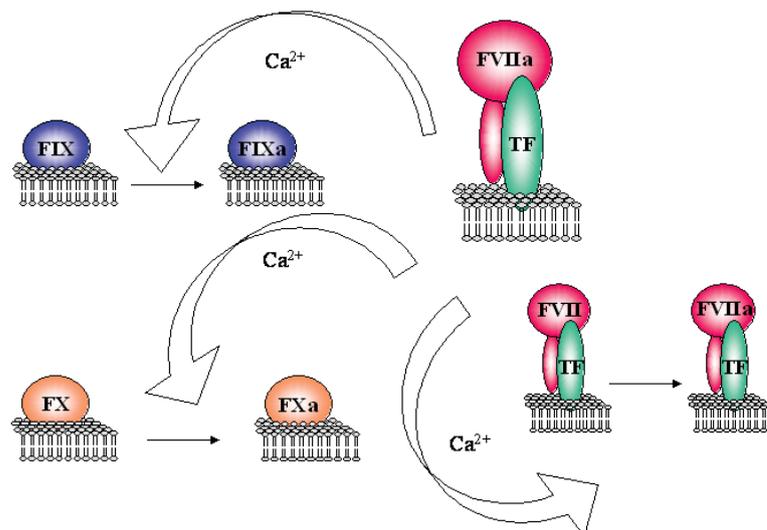


***Figure 6:*** *FVIIa activity*

The kinetics of substrate hydrolysis by FVIIa were extensively investigated (Komiyama, Pedersen et al. 1990). Binding of FVIIa to TF increases the catalytic efficiency of FVIIa of more than 100 fold. It was suggested that TF might induce an alteration in the catalytic site of FVIIa, which allows a more efficient hydrolysis of the small fluorescent substrate. Measurements conducted using various phospholipids and detergents demonstrated that the increase in the catalytic efficiency of FVIIa, when complexed to TF, is independent of the supporting surface (Lawson, Butenas et al. 1992).

The interaction of FVIIa with TF is $Ca^{2+}$-dependent. $Ca^{2+}$ saturation of the Gla domain is likely responsible for this increase in affinity, since deletion of Gla in FVIIa results in a loss of affinity for TF. $Ca^{2+}$ may stabilize energetically important hydrophobic contacts of Gla with TF (Kelly,

Dickinson et al. 1997). In addition, $Ca^{2+}$ increases FVIIa affinity for FX by conformational changes in FVIIa and FX that are essential for the interaction of these proteins with phospholipids (Bom and Bertina 1990). This conclusion is reasonable considering the membrane-binding capability of FX and the great enhancement in productive collisions between substrate and enzyme realized by an initial interaction of the substrate with the membrane surface.

## I.9 Half-life and degradation

Regulation of FVIIa activity is a key step and it includes inactivation by plasmatic inhibitors such as antithrombin and TFPI, internalization and degradation. TFPI directly inhibits FVIIa-TF complex and FXa activity and the formation of the quaternary complex promotes its internalization by about 3-fold (Warshawsky, Herz et al. 1996; Iakhiaev, Pendurthi et al. 1999). Alternatively, TFPI anchored to glycosyl phosphatidylinositol can mediate a transient down regulation of the quaternary complex through its translocation to glycosphingolipid-rich microdomains, unfavorable for FVIIa-TF activity (Ott, Miyagi et al. 2000). Hansen *et al*. also described a clathrin-independent mechanism of FVIIa-TF internalization not affected by the presence of TFPI (Hansen, Pyke et al. 2001). Internalized FVIIa can return to the cell surface, as recycled fully active FVIIa, or associate with nuclear fractions. Whether FVII/FVIIa is degraded in plasma prior to internalization is not known. *In vitro* degradation of FVIIa occurs through cleavage after $Arg^{290}$ and $Arg^{315}$ (Nicolaisen, Thim et al. 1993). These residues belong to Loop 140s and Loop 170s, respectively, solvent-exposed structures that in the activated form of FVII present a high degree of flexibility.

The circulating half-life of the zymogen FVII in humans has been reported to be approximately 5 hours (Hasselback and Hjort 1960), while it was about 2.5 hours for the activated form (Seligsohn, Kasper et al. 1979). Compared to other vitamin K-dependent coagulation proteases, the circulating half-life of FVIIa is extremely long.

# I.10 FVII deficiency

Inherithed FVII deficiency, first described by Alexander *et al.* in 1951 (Alexander, Goldstein et al. 1951), is the most frequent of the rare congenital coagulation disorders with an estimated prevalence of 1 in 300.000-500.000 individuals (Mariani, Lo Coco et al. 1998). It is usually transmitted in an autosomal recessive fashion and it is frequently associated with consanguinity. Triplett *et al.* have classified FVII deficiency in CRM⁻ (activity and antigen proportionally reduced), CRM⁺ (reduced activity, antigen normal) and CRM^red (antigen is reduced but not as much as activity) (Triplett, Brandt et al. 1985).

There is a considerable phenotypic (Triplett, Brandt et al. 1985) and molecular heterogeneity (Mariani, Herrmann et al. 2000; McVey, Boswell et al. 2001) in the congenital FVII deficiencies. The clinical bleeding tendency ranges in severity from lethal to mild, or even asymptomatic forms. Some individuals experience mild mucous membrane bleeding, menorrhagia, and post-surgical bleeding, but more significant events such as hemarthroses and soft tissue bleeds are documented. A higher prevalence of females was found among symptomatic subjects and in particular among moderate bleeders: much of the excess of bleeding tendency can be attributed to menorrhagia, the most frequent symptom in this gender. Life-threatening gastrointestinal and central nervous system bleeds are well recognized (GI or CNS, 20% of patients) and are characterized by early presentation and association with lower FVIIc levels. Life-threatening bleeds occur most frequently (70% of the cases) during the first 6 months of life and are associated to high morbidity and mortality rate. The greatest risk factor for CNS hemorrhage is trauma related to the birth process (Ragni, Lewis et al. 1981).

The potential severity of the clinical phenotype of FVII deficiency reflects its pivotal role in the initiation of coagulation. Mice with targeted disruption of their FVII gene show lethal hemorrhage in the peri-partum period: 70% suffered fatal intra-abdominal bleeding within the first 24 hours and most of the remaining neonates died from intracranial hemorrhage before the age of 24 days (Rosen, Chan et al. 1997). Interestingly, the FVII deficient embryos develop to term and do not exhibit the developmental lethality at mid-gestation experienced by TF deficient embryos. In fact, TF is also implicated in a variety of biological processes, from angiogenesis and tumor metastasis to vascular remodelling and signal trasduction (Zhang, Deng et al. 1994; Bromberg, Konigsberg et al. 1995; Rottingen, Enden et al. 1995; Khachigian, Lindner et al. 1996).

In human neonates that are homozygous for a Factor VII null allele development in utero is normal but mortality occurs shortly after birth due to intracranial haemorrhage. The minimal FVII level able to interact with TF to prevent lethal bleeding in human subjects has not yet been defined.

The clinical phenotype in patients with FVII deficiency correlates poorly with FVII coagulant activity (FVII:C) measured *in vitro*. This lack of correlation probably reflects the fact that only trace amounts of FVIIa are required to initiate coagulation *in vivo*, and *in vitro* tests fail to differentiate between a 'true' null mutation and one that results in very low but not-zero FVII:C levels, capable of initiating coagulation *in vivo* and resulting in a mild/moderate bleeding phenotype. Furthermore, FVII:C levels were usually measured using a non-human source of TF and this could generate values discrepant with those obtained with human TF.

## I.11 FVII levels

Plasma levels of FVII protein and procoagulant activity vary significantly in the general population (18% and 26% respectively) and are influenced by different environmental factors including sex, age, body mass index and diabetes (Balleisen, Assmann et al. 1985). In women, in whom the increase in FVII with age appears to be greater than for men (Meade, Haines et al. 1983), levels of FVII have also been linked to use of oral contraceptives (Balleisen, Assmann et al. 1985), reproductive status (Meade, Haines et al. 1983), and use of oestrogen hormone replacement therapy (Meilahn, Kuller et al. 1992). Variations in plasma FVII levels can also be attributed to genetic factors as demonstrated for several FVII polymorphisms.

*-402 and –401 polymorphisms.* The G to A substitution at position –402 and the G to T substitution at position –401 are two common, nonrelated, functional polymorphisms in the promoter region of the FVII gene. Both polymorphisms strongly influence the binding properties of nuclear proteins. The rare –401T allele is associated with a reduced basal rate of transcription of the FVII gene in human hepatoblastoma cells and with reduced plasma concentrations of total FVII and FVIIa molecules. In contrast, the rare –402A allele confers increased transcriptional activity and is associated with increased plasma FVII levels (van 't Hooft, Silveira et al. 1999).

*Decamer insertion at -323.* Studies of FVII levels in healthy individual have shown that insertion of the sequence CCTATATCCT at position –323 in the 5' UTR of the FVII gene is associated with a decrease of about 25% in FVII levels. The allele with the decanucleotide insertion is called the A2 allele while the one lacking the insertion is called the A1 allele. Clear evidence concerning the effect of this decanucleotide was provided through the examination of promoter strength by transfection experiments in HepG2 cells in which it was shown that the insertion reduced promoter

activity by 33% compared with the allelic sequence which lacks the decanucleotide (Pollak, Hung et al. 1996). The decamer insertion at –323 and the R353Q polymorphism have been shown to be in strong allelic association with each other (Bernardi, Marchetti et al. 1996; Bernardi, Arcieri et al. 1997).

*G73A polymorphism*. This polymorphism is located in intron 1a of the FVII gene and caused by the nucleotide change G to A at position 73. It is often associated with the promoter decamer insertion and the Q353 alleles, thus impairing the understanding of the A73 allele per se contribution to lowering FVII levels in plasma. The concomitant presence of A73 allele with both the decamer insertion and the Q353 alleles was associated with the lowest factor VII levels and might confer protection against myocardial infarction in the young (Peyvandi, Mannucci et al. 2000).

*Arg353Gln polymorphism*. This polymorphism results from a G to A transversion at position 10976 in exon 8 (Green, Kelleher et al. 1991). The Arg353 allele is referred as M1 allele, while the Gln353 allele is called M2. The M2 allele is associated with a decrease of about 25% in FVII:C and FVII:Ag levels (Green, Kelleher et al. 1991; Bernardi, Marchetti et al. 1996; Bernardi, Arcieri et al. 1997). The conformation of the Gln 353 molecule may be different from that of the Arg 353 protein, affecting its intracellular processing, secretion, turnover in plasma, or activity. *In vitro* expression studies in COS-1 cells have demonstrated that the Q353 variant was secreted with a significantly reduced efficiency (Hunault, Arbini et al. 1997). Analysis of the crystal structure of the soluble TF-FVIIa complex reveals a peripheral location for Arg[353]. It has been proposed that this residue may be involved in interaction between triglyceride-rich lipoproteins and FVII and the substitution to Gln may therefore alter the strength of this interaction that limits or slows cleavage to the active two-chain form, or rate of removal from the circulation (Humphries, Lane et al. 1992). The Gln variant occurs with a frequency of about 10% in various populations and this high frequency could indicate that the variant confers some benefit, for example protection against thrombosis or myocardial infarction (Hunault, Arbini et al. 1997).

*Variable number tandem repeat polymorphism (VNTR) in intron 7*. This polymorphism spans the exon 7-intron 7 boundary and is due to a variation in repeat copy number of a 37 bp element. This microsatellite is composed by the last four nucleotides of exon 7 and the first 33 base pairs of intron 7. In that way, the IVS7 repeated microsatellite provides multiple identical donor splice sites, that remain normally silent (O'Hara and Grant 1988). Four different alleles with 5 to 8 monomer repeats have been reported and the most common are those containing 6 and 7 repeats (Bernardi, Marchetti et al. 1996; Bernardi, Arcieri et al. 1997), designated as b and a, respectively. The allelic forms with

a lower number of repeats was found to be associated with a decrease in FVII levels (Bernardi, Arcieri et al. 1997), very likely caused by reduced efficiency of mRNA splicing (Pinotti, Toso et al. 2000).

*His115His*. The polymorphism is located within exon 5 and results from a C to T change at position 7880 (codon 115). This change is silent at the amino acid level (Chaing, Wallmark et al. 1994). The most common C allele is generally referred as H1 and the rare T allele is called H2.


## I.12 FVII mutations


A considerable number of mutations have been reported to date in FVII gene (F7 mutation database: http://www.hgmd.cf.ac.uk/ac/index.php) (Mariani, Herrmann et al. 2000; McVey, Boswell et al. 2001; Mariani, Herrmann et al. 2005; Krawczak, Thomas et al. 2007). The majority of individuals with mutations in their FVII gene are either asymptomatic or the clinical phenotype is unknown and have come to notice through pre-operative clinical tests.

Missense mutations were the most frequent and occurred in the 68% of subjects, followed by splicing-site (13%), promoter (8%) and nonsense (6%) mutations, small insertions and deletions (6%)(Mariani, Herrmann et al. 2000; Mariani, Herrmann et al. 2005). Many of these mutations have been identified as the cause of FVII deficiency only a limited number of them have been expressed and characterized. In some cases, naturally occurring FVII mutants constitute valuable tools to investigate single residues or to define regions important in the structure-function relationship in FVII and in the formation of macromolecular complexes responsible for coagulation initiation.

Among FVII deficient patients, the most severe cases are all either homozygous or doubly heterozygous for deleterious mutations resulting in FVII:C levels less than 2% of normal. The majority are mutations that disrupt appropriate expression: promoter, splice-junction or frameshift mutations caused by deletions. Only a few missense mutations have been described that result in a severe phenotype (McVey, Boswell et al. 2001).

Cases of mild/moderate FVII deficiency have *in vitro* FVII:C levels which range from <1% to 52% and these levels don't correlate with the reported clinical severity. Thus it's impossible to differentiate between the severe and mild/moderate cases only on the bases of FVII:C or FVII:Ag in plasma.

The asymptomatic cases have FVII:C ranging from 4% to 61% and FVII:Ag levels from 5% to 113% of normal. The mutations are all missense.
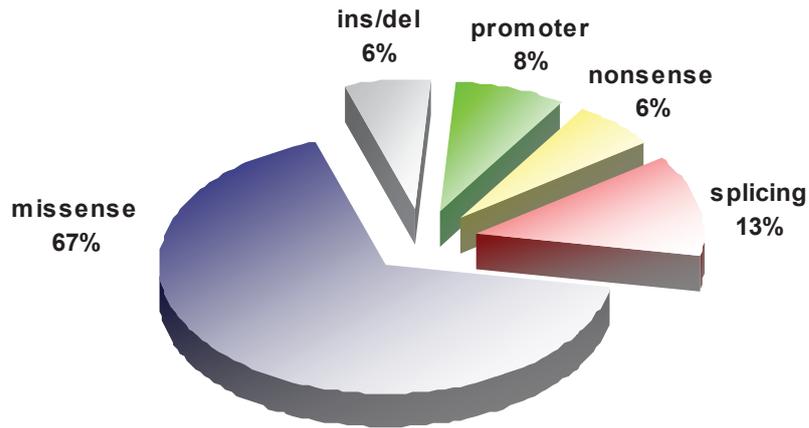
***Figure 7:*** *Pie chart showing the type of mutations reported*
*in the International Registry of FVII Deficiency (IRF7)*

## I.13 Conventional treatment strategies in FVII deficiency

As for the hemophilias, replacement of the deficient coagulation factor is the mainstay of treatment for FVII deficiency, but safe and efficacious products are fewer and experiences on their optimal use much more limited (Mannucci, Duga et al. 2004).

*Intermediate Purity Factor IX Concentrates and prothrombin complex concentrates (PCCs):* their main advantages are the small volume of infusion, fewer allergic reactions, and the adoption of virus-inactivation procedures during manufacturing.
These products are not calibrated for FVII concentrations and as the half- life of FVII is much shorter than that of other coagulation factors present, multiple doses of PCCs may result in a build-up of other factors, increasing thrombotic risk.

*Plasma-derived Factor VII concentrates:* FVII concentrates are prepared from pooled plasma. They are used for prophylactic treatment, as well as for controlling serious bleeding episodes, and bleeding during surgery. However, plasma-derived concentrates carry the risk of potential transmission of blood-borne pathogens

*Fresh frozen plasma (FFP):* single-donor FFP, that contains all coagulation factors, is relatively inexpensive and widely available. However, because of the very short half-life of FVII the risk of volume overload is real when repeated infusions are administered to raise and keep the deficient factor at hemostatic levels. Hence, concentrates should be preferred for major surgical procedures

or when the severity of the clinical manifestations predicts a long-lasting treatment. Most importantly, infectious complications with viruses as the hepatitis viruses or human immunodeficiency virus (HIV) are still perceived as a threat of FFP.

*Recombinant activated FVII (rFVIIa):* rFVIIa is indicated for the treatment of bleeding episodes and for the prevention of bleeding in patients with congenital FVII deficiency undergoing surgery procedures. It is free of human plasma and albumin, so there is no risk of human viral transmission, but it is very expensive and not available for all patients.

## I.14 FIX

## I.14.1 F9 gene

Coagulation factor IX gene (F9) is located on the X chromosome, on the long arm, more towards the centromere at Xq28. The gene is approximately 34 kb in length and contains only eight exons, the largest of which is only 1935 bp. The transcript is 2803 bases in length and comprises a short 5' UTR (29 bp), an open reading frame plus stop codon (1383 bp) and a 3' UTR (1390 bp). The open reading frame encodes a pre-pro-protein in which the signal sequence (pre-sequence) directs factor IX for secretion, the pro-sequence provides a binding domain for vitamin K dependent carboxylase, which carboxylates certain glutamic acid residues in the adjacent Gla domain, and the remainder represents the factor IX zymogen (Anson, Choo et al. 1984). Activation of factor IX involves cleavage of two peptide bonds, one on the C-terminal side of arginine 145 (α-cleavage) the other on the C-terminal side of arginine 180 (β-cleavage). These cleavages are caused by activated factor XI generated through the intrinsic pathway or via tissue factor/activated factor VII complex of the extrinsic pathway. The activation cleavages generate an N-terminal light chain and a C-terminal heavy chain, held together by a disulphide bond between cystein residues 132 and 279 (Bowen 2002) (Figure 8).
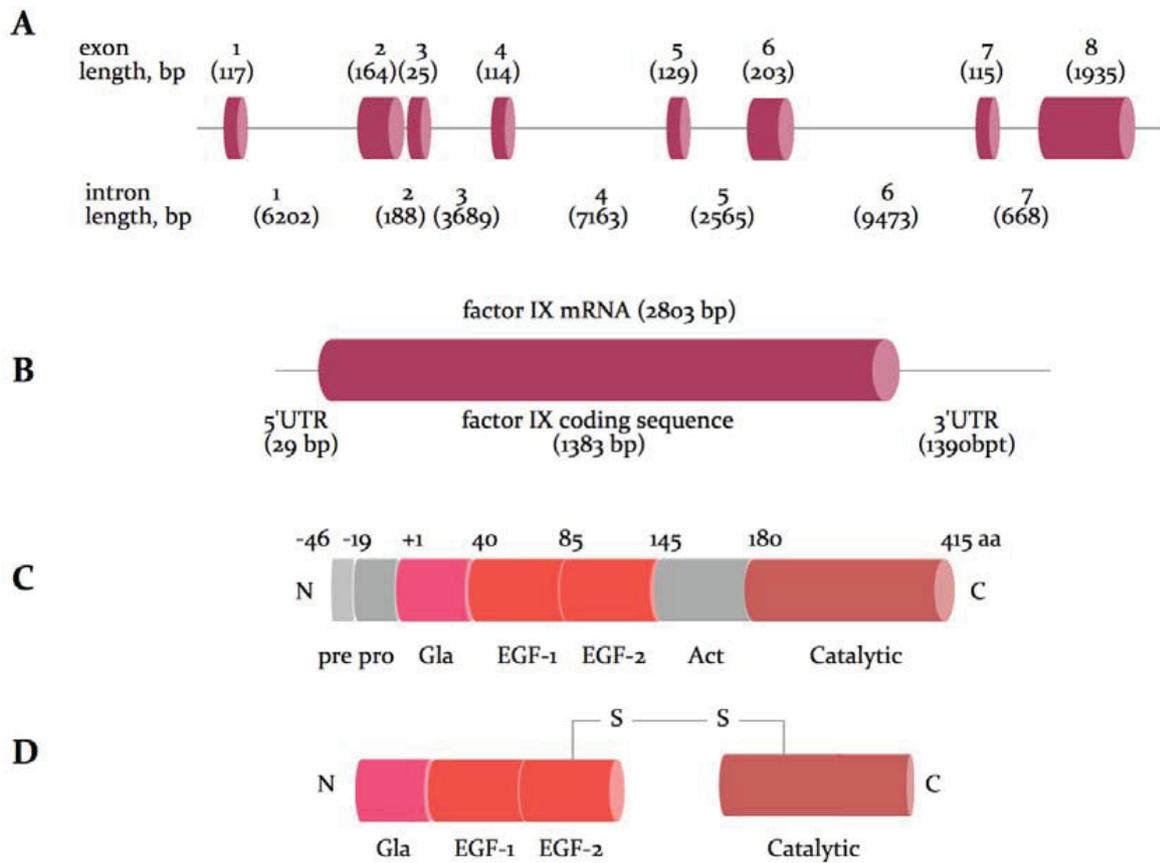
***Figure 8:*** *Schematic representation of F9 gene, mRNA and protein.*
***A,*** *genomic organization of F9 gene. Exon and intron number and size are reported above and below, respectively.*
***B,*** *FIX mRNA showing the relative size and location of the ORF*
***C,*** *the synthetized FIX protein comprising a pre-pro signal sequence and a mature peptide of 45 aminoacids.*
***D,*** *activated Factor IX comprising a N-terminal light chain and a C-terminal heavy chain held together by a disulphide bond between cysteine residue 132 and 279. Gla, Gla domain; EGF, epidermal growth factor-like domain; Act, activation peptide released after proteolitic cleavage; catalytic, the serine protease domain.*

## I.14.2 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, Undas et al. 2009). Factor VIII is a cofactor with no enzymatic activity per se; factor IX is a serine protease with an absolute requirement for factor VIII as cofactor. Upon activation, and in presence of calcium ions and phospholipid surfaces, factor VIII and factor IX form an active complex, which activates factor X. Subsequent stages of the cascade then proceed, culminating in the deposition of fibrin, the structural polymer of the blood clot (Bowen 2002).

## I.14.2 Molecular basis of hemophilia B

Hemophilia B (or Christmas morbus) is a coagulopathy X-linked (Green 1989 telethon ) caused by mutations in the F9 gene. Generally, only male are symptomatic ( incidence of 1:35000 live male births) due to the presence of two alleles in female sujects. Based on FIX levels (antigen and/or protein activity), patients experience severe hemorrhagic symptoms, not rarely life−threatening (central nervous system and gastrointestinal bleeds), or causing substantial handicap (hemarthrosis, muscle hematoma). Based on FIX levels on plasma, the phonotype of patients is classified as mild, moderate or severe:

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

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

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

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

In general nonsense mutations are associated with severe forms of haemophilia; exon skipping is a further possibility arising from a nonsense mutation and it is also extremely detrimental: an in frame skip will result in a protein lacking the aminoacids encoded by skipped exon, an out of frame skip will result in a frame shift (Dietz, Valle et al. 1993; Ketterling, Drost et al. 1999)(Figure 9).

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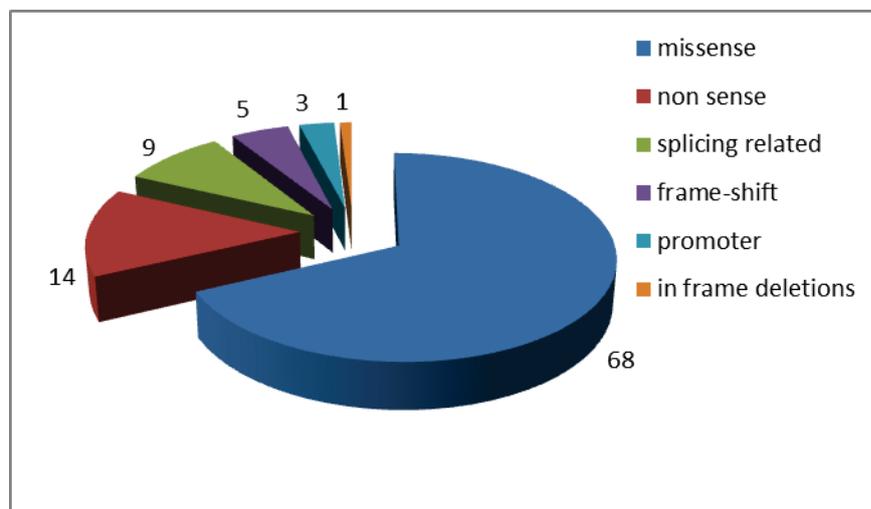


**Figure 9:** *Pie chart showing the type of mutations reported in the Hemophilia B Mutation Database*

## I.15 Conventional treatment strategies haemophilia B

Likely FVII deficiency, the administration of the missing protein (FIX) from various sources is the only therapy available. So FIX concentrates, Prothrombin complexes and fresh frozen plasma remain the only sources of FIX protein for patients. Due to the fact that these preparations contain other coagulation proteins, the continuous administration of FIX mix may result in a build-up of other factors, increasing thrombotic risk. Various drawbacks have been associated with prolonged infusion of FIX concentrates: disseminated intravascular coagulation, pulmonary embolus, and deep venous thrombosis. Moreover, the administration of a protein that the body is not able to synthetize

by itself results, in a long term therapy, in the development of neutralizing FIX antibodies. The costs of substitutive treatments (~50000 euro/year/person with severe disease) are prohibitive for the majority of the world hemophilia populations, so the demand for alternative therapies is growing up.

Enormous efforts have been pushed on the substitutive gene therapy that consists in the viral or non-viral mediated delivery of a copy of the defective gene (or better of the coding DNA sequence) into the patient's cells, thus triggering stable endogenous expression of the missing protein (Murphy and High 2008; Petrus, Chuah et al. 2010). Anyway, there are still numerous issues regarding the delivery method to the target site and safety of gene therapy (e.s risk to induce tumor formation after gene therapy).

## I.16 Bases for a mutation-specific therapy based on modified U1 snRNA.

## I.16.1 The splicing reaction.

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 and Reed 2002). As soon as transcription begins, the first nucleotide of the precursor messenger RNA (pre-mRNA) is linked to a G through a 5'-5'-triphosphate bond and further methylated to create a N7-methylguanosine cap (m1G-cap) (Shatkin 1976; Babich, Nevins et al. 1980; Coppola, Field 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, accounting for no more than 1% of the entire genome (Lander, Linton et al. 2001). Therefore the majority of the protein-coding genes are composed mostly by non-coding, intervening sequences (introns). Within the pre-mRNA, the shorter and less abundant exons must be identified, defined, and assembled in a

mRNA transcript which encodes ultimately the whole protein sequence whereas the introns are removed. This cut and paste reaction is called splicing. Recent findings indicate that almost every multiexon gene in the human genome undergo at least one alternative splicing event highlighting the central role that splicing plays in gene expression (Pan, Shai et al. 2008; Wang, Sandberg et al. 2008). Furthermore, splicing is a very-well conserved pre-mRNA processing mechanism, found from unicellular eukaryotic organisms as Saccharomyces cerevisae to metazoans. It displays increasing levels of regulation and complexity as the number and length of introns in multicellular eukaryotes increases (Ast 2004; Nilsen and Graveley 2010). Thus, splicing constitutes not only a crucial step for accurate transfer of the genetic information from DNA to RNA to protein, but also a step that allows for regulation of gene expression as well as increased protein diversity through alternative splicing decisions. In order to find the short exons among the sea of intronic regions some signals present at the exon/intron boundaries are crucial. In higher eukaryotes, these elements are short consensus sequences surrounding the 3' and 5' end of the introns, which are known as 3' and 5' splice sites (ss), respectively. Their sequences exhibit a variable degree of conservation nonetheless they are fundamental for proper intron recognition and splicing catalysis (Figure 10).

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

***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). The BP is characterized by the presence of a conserved A surrounded by a highly degenerated motif YNYURAY (Y=pyrimidine and R=purine)(Reed and Maniatis 1985). It is commonly found about 18-40 nucleotides upstream of the AG dinucleotide (Ruskin, Krainer et al. 1984; Reed and Maniatis 1985) although some exceptions can be found hundreds of nucleotides away (Reed 1989). The recognition of the branch site involves a base-pairing with the U2 snRNP in order to form the spliceosome A complex (Berglund, Chua et al. 1997).The PPT is a run of pyrimidines (eight bases in the average intron) located between the branch site and the terminal AG at the intron/exon junction (Reed 1989). It can display variable pyrimidine content, length and distance to the branch-point and the AG. The PPT is essential for efficient branch-point utilization and correct AG recognition as it has been shown that progressive deletion of the polypyrimidine tract impairs splicing while elongating its length can improve its efficiency (Roscigno, Weiner 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, Romfo et al. 1999). The sequences between the branch-point and the acceptor site are commonly devoid of AG dinucleotides (Gooding, Clark et al. 2006).
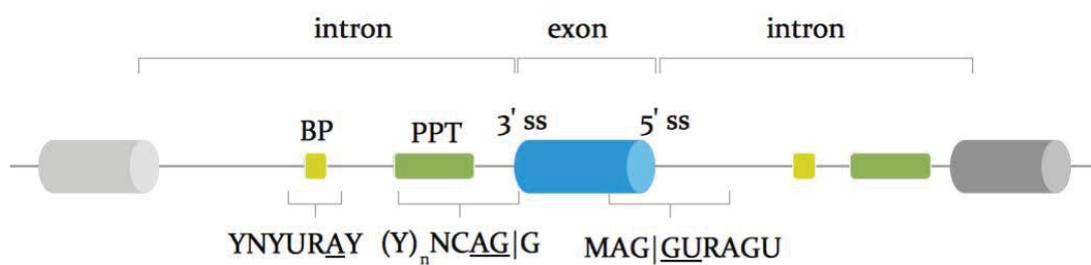


***Figure 10:*** *Schematic representation of exon‑intron boundaries*
*The cylinders represent the exons and the grey line depict the introns. The branch‑‑point (BP) and the polypyrimidine tract (PPT) are illustrated by the rectangles (yellow and green, respectively). Below the consensus sequences for each canonical signal are shown. The nearly invariant GU, AG at 5' and 3' ss, respectively and the conserved A at the branch‑‑point (BP) are underlined. The polypyrimidine tract (PPT) is also represented. Y, polypyrimidines (U or C); R, purines (A or G); M, A or C.; N, any nucleotide.*

## I.16.2 The spliceosome: Assembly of Spliceosomal snRNPs

*Structure of Spliceosomal snRNPs:*

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

The mono-snRNPs just described do not represent their *in vivo* functional forms; rather, they are organized into higher order particles. The U4, U5 and U6 snRNPs exist largely in their functional form as a U4/U6. U5 tri-snRNP (Behrens and Luhrmann 1991; Liu, Rauhut et al. 2006). The same holds true of the U4atac, U5 and U6atac snRNPs, which form the minor spliceosomal U4atac/U6atac.U5 tri-snRNP (Schneider, Will et al. 2002). Interestingly, minor spliceosomal counterparts of the U1 and U2 snRNPs, the U11 and U12 snRNPs, respectively, are known to assemble into the minor spliceosomal U11/U12 di-snRNP. Furthermore, penta-snRNP complexes, which consist of all five major splicing snRNPs and may represent a 'splicing holoenzyme', have been shown to exist in both yeast and humans (Nilsen 2002; Stevens, Ryan et al. 2002; Malca, Shomron et al. 2003) (Figure 11). The snRNPs, along with over 300 other splicing factors, assemble onto pre-mRNA to form the spliceosome, and it is this dynamic macromolecular machine that orchestrates the excision of introns and the ligation of exons through two successive transesterification reactions (Patel and Steitz 2003). Prior to participating in splicing, however, snRNPs must be assembled through a series of intricate steps that, in all organisms, begins in the nuclear compartment. In animalia, protista and plantae, a brief transit to the cytoplasm is essential

for the assembly of Sm snRNPs, but the assembly of the U6 snRNP is uninterrupted by a cytoplasmic phase.
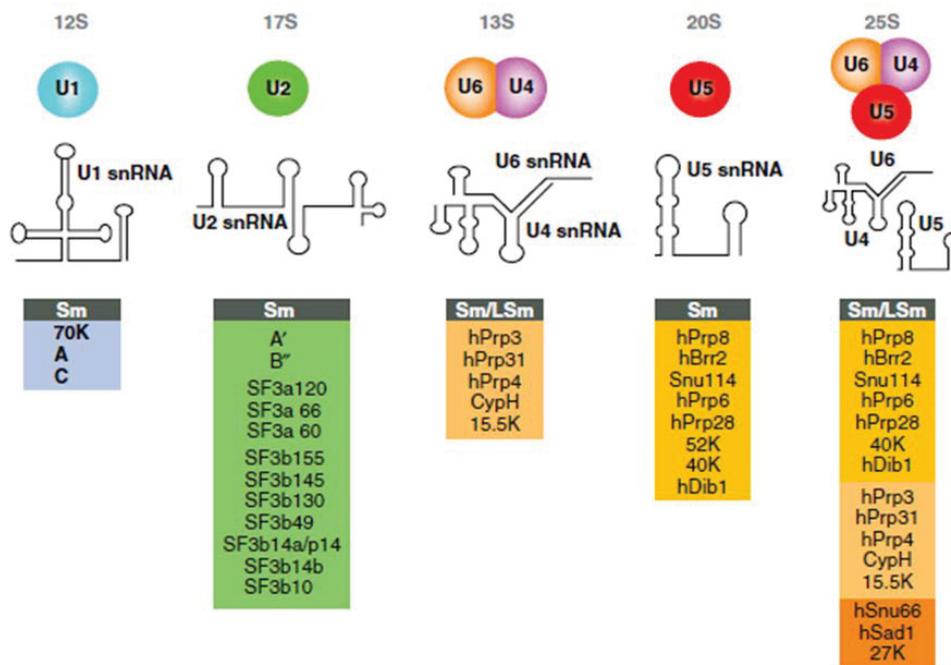


**Figure 11:***Protein composition and snRNA secondary structures of the major human spliceosomal snRNPs. All seven Sm proteins (B/B', D3, D2, D1, E, F, and G) or LSm proteins (Lsm2-8) are indicated by "Sm" or "LSm" at the top of the boxes showing the proteins associated with each snRNP. The U4/U6.U5 tri-snRNP contains two sets of Sm proteins and one set of LSm proteins.*

## I.16.3 The Biogenesis of Spliceosomal snRNPs

The assembly of all spliceosomal snRNPs begins with the transcription of a U snRNA. The genes for the U snRNAs reside in the nuclear genome and are transcribed by either RNA polymerase (RNAP) II or III. During evolution, multiple copies (20-100) of U1, U2, U4 and U5 genes have arisen by gene duplication; however, the U6 gene is present only in approximately five functional copies in the haploid human genome (Lund and Dahlberg 1984; Van Arsdell and Weiner 1984; Domitrovich and Kunkel 2003). Major clusters of the human U1 and U2 genes are present on chromosomes 1 and 17, respectively (Lund and Dahlberg 1984; Van Arsdell and Weiner 1984); whereas, the U6 genes are scattered throughout the genome (Domitrovich and Kunkel 2003). It is clear that the assembly of Sm snRNPs and the U6 snRNP follow two distinct pathways. The Sm snRNAs are transcribed as 3′-extended (2- to 10-nt longer) precursors by RNAPII, and like all other RNAPII transcripts, they co-transcriptionally acquire a 5′- to 5′-linked N7-methyl guanosine ($m_1G$) cap. The 3′-ends of the pre-snRNAs are generated by an RNA processing event that is coupled to

41

the proximal sequence element-directed transcription. A conserved 3′-box (GTTTN$_{0-3}$AAAPuNNAGA, where Pu = purine and N = any nucleotide) marks the cleavage site which resides ∼10-nt upstream and a heterododecameric metallo β-lactamase complex (Integrator) contains the enzymatic activity for 3′-end formation (Baillat, Hakimi et al. 2005). The newly transcribed pre-snRNAs must be transported to the cytoplasm to continue their maturation, necessitating the assembly of an export competent complex. First of all, the nuclear CBC, consisting of CBP20 and CBP80, first associates with the m$_1$G cap of the RNA (Izaurralde, Stepinski et al. 1992; Izaurralde, Lewis et al. 1995). Next, the phosphorylated adaptor for RNA export (PHAX) binds the CBC–RNA complex (Segref, Mattaj et al. 2001). The export receptor, Exportin 1/Chromosome Region Maintenance 1 (Xpo1/CRM1), recognizes the export adaptor, PHAX, in its phosphorylated form bound to its CBC/pre-snRNA cargo and binds to this complex together with RanGTP (Segref, Mattaj et al. 2001). While all of the above interactions are individually quite weak, cooperative binding ensures the formation of a stable export complex. It was shown recently that the assembly of an export competent mRNA begins at the transcriptional unit (Kohler and Hurt 2007; Patel, Novikova et al. 2007). After assembly, the entire complex translocates through the nuclear pore complex (NPC). The transcription of nascent snRNAs, 3′-end formation, assembly into an export complex, putative transit through CBs and translocation through the NPC are rapid processes, as revealed by radiolabeled experiments, requiring only 4 minutes (Eliceiri and Sayavedra 1976; Eliceiri and Gurney 1978). Upon entry in the cytoplasm, PHAX is dephosphorylated by protein phosphatase 2A, but remains associated with the CBC/pre-snRNA complex, presumably until the m$_1$G cap is hypermethylated (Ohno, Segref et al. 2000; Kitao, Segref et al. 2008). CBC and PHAX are subsequently recycled to the nucleus, where the latter is phosphorylated by casein kinase 2 (CK2) to initiate another round of pre-snRNA export. The cytoplasmic phase of snRNP maturation is orchestrated by a large 20S assemblysome called the survival of motor neuron protein (SMN) complex, which consists of the SMN, seven distinct Gemin proteins (Gemin 2-8), and several other protein factors (Battle, Kasim et al. 2006; Battle, Lau et al. 2006; Kolb, Battle et al. 2007). The SMN complex participates in all three snRNP maturation events in the cytoplasm: (i) the assembly of an Sm ring onto the Sm site; (ii) the hypermethylation of the m$_1$G cap; and (iii) the trimming of the pre-snRNA's 3′-end. Indeed, the SMN complex associates with a distinct set of snRNP populations, each representing different stages in their cytoplasmic maturation: a disassembled export complex, the core Sm snRNP and an import complex (Massenet, Pellizzoni et al. 2002) (Figure 12).

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

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

(iii) Nucleolytic trimming of the 3′-end of the pre-snRNA generates the mature length snRNA. While the molecular mechanisms regulating the cytoplasmic maturation events of snRNAs have been extensively studied, the spatial arrangement of these events within the cytoplasm remains poorly understood. A recent study suggests that snRNP maturation might partly occur in discrete cytoplasmic bodies (Liu and Gall 2007). These organelles were named the 'U bodies' because they contain the major U snRNPs.

The core snRNP must be brought into the nucleus to continue its maturation and, afterwards, participate in pre-mRNA splicing. The requirements for nuclear import vary depending on the particular snRNP and on the cell system. In general, the $m_3G$ cap and the Sm core are considered to be nuclear localization signals (NLS) that utilize the same import receptor importin β (Imp β) but distinct import adaptors, snurportin-1 (SPN1) and possibly SMN, respectively. The nuclear phase is

the least understood part of the entire snRNP biogenesis pathway, and it involves a multiplicity of processes and factors, as well as trafficking to several subnuclear domains. In particular, extensive internal modifications of the U snRNAs by 2′-O-methylation and pseudouridylation represent a critical step in the making of a fully functional snRNP.
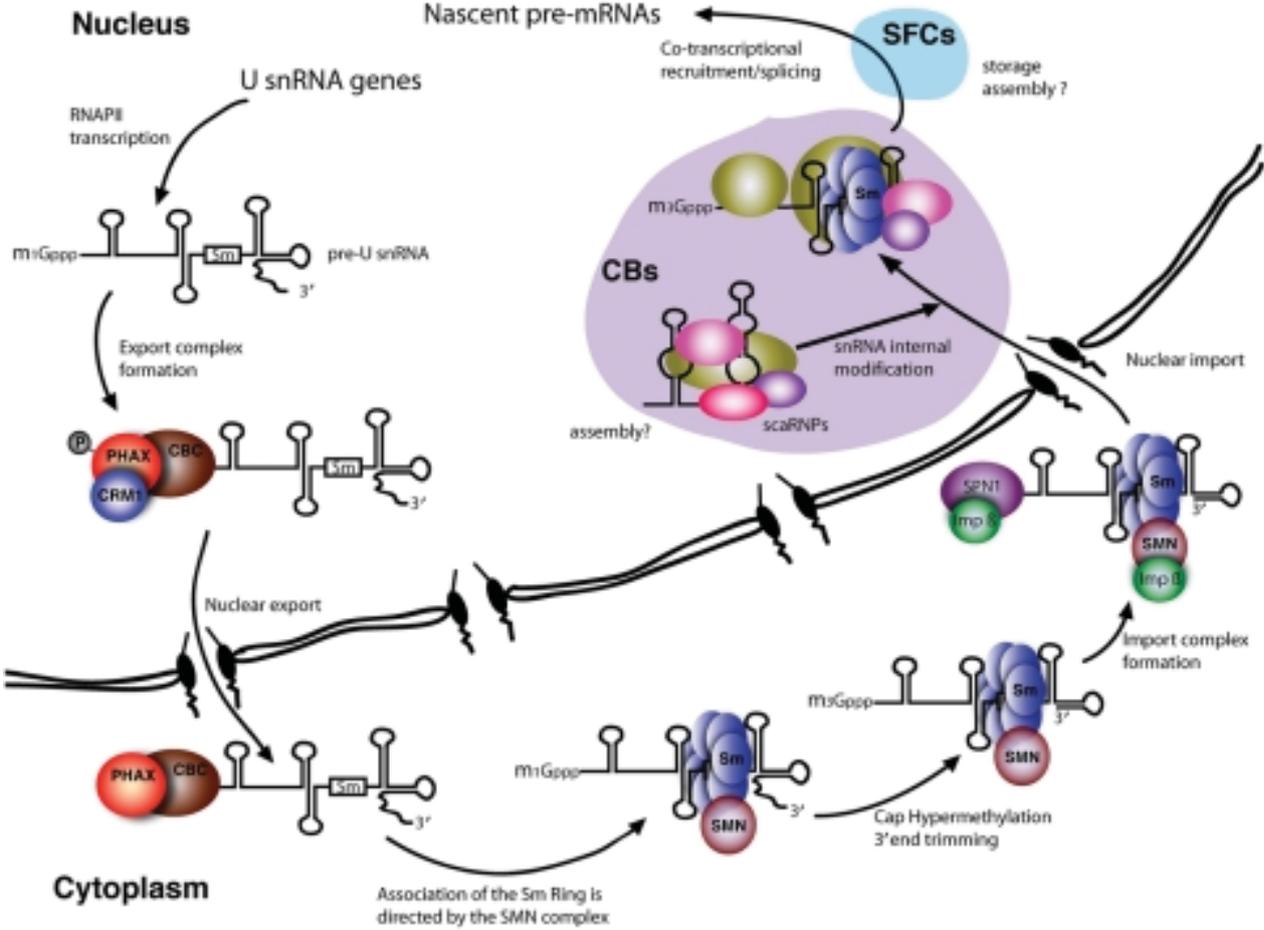


***Figure 12:****The Sm snRNP assembly and maturation pathway. The U2 snRNA was used here as a representative member of the Sm snRNAs. While discussed in the text, the U bodies are not shown here because of a lack of understanding of their role(s) in snRNP maturation, if any. scaRNP = small Cajal body specific ribonucleoprotein particle.*

## I.16.4 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 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 (Bringmann and Luhrmann 1986; Bindereif and Green 1987; Jamison and Garcia-Blanco 1992; Hong, Bennett et al. 1997; Das and Reed 1999; Kent, Ritchie et al. 2005; Tardiff and Rosbash 2006). The assembly of the spliceosome (E –*early*- complex or commitment complex) begins with the recognition of the 5' ss sequence by the U1 snRNP through its 5'-tail (Rossi, Forne et al. 1996; Will, Rumpler et al. 1996). U1 snRNP-associated proteins U1-70k and U1C stabilize this transient interaction. Another important step following the U1 snRNP-5' ss recognition is the recognition of the 3' ss: the U2 Auxiliary Factor (which is an heterodimer made of U2AF65/35) identifies the AG dinucleotide at the intron/exon junction together with the PPT and SF1/mBBP protein binds at the BP site. Mutual stabilization of contacts with the U2AF bound to the 3' ss and the downstream U1 snRNP at the 5' ss can be mediated by members of the serine/arginine-rich (SR) protein family. The establishment of multiple weak interactions from the 3' ss to the 5' ss defines an exon, and constitutes the commitment step towards the splicing pathway (Robberson, Cote et al. 1990; Berget 1995). Subsequent to E complex formation, the A (prespliceosome) complex is built. The recruitment of U2 snRNP to the BP site, in an ATP-dependent fashion, with the concomitant displacement of SF1/mBBP from the BP site (Hong, Bennett et al. 1997). U2 snRNP base pairing to the BP sequence is facilitated by the U2AF65 subunit bound at the PPT (Ruskin, Zamore et al. 1988). This base-pair interaction is further stabilized by heteromeric complexes of the U2 snRNP, namely SF3a and SF3b (Gozani, Feld 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, Norman 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, Mohammadi 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, Will et al. 2009). When also the second catalytic reaction has occurred, U2, U5 and U6 are released and these snRNPs are recycled for additional rounds of splicing. Spliceosome assembly and function also appear to be regulated by protein kinases and phosphatases. Phosphorylation/desphosphorylation cycles of constitutive and alternative splicing factors have been observed during assembly and catalytic steps (Tazi, Kornstadt et al. 1993; Shi and Manley 2007; Stamm 2008; Heyd and Lynch 2010; Tripathi, Ellis et al. 2010) (Figure 13).
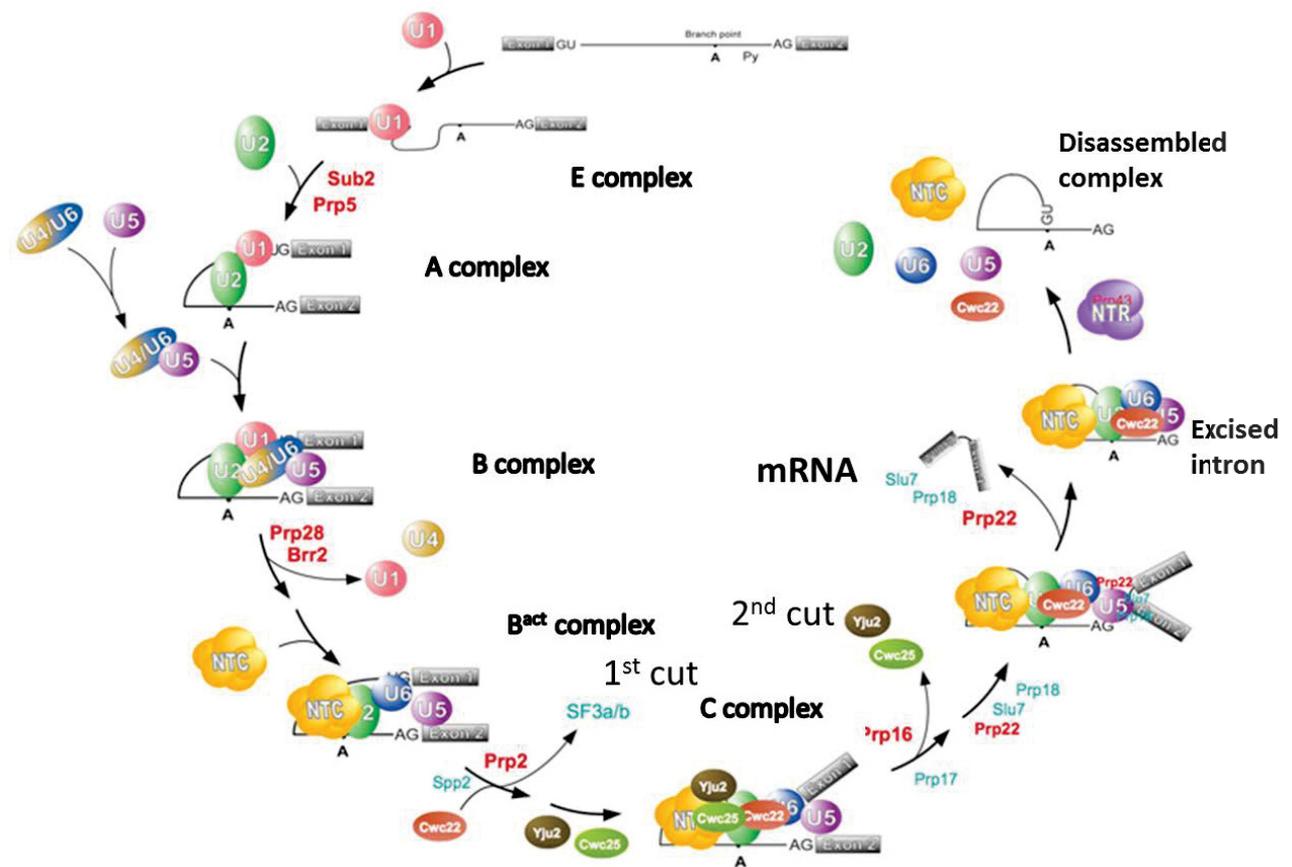


*Figure 13: A schematic view of spliceosome assembly*

## I.16.5 RNA-RNA Interactions in the Spliceosome and the catalytic role of RNA

During spliceosome assembly, an intricate RNA–RNA interaction network is formed that is extensively rearranged during catalytic activation of the spliceosome and the catalytic steps of splicing. At the earliest stages of spliceosome assembly, U1 snRNA base pairs with the 5′ss. U2

snRNA then base pairs with the branch site (BS), forming a short U2-BS duplex in which the branch adenosine is bulged out, specifying its 2′ OH as the nucleophile for the first catalytic step of splicing. Within the U4/U6.U5 tri-snRNP, the U6 and U4 snRNAs are extensively base paired with each other. After association of the tri-snRNP with the A complex, the U4/U6 interaction is disrupted, and the 5′ end of U6 snRNA base pairs with the 5′ss, displacing the U1 snRNA in the process. In addition, an extensive base pairing network is formed between U6 and U2, which juxtaposes the 5'ss and BS for the first step of splicing. Furthermore, a central region of the U6 snRNA forms an intramolecular stem-loop structure (U6-ISL) that appears to play a crucial role in splicing catalysis.

The precise nature of the U6 and U2 snRNA interaction network is the subject of some debate, with two different models currently proposed. In the first, U2 and U6 form three helices (Ia, Ib, and II) , with the conserved U6 triad AGC forming three base pairs with U2 (corresponding to helix Ib) (Madhani and Guthrie 1992). In an alternative model the AGC triad no longer base pairs with U2 but rather with other U6 nucleotides, extending the U6-ISL and allowing for an intramolecular U2 stem-loop, thereby generating a U2–U6 four-way junction (Sun and Manley 1995; Sashital, Cornilescu et al. 2004). Recent data have revealed that U2–U6 interactions appear to be highly dynamic, and these snRNAs likely adopt different conformations at different stages of splicing. Rearrangements are required after step 1 of splicing to reposition the splicing intermediates for the second step of splicing and allow nucleophilic attack of the 5′ exon at the 3′ss (Smith, Query et al. 2008). The precise timing of these changes and the conformation of the RNA–RNA interaction network at this stage is not clear. Recent data have shown that the U6/5′ss interaction must be disrupted before step 2 (Konarska, Vilardell et al. 2006), The U2/BS interaction is not strictly required for the second step and thus it has been proposed that this interaction is also disrupted between the first and second step of splicing (Smith, Query et al. 2007). The conformation of the U2/U6 interaction is also not entirely clear, but likely helix Ia, Ib, and II are formed at this stage (see above). Before step 2, U5 also contacts exon nucleotides just downstream of the 3′ss, and not only tethers the 5′ exon to the spliceosome after step 1, but also aligns both exons for the second catalytic step (Turner, Norman et al. 2004) (Figure 14).

A large body of evidence supports the idea that catalysis of pre-mRNA splicing is at least partially RNA-based with U2 and U6 playing key roles (Valadkhan 2005; Wachtel and Manley 2009). Reactions carried out by UsnRNAs *in vitro* in the absence of protein (Valadkhan and Manley 2001; Valadkhan, Mohammadi et al. 2007; Valadkhan, Mohammadi et al. 2009) are functional but the kinetics and efficiency of these reactions are slow, suggesting that Proteins play a crucial role.
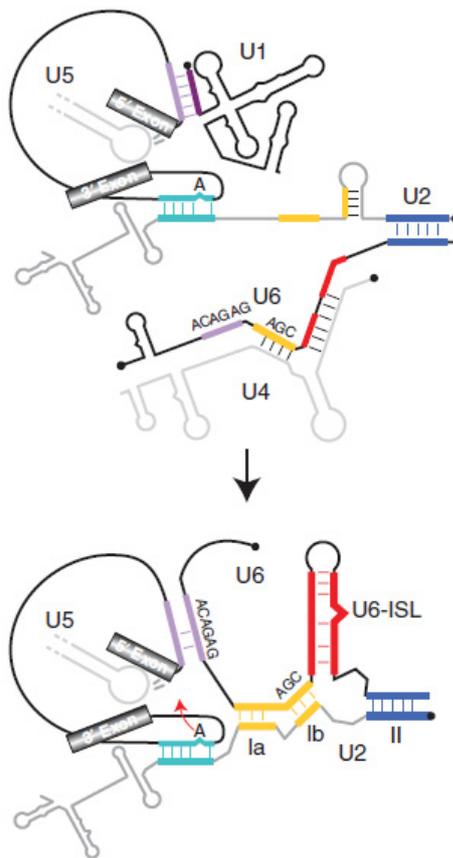
*Figure 14: Dynamic network of RNA-RNA interactions in the spliceosome. Exon sequences are indicated by grey boxes and intron sequences by a thin black line. snRNAs are shown schematically (secondary structure as observed in mammals) in grey or black, with those regions engaging in base pairing interactions (indicated by short lines) highlighted in color (not drawn to scale). The 5' end of the snRNAs is indicated by a black dot. Solely loop 1 of the U5 snRNA is shown. During the transition from a precatalytic spliceosome (upper diagram) to a catalytically activated spliceosome (lower diagram) U1 and U4 are displaced, and U6 and U2 engage in novel base pairing interactions.*

## I.16.6 U1 small nuclear RNA (U1 snRNA)

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

The true genes of U1snRNA, considering the high level of transcription, are found to be not methylated to any appreciable extent, while pseudogenes tend to be more highly methylated than true genes. The human U1 snRNA gene (RNU1-1) is repeated many times in the human genome as clustered repeat units of 45 Kb on chromosome 1 (Lund and Dahlberg 1984; Bernstein, Manser et al. 1985). These genes have a short TATA-less promoter, composed by a distal sequence element (DSE) that serves as a transcription enhancer and a proximal sequence element (PSE), located in the core promoter region upstream from the transcription start site. The PSE sequence is common to all human snRNA genes whereas the absence of a TATA box specifies the recruitment of RNA pol II and its associated transcription apparatus. The PSE is recognized by the snRNA activating protein complex, SNAPc (snRNA activator protein complex)(Sadowski, Henry et al. 1993), which serves as a target for transcription activators and repressors, such as Oct-1 ,p53 and RB (Ford, Strubin et al. 1998; Hirsch, Gu et al. 2000; Gridasova and Henry 2005). SNAPc binding to the U1 PSE is necessary for the recruitment of general transcription factors such as TBP (transcription factor IIB), TFIIA, TFIIB, TFIIE and TFIIF (Kuhlman, Cho et al. 1999)(Figure 15). 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, Murgo et al. 1987; Schaub, Myslinski et al. 1997).
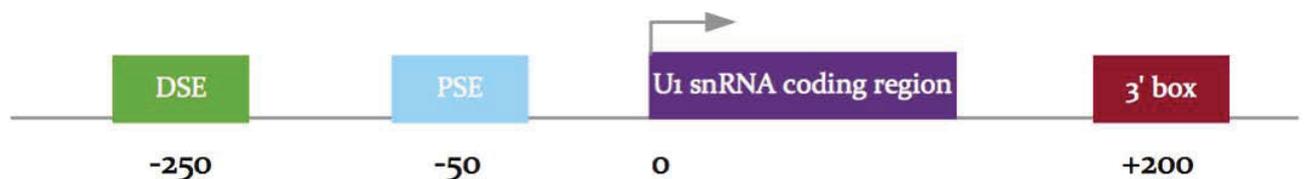


**Figure 15:** *The structure of human U1 snRNA gene transcribed by RNA pol II.*
*The diagram shows the DSE and PSE cis-‐acting promoter elements and the 3'box cis-‐acting RNA‐processing element of pol II-‐transcribed U1-‐snRNA gene boxed, with their position relative to the transcription start site noted below. The start site of transcription is marked with an arrow above the line.*

The transcriptional stimulatory activity of these factors relies on multiple interactions with the general transcription machinery components associated to the SNAPc (Mittal, Cleary et al. 1996; Ford, Strubin et al. 1998). U1 snRNA gene transcript is not spliced and the 3' end is not polyadenilated: probably this feature prevent the association with the translation machinery (Hernandez 2001). Nevertheless the snRNA gene-specific 3'-box (9-19 bp downstream the RNA-encoding region) is required for correct 3'-end formation of U1 snRNA (Egloff, O'Reilly et al. 2008). The 3'-end formation occurs in a step manner: first there is the recognition of the *cis*-acting 3'-box. The 3'-box is a 13-16 nucleotide long element that directs the production of a 3'-extended pre-snRNA which is subsequently processed, leading to a formation of the mature 3'end after transport to the cytoplasm (Huang, Jacobson et al. 1997; Kiss 2004). Moreover, recently

a large complex termed Integrator (constituted of Int11 and Int9) has been shown to play a role in pre-snRNA 3'-end formation (Baillat, Hakimi et al. 2005). Thus, these findings indicate that the 3'-box is an RNA-processing element analogous to the polyadenilation signal commonly found in protein coding genes (Uguen and Murphy 2003; Egloff, O'Reilly et al. 2008). Recent works has demonstrated that the phosphorylation of the CTD (C- terminal domain) of the large subunit of RNA polymerase II is necessary for the 3'-box-dependent RNA 3'-end formation *in vivo*, indicating that processing occurs co-transcriptionally (Medlin, Uguen et al. 2003; Jacobs, Ogiwara et al. 2004). In particular it has been demonstrated that the CTD phosphorylation is fundamental for recruiting Integrator complex which binding is crucial for a correct 3'-end processing (Egloff, O'Reilly et al. 2007).

## I.16.7 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, Dube 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, Oubridge et al. 2009). The association of the two largest U1-specific proteins U1-70K and U1-A depends on the presence of stem-loop I and II, respectively. Deletion or sequence alteration of these stem-loops alter significantly binding of U1-70k and U1-A and may negative consequences on U1 snRNP assembly and function (Surowy, van Santen et al. 1989; Yuo and Weiner 1989; Hamm, Dathan et al. 1990). The smaller U1-C protein is probably attached by protein-protein interaction with the first 97 residues of U1-70k (Nelissen, Will et al. 1994). The characteristic Sm site, to which the Sm proteins bind (sequence *AAUUUGUGG*) is located between stem-loop III and IV (Raker, Hartmuth 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, Pettersson et al. 1983). This interaction occurs via short RNA-RNA interactions between the consensus 5' ss at the exon/intron boundary and the 5'-tail of the U1 snRNP. The role of U1 snRNP in 5' ss recognition has been established long ago through different experimental approaches including

targeting U1 snRNP 5'-tail with RNA oligos or its degradation by ribonucleases, all of which inhibit splicing processing (Mount, Pettersson et al. 1983; Rinke, Appel et al. 1984; Black, Chabot et al. 1985). Furthermore mutations at the 5' ss can be suppressed through the introduction of compensatory changes into the U1 snRNA 5'-tail, further highlighting the role of the U1 snRNP in 5' ss recognition (Zhuang and Weiner 1986). Although 5' ss recognition can occur in absence of a U1 snRNP 5'-tail (Du and Rosbash 2002), the splicing efficiency is compromised. U1C protein contributes to 5' ss reconigition and its stabilizes the RNA duplex between the 5' ss and the U1 snRNP 5'-tail (Du and Rosbash 2002; Pomeranz Krummel, Oubridge et al. 2009), enhancing the formation of E complexes (Will, Rumpler 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, Bach et al. 1990; Will, Rumpler et al. 1996)(Figure 16 ,17 and 18).

In addition to U1C, U1-70K plays an important role in U1 snRNP stabilization and promotion of E complex formation. U1-70k is known to interact with several SR proteins such as ASF/SF2, SC35 and ZRANB2 (Wu and Maniatis 1993; Jamison, Pasman et al. 1995; Cao and Garcia-Blanco 1998; Wang, Lin et al. 1998; Adams, van der Weyden et al. 2001) through its RS domain. These interactions are expected to facilitate U1snRNP recruitment to the 5' ss and subsequent formation of the E complex (Cho, Hoang et al. 2011). The importance of U1-70k behind U1 snRNP activity is further evidenced since it is the only U1 snRNP-associated protein which is regulated by phosphorylation (Tazi, Kornstadt et al. 1993). Interestingly, U1-70k phosphorylation do not impede spliceosome formation but blocks splicing catalysis (Mermoud, Cohen et al. 1994) (Figure 16 ,17 and 18). Even though complementarity between the 5' ss and U1 snRNP 5'-tail is a major determinant in identification of the 5' ss, but base-pairing alone is not sufficient to specify the site of nucleophilic attack (Zhuang and Weiner 1986; Seraphin, Kretzner et al. 1988; Siliciano and Guthrie 1988; Liao, Kretzner et al. 1990). Increased base pairing to a mutant 5' ss can nevertheless maintain usage of cryptic splice sites located nearby (Siliciano and Guthrie 1988; Cohen, Snow 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, Snow et al. 1994). Recently a group of natural 5' ss with a shifted complementarity to U1 snRNP 5'-tail has been described, supporting the notion that U1 snRNP is required for splicing commitment of a 5' ss but do not specify itself 5' ss activation and cleavage site. U5 and U6 snRNP contribute greatly to ensure proper 5' ss selection, as they interact directly with the donor site sequence after U1 snRNP release from the pre-mRNA. In addition, some cases of U1- independent splicing have been reported, in which an abundance of SR proteins compensate the absence of a functional U1 snRNP 5'-tail (Tarn and Steitz 1994). However, the SR proteins themselves cannot discriminate real from cryptic splice sites as the successive interactions

of U1, U5 and U6 snRNPs on the pre-mRNA do (Kandels-Lewis and Seraphin 1993; Lesser and Guthrie 1993; Tarn and Steitz 1994).



***Figure 16:*** *Schematic structure of U1 snRNP*
*The secondary structure of the U1 snRNA and the sites at which the proteins U1-70K and U1-A bind the RNA are represented. The Sm site is represented with the Sm proteins. U1-70k physically interacts with U1-C and Sm-D1/D2.*



***Figure 17:*** *Secondary and 3D structural models of the U1 snRNA.*
*The four stem-loops, four-way junction and Sm site are color coded.*

*Figure 18: Model of the 3D arrangement of the RNA and proteins in the human U1 snRNP.*
*A surface representation of the Sm protein ring was fitted into the 3D structure.*
*The precise positions of stem-loops I and II of the U1 snRNA are not known*

## I.16.8 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 etiology (Baralle, Lucassen et al. 2009; Tazi, Bakkour 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 and Baralle 2004; Buratti, Baralle 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. 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, Wei et al. 1998; Ohno, Tsujino et al. 2001; Aznarez, Chan et al. 2003). Thus the primary mechanism of disease behind most pathological exonic mutations is a catastrophic splicing abnormality rather than a direct effect on coding potential (Lopez-Bigas, Audit et al. 2005). Nevertheless, the clear correlation between the suspected mutation and the disease is hard to be shown (Baralle and Baralle 2005; Wang and Cooper 2007). As research has progressed, it has become clear that genomic

variants - even if found in intronic region - should be considered as a potential disease-causing mutation affecting splicing (Baralle, Lucassen 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, Kornblihtt et al. 1984) has allowed a relatively fast approach to identify splicing spoilers and to study their underlying functional mechanism (Baralle and Baralle 2005; Cooper 2005).


## I.16.9 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 generally 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 and frequently abolish splicing (Krawczak, Thomas 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 genetic variant (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 genetic variants tend to cluster at certain positions, particularly -1 and spanning positions +3 to +6 for donor sites and -3 for acceptor sites (Krawczak, Thomas et al. 2007). Pathological polypyrimidine tract mutations are less abundant than those affecting splice sites and show a homogeneous distribution (Krawczak, Thomas 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 (Krawczak, Reiss et al. 1992; Pohlenz, Dumitrescu et al. 2002; Krawczak, Thomas et al. 2007). 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 spliceosome assembly on nascent pre-mRNA. This results usually in exon skipping (Krawczak, Thomas et al. 2007), due to the utilization of the downstream donor splice site and the upstream acceptor site, 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 a weak splice site recognition caused by a genetic mutation leads to a decreased exon inclusion, yet some traces of fully and functional mRNA is produced, with reduced levels of functional protein. 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 (Hartmann, Theiss et al. 2008; Houdayer, Dehainault et al. 2008; Spurdle, Couch et al. 2008; Tournier, Vezain 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,SMN and FVII deficiency (Pagani and Baralle 2004; Pinotti, Rizzotto et al. 2008; Pinotti, Balestra et al. 2009; Pinotti, Bernardi et al. 2011). At the same time, this method allows to discover novel regulatory elements, either intronic or exonic, which may participate in splice site selection (Cooper 2005).

## I.16.10 Corrective therapies for splicing defects

*U snRNP-based therapies for splicing modulation*. The use of modified U snRNPs as therapeutic molecules represents an attractive tool for splicing diseases for several reasons. First, U snRNAs are naturally imported to the nucleus to target pre-mRNAs, ensuring higher efficient delivery respect to other strategies (i.e., antisense oligonucleotide). Second, delivery of U snRNAs in a proper recombinant vector (i.e., the widely used AAV) would require few, if not only one, administrations to prospective patients. Third, the risk of degradation of administered U snRNA is very low, because the injected molecule mirrors the endogenous one.

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 (Muller and Schumperli 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, Stefanovic et al. 1993; Stefanovic, Hackl et al. 1995). The natural ant-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, Suter et al. 1998; Goyenvalle, Vulin et al. 2004; Madocsai, Lim et al. 2005; Asparuhova, Marti et al. 2007; Uchikawa, Fujii 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, Rosa et al. 2006). More recently, investigators used engineered U1 snRNA to correct splicing through induction of exon skipping in human *DMD* pre-mRNA in primary patient fibroblasts. The engineered U1 snRNA used to induce exon skipping of exon 51 bore antisense sequence replacing the its natural 5' tail. In that way, the mutated U1 snRNA behaved like an antisense oligonucleotide. Similarly, U7 SmOPT snRNA, containing antisense sequence targeting the SMN2 intron 7/exon 8 junction, was able to induce exon 7 inclusion in up to 80% of transcripts in SMA patient-derived fibroblasts (Geib and Hertel 2009). The use of modified U1 snRNAs suppress the impact of disease-associated splice site mutations, complementing the loss of the normal U1 snRNP activity due to its inability to recognize and base pair with the mutated site. 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. That approach bore by the previous observations that compensatory changes in U1 snRNAs were able to modulate 5' splice site recognition mutants in Drosophila, adenovirus E1a and albumin gene (Hitomi, Sugiyama et al. 1998). These approaches have been recently proposed for the correction of splicing defects in human malignant infantile osteoporosis (arOP) (Susani, Pangrazio et al. 2004), retinitis pigmentosa (Tanner, Glaus et al. 2009; Glaus, Schmid et al. 2011), Fanconi anemia (Hartmann, Neveling et al. 2010) and Bardet-Biedl syndrome (Schmid, Glaus et al. 2011).

# II. Material and Methods

## 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.

## 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.

## Preparation of bacterial competent cells

Bacterial chemical competent cells were prepared as follow: 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 OD600 was 0.3-0.38. The cells were transferred into a Falcon 50ml tube and then placed in ice and centrifuged at 4 °C and 1000g 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 Cl2, 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.

## Plasmid DNA purification

Rapid isolation of plasmid DNA from 5 ml of E. coli culture grown over night was achieved by using the Wizard® Plus Minipreps DNA Purification System (Promega Corporation, Madison, WI). Bacterial culture was harvested by centrifugation for 5 min at 10,000 x g, and the pellet resuspended in 250 µl of Cell Resuspension Solution. Cell Lysis Solution (250 µl) was then added and the tube was inverted 4 times. After adding 10 µl of Alkaline Protease Solution and inverting the tube 4 times, the sample was let incubate at room temperature for 5 min. Neutralizing Solution (350 µl) was then added and the tube was inverted 4 times before centrifuging the bacterial lysate at 14,000 x g for 10 min at room temperature. The plasmid DNA purification unit was prepared by inserting the Spin Column into a 2 ml Collection Tube. After centrifugation, the cleared lysate was transferred to the prepared Spin Column and centrifuged at 14,000 x g for 1 min at room temperature. DNA bound to the Spin Column was then washed with 750 µl of Column Wash Solution and newly centrifuged at 14,000 x g for 1 min at room temperature. The same procedure was repeated using 250 µl of Column Wash Solution. Finally DNA was eluted from the Column into a new tube adding 100 µl of Nuclease-Free Water to the Spin Column and by centrifuging the system at 14,000 x g for 1 min at room temperature.

High-quality DNA for transfection experiments was obtained from 70-100 ml of E. coli culture grown over night by using the GenElute™ HP Plasmid Midiprep Kit (Sigma-Aldrich, St. Louis, MO).

## Genomic DNA extraction from tissue

For the purification of total DNA from animal tissues we used the DNeasy Blood & Tissue Kit (QIAGEN). Basically samples are first lysed using proteinase K. Buffering conditions are adjusted to provide optimal DNA binding conditions and the lysate is loaded onto the DNeasy Mini spin column. During centrifugation, DNA is selectively bound to the DNeasy membrane as contaminants pass through. Remaining contaminants and enzyme inhibitors are removed in two efficient wash steps and DNA is then eluted in water or buffer, ready for use. Concentration and purity of the extracted DNA was evaluated by reading the absorbance of the DNA suspension at

260 nm (DNA) and 280 nm (proteins) with an Ultrospec 2000 spectrophotometer (Pharmacia Biotech, Uppsala, Sweden). An OD260/OD280 ratio equal or greater than 1.8 indicates pure DNA. A rough estimate of the quality and quantity of the DNA was also obtained by running a few microliters on a 0.8% agarose gel. DNA samples were stored at –20°C.

## Restriction analysis

Restriction enzymes are endonucleases of bacterial origin able to recognize specific nucleotide sequences (recognition sites) and cleave both strands of the DNA containing those sequences. Single-nucleotide substitutions can create/disrupt a restriction endonuclease recognition site, making it possible to distinguish the two alleles by restriction analysis of suitable PCR products spanning the substitution.

Restriction enzymes were purchased from New England BioLabs, Inc. and Fermentas AB, Vilnius, Lithuania. Amplified DNA (100 ng) was incubated with 3-4 units restriction enzyme in the proper buffer for 1 hour or longer at the recommended temperature. The products of DNA digestion were then analyzed by gel electrophoresis.

| Enzyme | Target sequence | Working temperature |
|--------|-----------------|---------------------|
| ClaI | AT\|CGAT | 37°C |
| SacII | CCGC\|GG | 37°C |
| XbaI | T\|CTAGA | 37°C |
| BclI | T\|GATCA | 50°C |
| BglII | A\|GATCT | 37°C |
| EcoRI | G\|AATTC | 37°C |
| XhoI | C\|TCGAG | 37°C |

## Direct sequencing

Sequencing was performed with the dideoxy-mediated chain termination method.

*PCR fragment purification.* PCR products were purified from contaminating primers, free nucleotides, *Taq* polymerase and salts using the Genomics Millipore filters (Bedford, MA, USA), which is based on the property of glass fibers to specifically bind nucleic acids. 50µl PCR reaction were mixed well with 350 µl distilled water and spun at 1000×g for 15 sec. Subsequently, the filter

tube was inverted, applied to a clean collection tube and filled with 20 µl distilled water, before spinning again at 1000×g for 2 min.

*Cycle sequencing.* The sequencing reactions were performed using the BigDye™ Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA polymerase, FS (PE Applied Biosystems, Foster City, CA, USA). This kit makes use of a genetically modified *Taq* polymerase that does not discriminate di-deoxynucleotides and includes a set of terminators labeled with high-sensitivity dyes. Reaction mixture (10 µl) was prepared by mixing 3.5 µl purified PCR product (corresponding to 30-90 ng, template DNA), with 0.5 µl of 6.8 µM primer, 3.0 µl of Terminator Ready Reaction Mix from the kit (comprising A-dye, C-dye, G-dye, T-dye, deoxynucleoside triphosphates, the thermostable DNA polymerase, $MgCl_2$ and Tris-HCl buffer, pH 9.0) and 3.0 µl of distilled water. Cycle sequencing was carried out by subjecting this mixture to 45 cycles of denaturation (10 sec at 96 °C), annealing (5 sec at 45-60 °C, according to the primer used), extension (4 min at 60 °C).

*Sample preparation.* Prior to electrophoresis, extension products were subjected to ethanol/sodium acetate purification to remove unincorporated dye-labeled terminators. 2 µl of 3M sodium acetate (pH 4.6) and 50 µl of 95% ethanol were added to the sequencing reaction (10 µl) and mixed well. Following centrifugation at 16000 rpm for 30 min, the supernatant was discarded and 250 µl of ice-cold 70% ethanol were added to wash the precipitate, which was vortexed gently and spun again at 18000 rpm for 5 min. After discarding the supernatant, the pellet was air-dried, resuspended in 15 µl of highly deionized formamide (Hi-Di™ Formamide, Applied Biosystems, Foster City, CA) and denatured at 95 °C for 3 min before loading on the gel.

## Gel electrophoresis

Due to their numerous phosphate groups, nucleic acids are negatively charged at neutral pH and tend to migrate towards the anode if subjected to an electric field. Their migration rate is inversely proportional to the logarithm of their length in bp. These properties make it possible to separate DNA fragments according to their size. At the end of the electrophoretic run, the positions of the DNA fragments in the gel are visualized by ethidium bromide, an intercalating dye that fluoresces when bound to DNA.

Agarose gels (able to separate fragments ranging from 200 bp to 50 kb) were prepared by dissolving the desired amount of agarose in 1× TAE buffer (40 mM Tris-acetate, 1 mM EDTA) and

heating this mixture in a microwave oven till complete clarification. A 1% agarose gel contains 1 g agarose in 100 ml buffer. Ethidium bromide was added directly to the melted gel before casting, in the proportion of 5 µl of a 10 mg/ml stock to 100 ml gel. Agarose gels were run horizontally in 1× TAE buffer, by applying a voltage of 5 V/cm.Polyacrylamide gels (able to resolve fragments ranging from 5 bp to 500 bp) were prepared by mixing the desired volume of an acrylamide stock (40% 1/19 N,N'-methylenebisacrylamide/acrylamide) in 1× TBE buffer (90 mM Tris-borate, 2 mM EDTA). Immediately before pouring the gel, appropriate amounts of ammonium persulfate (0,66% of the final gel volume of a 10% solution) and TEMED (0,066% of the final gel volume) were added to allow matrix polymerization (about 1 hour). Polyacrylamide gels were run vertically in 1× TBE buffer, by applying a voltage of 25 V/cm. At the end of the run, the gel was recovered and stained by soaking it in 1× TBE buffer containing 2 µg/ml ethidium bromide for about 15 min.Stained gels were viewed under UV transillumination at 254 nm, the picture was imported with a GelDoc 1000 UV-gel camera (Bio-Rad Laboratories, Hercules, CA, USA) and stored on a computer as an image file. Gel images were manipulated with the software Molecular Analyst™, version 1.3 (Bio-Rad Laboratories, Hercules, CA, USA).

## Polymerase chain reaction (PCR)

DNA amplification by PCR (Mullis and Faloona 1987) takes advantage of a natural enzyme (the thermostable DNA-polymerase I of the thermophilic bacterium *Thermus aquaticus*, *Taq* polymerase) to produce a large number of copies of the same DNA fragment. The reaction requires a template (usually purified DNA), two primers (single-stranded oligonucleotides that frame the target sequence) and all four deoxyribonucleoside 5'-triphosphates (dNTPs) in the presence of $MgCl_2$.

The amplification reaction is performed via 25-30 DNA replication cycles, each comprising three steps: 1) denaturation (separation of the two strands of template DNA); 2) annealing (hybridization of the primers to their complementary sequences on the template); 3) extension (elongation of the primers). These steps require different temperatures (denaturation: 95 °C; annealing: 45-60 °C, according to the characteristics of the primers; extension: 72 °C).

Primers for the PCR-amplification of all exons and splicing junctions of the FVII gene (Table 2.3.2) were designed on the basis of the published FVII gene sequence using the OLIGO 4.1 Primer Analysis Software. Amplification reactions (25-50 µl total volume) were carried out using 1 unit of *Taq* polymerase (BioTherm™ Taq DNA Polymerase, eEnzyme LLC, Gaithersburg, MD) in the buffer provided by the supplier. Reaction conditions were as follows: ~100 ng template DNA, 272

nM each primer, 50 nM each nucleotide precursor, 1.5-2.0 mM $MgCl_2$ and, in some cases, 1-10% of Dimethylsulphoxide (DMSO) to improve amplification efficiency and specificity. A negative control (*i.e.* a reaction carried out in the absence of template DNA) was always included to check for reagent contamination with template DNA. Thermal cycles comprised 5 min of initial denaturation at 95 °C, 10 min hot start at 65 °C (during which the enzyme was added to the reaction mixtures), 30 cycles of denaturation, annealing, extension as previously described and 10 min of final extension at 72 °C. All PCR reactions were performed with a GeneAmp® PCR System 2400 thermal cycler (Applied Biosystems, Foster City, CA). The qualitative and quantitative outcome of the amplification reaction was checked by running 3-5 µl of PCR product on agarose gel in parallel to an appropriate molecular weight marker.

*Oligonucleotides used in PCRs or RT-PCRs*

| Name | Sequence (5' →3') |
| --- | --- |
| 6F | AAACCCCAAGGCCGAATT |
| 7F[Fam] | cattcaAGGTCCTGTTGTTGGTGAAtG |
| 7F | ACCCTGATCAACACCATCTGG |
| 8R' | GCGATGTCGTGGTTGGTGGT |
| 8R | GCCCTCTAGATGCATGCTCGAGCGG |
| 6Fbis | GCATCTTTCTGACTTTTGTT |
| 8R XbaI | ATATTACTCTAGAACAGGCCAGGGCTGCTG |
| 8Rbis | ATCCGAGTAGCCGGCACAGAACATGTACTC |
| hFVIIex6F | TAGAAAAAAGAAATGCCAGCAAACC |
| hFVIIex7F | GTCCTGTTGTTGGTGAATGGAGCTCA |
| hFVIIR8 | GCTGACCAATGAGAAGCGCACGAA |
| FIXex5 F | CATATGGTTATACATTAATAAATAG |
| FIXex5 R | CATATGCAGAAATCACACAAATTAATTGC |
| Alfa2-3 | CAACTTCAACTCCTAAGCCACTGC |
| Bra2 | GTCACCAGGAAGTTGGTTAAATCA |
| pBsK-FIX F | GGTAAATTGGAAGAGTTTGT |
| pBsK-FIX R | ATTAACGATAGAGCCTCCAC |

## RNA extraction from tissues and cells

RNA was isolated from cells and animal tissues exploiting Trizol reagent (Invitrogen Corporation, Carlsbad, CA). Withdrew tissues were wrapped in aluminum foil and fast cooled through liquid nitrogen. For long term cryopreservation samples were stored to -80°C freezer. Trizol, a mono-

phasic solution of phenol and guanidine isothiocyanate, during sample homogenization or lysis, maintains the integrity of the RNA, while disrupting cells, RNase and dissolving cell components. Addition of chloroform followed by centrifugation, separates the solution into an aqueous phase and an organic phase. RNA remains exclusively in the aqueous phase. After transfer of the aqueous phase, the RNA is recovered by precipitation with isopropyl alcohol. The pellet resulting from the centrifugation is washed once with 75% ethanol and the RNA pellet is briefly air-dried and dissolved in sterile water containing an RNase inhibitor. The sample is stored to -20°C for short term storage, at -80°C for long term storage.

## Reverse transcription of RNA and amplification (RT-PCR)

Reverse-transcriptase polymerase chain reaction (RT-PCR) was carried out by using the SuperScript™ One-Step RT-PCR with Platinum® Taq kit (Invitrogen Corporation, Carlsbad, CA), designed for the coupled reverse transcription (RT) and polymerase chain reaction (PCR) amplification of RNA from either total RNA or mRNA. The system uses a modified version of Moloney Murine Leukemia Virus (M-MLV) RT, engineered to reduce RNase H activity and increase thermal stability, for first strand cDNA synthesis and the thermostable Taq DNA polymerase from *Thermophilus aquaticus* for second strand cDNA synthesis and DNA amplification. The Taq DNA polymerase in the kit is a Taq DNA polymerase complexed with a proprietary antibody that inhibits polymerase activity at ambient temperatures. The antibody is denatured and polymerase activity is restored during the denaturation step in PCR cycling at 94°C. This provides an automatic "hot start" in PCR, increasing sensitivity, specificity, and yield. Briefly, 2 µl RNA were mixed with 25 picomoles (each) primers in 2X buffer (a buffer containing 0.4 mM of each dNTP, 2.4 mM MgSO4) and 1 µl of RT/Taq mix at a final volume of 25 µl.

Reverse transcription reaction was performed at 45°C for 45 minutes. Inactivation of the M-MLV RT and denaturation of the RNA/cDNA hybrid was carried out by incubation at 94 °C for 2 minutes and followed by 40 thermal cycles (denaturation: 30 sec at 94 °C, annealing: 1 min at 60 °C, extension: 2 min at 68°C) and by 7 min of final extension at 68 °C. All RT-PCR were performed with the GeneAmp® PCR System 2400 thermal cycler. Reverse transcription and amplification reactions were performed with the primers indicated in Table. The outcome of the RT-PCR was checked by running 5 µl of product on agarose gel in parallel to a molecular weight marker.

## Construction of plasmids:

### *pFVII-wt and pFVII IVS7+5g/a*

The genomic area of human Factor VII gene spanning from exon 6 through 8 was PCR-amplified from genomic DNA of normal subject (carryng six IVS7 repeats as the mutant allele) and from a IVS7+5 g/a homozygote using oligonucleotides 6Fbis and 8Rbis. PCR was performed as previously described on 100 ng of genomic DNA in a standard 50 µL volume on a GeneAmp® PCR System 2400 thermal cycler. PCR amplification products were purified using the Genomics Millipore filters and TA cloned into a pGEM-T Easy vector (Promega Corp., Madison, WI). Positive white clones were propagated and DNA was extracted using the Wizard® Plus Minipreps DNA Purification System. The region of interest was excided from pGEM-T Easy vector by EcoRI sites and cloned into the mammalian expression vector pcDNA3 (Invitrogen Corporation, Carlsbad, CA) previously cut with the same enzyme. The two plasmids (pFVII-wt and pFVII IVS7+5a) were checked by sequencing.
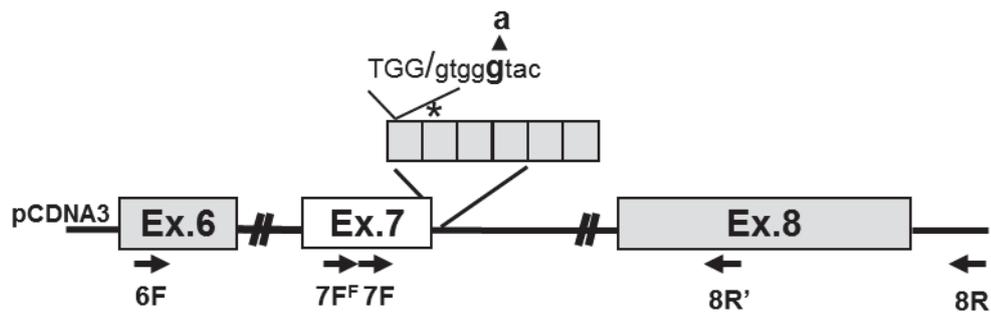


***Figure 19: Schematic representation of pFVII plasmid***
*The splicing cassette contains the FVII genomic region spanning exon 6 to exon 8. Minisatellite repeats are represented by greu squares. Asterisk indicates the cryptic donor splice site predicted to be used by the bioinformatics tool. The exon 7-intron 7 boundary is shown in the upper part of the picture. 3' exon 7 end is represented in bold.*

### *pSCFVII-wt and pSCFVII IVS7+5g/a*

To create a full-length splicing-competent constructs, the FVII gene region spanning exon 6 through 8 ( from nucleotides 8926 to 11157) from previous plasmids pFVII-wt and pFVII IVS7+5g/a was amplified using the primers 6Fbis and 8R XbaI. The amplified fragment was inserted into the coding sequence of FVII (cDNA) cloned into the pcDNA3 plasmid using the XbaI sites located in exon 6, downstream the FVII coding region spanning exon 1 to 6, and in pcDNA3 plasmid. Both plasmids were checked by sequencing.
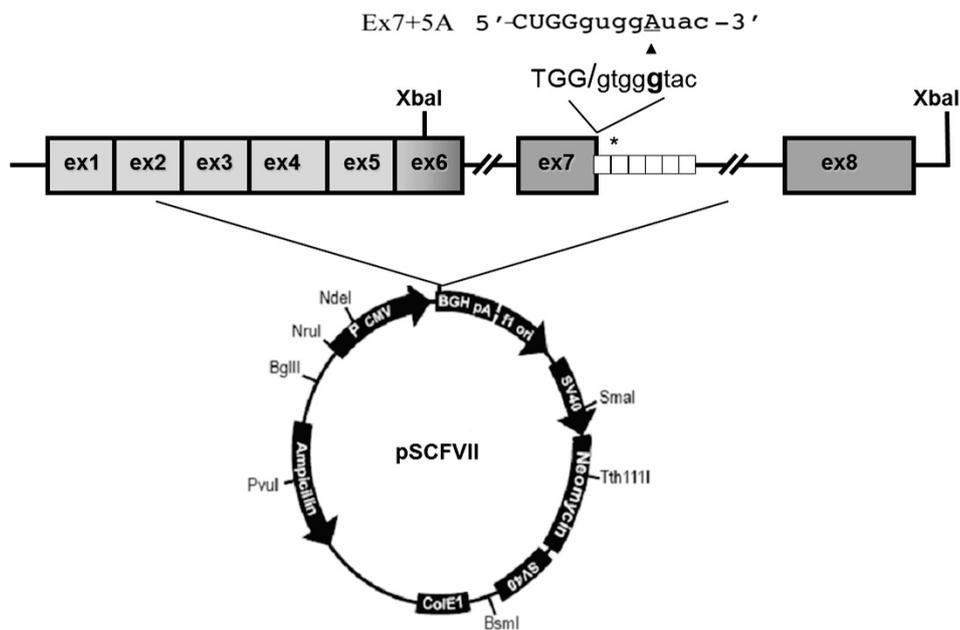
**Figure 20: Schematic representation of pSCFVII plasmid**

*It contains at the 5' end CMV gene promoter to allow polymerase II transcription in the transfected cell lines. At the 3' end a functionally competent polyadenylation site, derived from bovine growth hormone gene is present. The full-length splicing competent cassette contains the coding region (cDNA) of FVII from exon 1 to exon 6. Intron 6 and 7 are inserted through amplification of the region spanning exon 6 through exon 8 and cloned into the cDNA using a unique XbaI site located into exon 6.*

### pFIX ex5

It comes from the modification of the pTB plasmid. The pTB minigene, containing exons from α-globin and fibronectin, under the control of the α-globin promoter and SV40 enhancer. The intronic region between the two fibronectin exons contains a unique NdeI 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.

To obtain the pFIX exon 5 minigenes, the FIX exon 5 fragment consisting of the last 314 bp of intron 4, exon 5 (129 bp) and the first 278 bp of intron 5 was amplified from normal genomic DNA using FIXex5dir and FIXex5rev and cloned into the pTB NdeI – minigene.
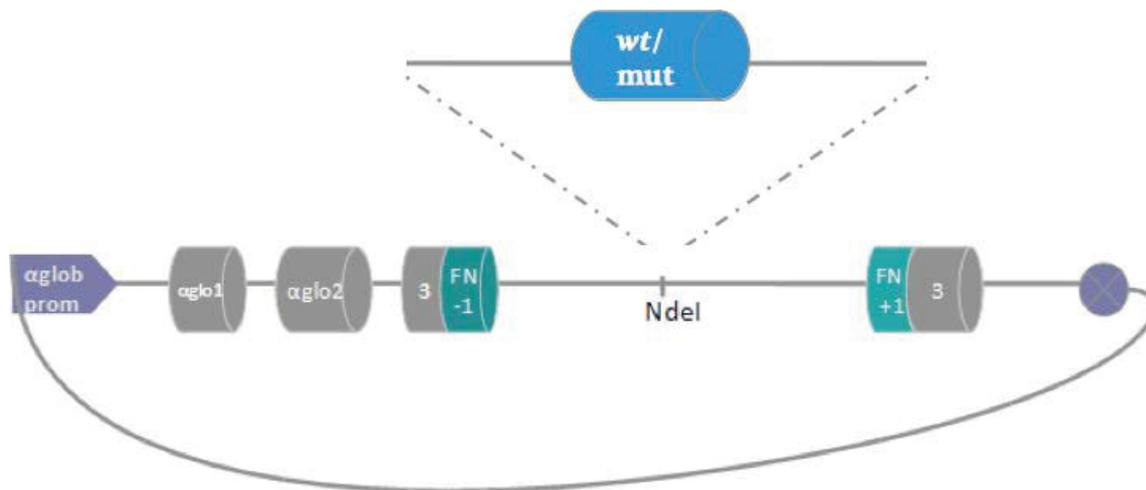
***Figure 21: Schematic representation of pTB NdeI - minigene***

*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 NdeI is used for subcloning of wt or mutant exon and flanking intronic regions.*

### pSC-FIX

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 BglII and ApaI restriction sites. It carries a simian virus 40 promoter, factor IX cDNA sequence from exons 1 to 4, first 544 bp of intron 4, NdeI restriction site, last 312 bp of intron4, exon 5 (129 bp), first 287 bp of intron 5, NdeI 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 NdeI-NdeI cassettes were taken from pFIX exon 5 minigenes and

subcloned into pSC-FIX create wt and mutant factor IX expression vectors.

### U1 construction

The vector used for the production of U1 snRNAs was the pGEM3, a standard cloning vector. The wt U1 snRNA gene was cloned in BamHI restriction site (Figure 22) and the region between BglII and BclI 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 the Table below. The identity of all modified U1 snRNAs expression vectors was ultimately confirmed through sequencing analysis.
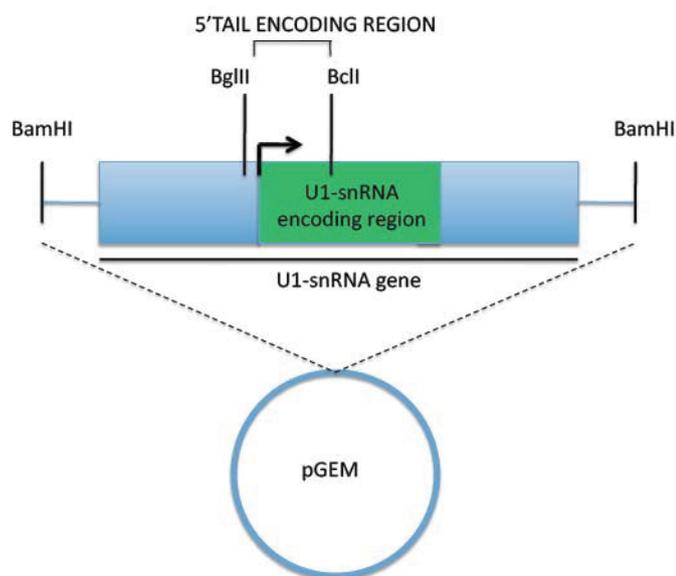
***Figure 22***: *Schematic representation of cloning strategy for creation of modified-U1 snRNAs. The wild-type U1 snRNA gene was cloned in BamHI restriction site located in the multiple cloning site of pGEM3 vector. The region between BglII and BclI sites encoding for the 5'-tail of the U1 was replaced by specific annealed oligos for create engineered U1 snRNAs able to recognize specific target sequences with their modified 5'-tails.*

| Name | Sequence (5' →3') |
|---|---|
| **FVII U1snRNAs** | |
| U1Mut1 F | GATCTCATCAGCACCGCGCAGGGGAGATACCAT |
| U1Mut1 R | GATCATGGTATCTCCCCTGCGCGGTGCTGATGA |
| U1Mut2 F | GATCTCATTCGGACAGGGCAGGGGAGATACCAT |
| U1Mut2 R | GATCATGGTATCTCCCCTGCCCTGTCCGAATGA |
| U1+5a F | GATCTCAATATCTACCTGGCAGGGGAGATACCAT |
| U1+5a R | GATCATGGTATCTCCCCTGCCAGGTAGATATGA |
| **FIX U1snRNAs** | |
| ex5 wt R | GATCATGGTATCTCCCCTGCCAGGTCATATATGA |
| ex5 wt F | GATCTCATATATGACCTGGCAGGGGAGATACCAT |
| ex5 -2C R | GATCATGGTATCTCCCCTGCCCGGTCATATATGA |
| ex5 -2C F | GATCTCATATATGACCGGGCAGGGGAGATACCAT |
| ex5 -2G R | GATCATGGTATCTCCCCTGCCGGGTCATATATGA |
| ex5 -2G F | GATCTCATATATGACCGCGCAGGGGAGATACCAT |
| ex5 -2T R | GATCATGGTATCTCCCCTGCCTGGTCATATATGA |
| ex5 -2T F | GATCTCATATATGACCGAGCAGGGGAGATACCAT |
| ex5 -1T R | GATCATGGTATCTCCCCTGCCATGTCATATATGA |
| ex5 -1T F | GATCTCATATATGACATGGCAGGGGAGATACCAT |
| ex5 +1A R | GATCATGGTATCTCCCCTGCCAGATCATATATGA |
| ex5 +1A F | GATCTCATATATGATCTGGCAGGGGAGATACCAT |
| ex5 +2C R | GATCATGGTATCTCCCCTGCCAGGCCATATATGA |

68

| | |
|---|---|
| ex5 +2C F | GATCTCATATATGGCCTGGCAGGGGAGATACCAT |
| fix 1 R | GATCATGGTATCTCCCCTGCGTCATAATCTATGA |
| fix 1 F | GATCTCATAGATTATGACGCAGGGGAGATACCAT |
| fix 7 R | GATCATGGTATCTCCCCTGCATCTGAATAAGATGA |
| fix 7 F | GATCTCATCTTATTCAGATGCAGGGGAGATACCAT |
| fix 9 R | GATCATGGTATCTCCCCTGCCTGAATAAGAATGA |
| fix 9 F | GATCTCATTCTTATTCAGGCAGGGGAGATACCAT |
| fix 10 R | GATCATGGTATCTCCCCTGCTGAATAAGATATGA |
| fix 10 F | GATCTCATATCTTATTCAGCAGGGGAGATACCAT |
| fix 13 R | GATCATGGTATCTCCCCTGCATAAGATTTTATGA |
| fix 13 F | GATCTCATAAAATCTTATGCAGGGGAGATACCAT |
| fix 16 R | GATCATGGTATCTCCCCTGCAGATTTTTTATATGA |
| fix 16 F | GATCTCATATAAAAAATCTGCAGGGGAGATACCAT |
| fix 22 R | GATCATGGTATCTCCCCTGCTTTAAAGAAATATGA |
| fix 22 F | GATCTCATATTTCTTTAAAGCAGGGGAGATACCAT |
| fix 33 R | GATCATGGTATCTCCCCTGCTCTGTATCTGAATGA |
| fix 33 F | GATCTCATTCAGATACAGAGCAGGGGAGATACCAT |
| fix 38 R | GATCATGGTATCTCCCCTGCATCTGAAACTATGA |
| fix 38 F | GATCTCATAGTTTCAGATGCAGGGGAGATACCAT |
| fix 63 R | GATCATGGTATCTCCCCTGCACCTACATAAATGA |
| fix 63 F | GATCTCATTTATGTAGGTGCAGGGGAGATACCAT |

*Table 2: List of oligonucleotides used for construction of modified U1 snRNAs*

### *pAAV-hAAT-FVII-wt, pAAV-hAAT-FVII IVS7+5g/a and pAAV-U1+5a*

In laboratory pAAV-CMV and pAAV-hAAT were just available (gifted by Prof. Arruda V, Children's Hospital of Philadelphia, PA, USA). The entire full-length splicing-competent constructs from pSCFVII-wt and pSCFVII IVS7+5a plasmids were isolated using BamHI sites, located at the extremities of the inserts, and ligated in the previously cut pAAV-CMV to create pAAV-CMV-FVII (wt and IVS7+5a). Those plasmids were cut with ClaI and XhoI restriction enzymes to isolate the inserts with new extremities that were successively inserted into ClaI-XhoI previously cut pAAV-hAAT plasmid. hAAT promoter, previously isolated from pAAV-hAAT plasmid through ClaI digestion, was inserted upstream the cloned construct using the ClaI site to generate the full length construct.

To generate pAAV-U1+5a plasmid, the cassette driving the expression of the modified U1(containing the endogenous promoter of U1 gene) was isolated from pU1+5a using BamHI restriction enzyme and inserted into BamHI digested pAAV-CMV plasmid (BamHI digestion removes the CMV promoter from plasmid, leaving the empty backbone).

## Transfection protocol

Hep3B (human hepatoma) and COS-1 (African green monkey kidney fibroblasts) cells were transiently transfected into 30-mm plates with 3 μg of each minigene vector and, in complementation assays, with 0.5×, 1×, 1.5×, 2×, and 2.5× molar excess of pU1-snRNA vectors by exploiting Lipofectamine 2000 (Invitrogen Corporation, Carlsbad, CA). Basically in an eppendorf tube, 3 μl (in ration 1:1 to the amount of DNA transfected) of Lipofectamine were added to 150 μl of serum-free medium (GIBCO-BRL™ OptiMEM®, Gaithersburg, MD) and incubated for 5 min at room temperature. DNA vector (3 μg) was placed in a separate tube with other 150 μl of serum-free medium. The solutions were joined after 5 minutes. The mixture was incubated at room temperature for 20 min and then added dropwise to the cells. The medium was changed 4 hours after transfection to reduce toxicity effects of Lipofectamine. Free FBS- Culture medium (OptiMEM; Invitrogen Corporation, Carlsbad, CA) was supplemented with 5 μg/mL vitamin K (Konakion; Roche, Welwyn Garden City, United Kingdom) to allow proper FVII biosynthesis. Lysis of cells and collection of media, for mRNA and protein studies, were conducted 72 hours after transfection.

## Computational analyses

Wild type and mutated sequences of Factor VII gene were analyzed to search for donor splice sites using the NNSPLICE 0.9 splice site predictor program, available at the Berekeley Drosophila Genome Project page (http://www.fruitfly.org/seq_tools/splice.html). The software is a neural network based program that scans for putative splice sites and gives them a splice site score (SSS). The best score is 1.00.

## FVII IHC

Immunohistochemistry or IHC refers to the process of detecting antigens (e.g., proteins) in cells of a tissue section by exploiting the principle of antibodies binding specifically to antigens in biological tissues. IHC takes its name from the roots "immuno," in reference to antibodies used in the procedure, and "histo," meaning tissue (compare to immunocytochemistry). Immunohistochemical staining is widely used in the diagnosis of abnormal cells such as those found in cancerous tumors or to specifically detect cells that are secreting a target protein.

The detection of the human FVII expression in mouse tissue is an hard goal to achieve. In fact, being the homology between human and mouse proteins very high, the quality of that assay mainly depends on the ability of the antibodies to specifically bind the human antigen. Furthermore, the king of blocking reagent, used to avoid aspecific bindings, is crucial for that assay. To reduce the possibility of aspecific interactions, we modified the classical protocol of immunohistochemistry. Briefly, a specifically designed anti mouse blocking buffer and a single recognition-detection step were used to reduce not specific interactions. 8um thickness sections ( fixed with 4%PFA in 1XPBS) were incubated with blocking buffer (10% sheep serum, 0,1% Triton X100, 100mM glycine in 1X PBS) for 2 hours at room temperature. Sections were rinsed 3 times in 1X PBS for 5 minutes each and further incubated with 0,1mg/ml unconjugated Affinipure Fab fragment Anti-Mouse IgG solution H+L (diluted in 1% BSA and 1X PBS) (Jackson ImmunoResearch laboratories, Baltimore, PA) for 1 hour at room temperature. Sections were rinsed again 3 times with 1X PBS for 10 minutes each and further incubated overnight at 4°C with  a polyclonal sheep anti-human FVII FITC conjugated antibody (Affinity Biologicals, Ancaster, Canada) diluted to 2ug/ml in 1% sheep serum/1X PBS. The day after, sections were washed again with PBS to remove unbound antibodies (3 times for 5 minutes each) and fixed with 4%PFA/PBS to fix conjugated antibodies. Sections were rinsed again and mounted with a mounting media containing a DAPI solution to counterstain nuclei.

## FVII antigen determination

The quantitative determination of FVII antigen levels was performed using the Asserachrom® VII:Ag kit (Diagnostica Stago, Asnieres-sur-Seine, France) that exploited the Enzyme-Linked ImmunoSorbent Assay (*ELISA*) technique.

A plastic support coated with specific monoclonal rabbit anti-human FVII antibodies captures FVII to be measured. The sample (200 µl) is placed into the well and incubated for 2 hours at room temperature (18-25°C). Next, after washing the plate 5 times with 200 µl of phosphate buffer, polyclonal rabbit anti-FVII antibodies coupled with peroxidase (200 µl) bind to the remaining free antigenic determinants of FVII, forming a sort of "sandwich", during an incubation of 2 hours at room temperature. After washing the plate with 200 µl of phosphate buffer, the bound enzyme peroxidase is revealed by its activity on the substrate ortho-phenylenediamine (OPD) in the presence of hydrogen peroxide. The reaction is stopped with a strong acid (50 µl of $H_2SO_4$ 3 M) and read at 492 nm on a SpectraFluorPlus microplate reader (TECAN, Salzburg, Austria). In these experimental conditions, the intensity of the color produced bears a direct relationship with the FVII concentration initially present in the sample. In plasma assays a calibration curve was constructed

using serial dilutions of pooled normal plasma. To quantify antigen levels of recombinant FVII, beside the standard curve with PNP, a reference curve with serial dilution of recombinant Wt-FVII in conditioned medium was made.

## FVII activity determination

FVII activity was assessed using two functional assays. The first one is able to measure the activity of FVII toward its physiologic substrate, factor X, using a specific FXa fluorogenic substrate. The second one instead evaluate the ability of FVII to trigger the coagulation cascade in a like-physiological contest, resulting in the formation of the blood clot.

FXa assay is basically performed as follow: Factor VII in conditioned medium was incubated for 10 minutes at 37°C with an activation mixture containing 0.5 nM human Factor Xa (hFXa) (Haematologic Technologies Inc., Essex Junction, VT, USA) and Innovin (Dade Behring, Marburg, Germany). Then, a second mixture containing 100 nM hFX (Haematologic Technologies Inc.) and 500 µM FXa fluorogenic substrate was added to activated samples. Fluorescence (360 nm excitation, 465 nm emission) was measured on SpectraFluorPlus microplate reader (TECAN, Salzburg, Austria).

FVII coagulant activity in plasma was assessed by Prothrombin Time (PT)-based assays, using the conventional one-stage clotting test. Innovin (Dade-Behring, Marlburg GmbH, Germany) was used as a source of tissue factor and phospholipids. Mouse plasma or medium (50 µl) was added to 50 µl of FVII deficient plasma (Dade-Behring, Marlburg GmbH, Germany); 200 µl of thromboplastin were added to trigger the reaction and the clotting time was immediately recorded using a fibrometer. Standard curve with serial dilution of PNP in FVII depleted plasma was prepared in parallel.

## FIX antigen determination

FIX antigen levels in conditioned medium were evaluated by ELISA (Factor IX antigen F.IX; Affinity Biologicals, Ancaster, Canada).

## Western blot analysis for FIX protein

For Western blotting 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 over night 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 FIX or rFIX-wt were used to optimize the assay.

## Factor IX activity and protein assays

Factor IX coagulant activity was assessed by aPTT assays. Basically 25 μl of medium is placed on test cuvette and warmed to 37°C for few minutes. 25 μl of aPTT reagent (Teco Diagnostics, Anaheim, CA) is added to the medium and 25 μl of 37°C warmed Calcium Chloride solution (0.02M) to trigger the solution. The clotting time is recorded in seconds. 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 Corporation, 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.

# MTT

The MTT assay is a colorimetric assays for measuring the activity of enzymes that reduce MTT or close dyes (XTT, MTS, WSTs) to formazan dyes, giving a purple color. A main application allows to assess the viability (cell counting) and the proliferation of cells (cell culture assays). It can also be used to determine cytotoxicity of potential medicinal agents and toxic materials, since those agents would stimulate or inhibit cell viability and growth. To perform this assay, it is important seed known number of cells in each well. Basically a sub confluent amount of cells are seeded into 96-wells plate. Upon the adhesion of cells to the bottom of the well, cells are transfected following the tranfectant reagent instruction using OPTIMEM medium (Life Technologies, USA). Medium is replaced with fresh DMEM after 4 hours to allow properly cell grow. 24 hours after the transfection, MTT working solution (0,5mg/ml MTT in RPMI without phenol red) is added to the cells and incubate 4 hours at 37°C. The converted dye is solubilized with acidic isopropanol (0,04 M HCl in absolute isopropanol) and mixed well by several pipetting up and down. Absorbance of converted dye is measured by spectrophotometer at 570nm with background subtraction at 650nm.

## Annexin V protocol

Annexin V, a 35.8-kDa protein, has a strong affinity for phosphatidylserine (PS), a phospholipid component that is early exposed in the outer layer of membranes during apoptosis process. Basically cells are trypsinized and washed twice with 1X PBS. Cells are then resuspended in 200 ul of 1X Binding buffer and 5 ul of FITC labeled Annexin V and 10 ul of Propidium Iodide are added to the mixture. After 15 minutes of incubation, cells are analyzed by flow cytometry using a single laser emitting excitation light at 488 nm.

## Cell cycle protocol

Propidium iodide (or PI) is an intercalating agent and a fluorescent molecule with a molecular mass of 668.4 Da that can be used to stain cells. PI binds to DNA by intercalating between the bases with little or no sequence preference and with a stoichiometry of one dye per 4–5 base pairs of DNA. Once the dye is bound to nucleic acids, its fluorescence is enhanced 20- to 30-fold. PI is membrane impermeant and generally excluded from viable cells. To allow DNA staining with propidium iodide, cells are permeabilised with a buffer containing a detergent agent (such as Triton X-100 or NP-40) and treated with PI to stain the DNA. The fluorescence intensity of the stained cells at certain wavelengths therefore correlate with the amount of DNA they contain and can be revealed

by flow cytometry . As the DNA content of cells duplicates during the S phase of the cell cycle, the relative amount of cells in the $G_0$ phase and $G_1$ phase (before S phase), in the S phase, and in the $G_2$ phase and M phase (after S phase) can be determined, as the fluorescence of cells in the $G_2$/M phase will be twice as high as that of cells in the $G_0$/$G_1$ phase.

## 2D electrophoresis

This technique consists in two different protein separation phases, by isoelectrofucusing point and by mass.

The first step is the complete solubilization of 100ug of the samples by adding the solubilization buffer (8M urea, 0,5M tiourea, 4% CHAPS, 65mM DTT, 40mM Tris, 0,1 mM EDTA) up to a final volume of 150ul at room temperature (RT) for 1.5 hour under gentle mix. The solubilization is followed by a reduction step, for reducing all the disulfide bonds, adding 5mM TBP for 30 minutes at RT; 15mM iodoacetamide (alkylating agent for the sulfhydryl groups) is added to the reduced samples for 1 hour at RT; to avoid an excess of reactivity of the iodoacetamide, a second treatment with 5mM TBP for 15 minutes at RT is performed.

The last step before the isoelectrofocalization process (IEF) is the precipitation of the insoluble fraction at 14000g for 5 minutes; supernatants are then transferred in new tubes.

The samples are loaded on strips (non-linear pH 3-10 range, previously hydrated over night with hydratation buffer composed by 8M urea, 0,5M tiourea, 2% CHAPS, 10mM DTT, anpholine ph 4-8 - 0,36%, ph 3-10 - 0,24%, blue-bromophenol) and the IEF is performed in the IEF cell (Bio-Rad Laboratories, Hercules  California, USA) until a final voltage/hour of 120000VH. During this procedure proteins are separated by isoelectrofocusing point.

At the end of the IEF, the strips are placed in a specific plate and soaked with the equilibration buffer (6M urea, 1% SDS, 50mM Tris ph 8,8, 30% Glycerol, blue-bromophenol) for 30 minutes under gentle mix at RT; this step is necessary for conferring a unique negative charge to all the proteins  separated into the strip.

The strips are placed onto a 12% SDS polyacrylamide gel and soaked with agarose. When the gel is ready, it is placed on a XL cell (Bio-Rad Laboratories, Hercules  California, USA) and the second dimension is performed at 35mA for 6 hours, inducing the complete separation of proteins according to their mass.

The final step is the staining of the bidimensional gels with a mass compatible silver stain (Invitrogen Corporation, Carlsbad, CA), followed by the image acquisition with Pharos-FX

molecular image scanner (Bio-Rad Laboratories, Hercules  California, USA) and the final analysis of the qualitative/quantitative variations of the spots is accomplished with a specific software (PD Quest, Bio-Rad Laboratories, Hercules  California, USA)


## AAV production

The minimum regions in helper adenovirus that mediate the replication of the AAV vector are E1, E2A, E4, and VA1. The 293 line of human embryonic kidney cells encodes the E1 region of the Ad5 genome2. When a helper plasmid that encodes the E2A, E4, and VA regions (Ad-helper plasmid) is used to transfect 293 cells, together with plasmids that encode the genome of the AAV vector (vector plasmid), as well as the rep and cap genes (AAV-helper plasmid), the AAV vector is produced as efficiently as when infection by adenovirus is used as a source of helper virus. Elimination of the heat-inactivation step directed against contaminating adenovirus can improve the yield of the virus. Furthermore, contamination of the AAV vector stock by most adenoviral proteins can be avoided when this helper virus-free method is used.

*Reagents:*

· Helper plasmid DNA (pHLP, pAdeno)
· Vector plasmid harboring the gene of interest flanked by inverted terminal
· repeats (ITRs )
· 293 cells (human embryonic kidney cells)
· DMEM/F12 culture medium (Invitrogen Corporation, Carlsbad, CA)
· Fetal bovine serum
· 2 x HBS buffer, containing 290 mM NaCl, 50 mM HEPES buffer and 1.5
· mM Na2HPO4, pH 7.1
· 300 mM CaCl2
· Phosphate-buffered saline (PBS)
· 1 M HEPES buffer, pH 7.4
· 100 mM Tris-HCl (pH 8.0) plus 150 mM NaCl (TBS)
· 0.5 M EDTA (pH 8.0)
· 40% sucrose plus 0.01% BSA in TBS
· DNase buffer, containing 50 mM Hepes (pH 7.6), 0.15 M NaCl and 10 mM
· MgCl2

- HNE buffer, containing 50 mM Hepes (pH 7.4), 0.15 M NaCl and 25 mM
- EDTA
- A solution of CsCl in HNE (1.25 g/ml) in HNE
- A solution of CsCl in HNE (1.50 g/ml) in HNE

*Plasmids*

The AAV vector plasmid harbors the desired expression cassette flanked by inverted terminal repeats (ITRs). The AAV-helper plasmid pHLP, harboring rep and cap, has been descriped previously as pHLP19. The Ad-helper plasmid pAdeno is identical to pVAE2AE4-5 and encodes the entire E2A and E4 regions plus the VA RNA I and II genes.

*Transfection and Extraction of Virus:*

That protocol works fine for the transfection of cells in one 225-cm$^2$ flask. For cultures of other sizes, multiply volumes on a linear basis. Trypsinized 293 cells were seeded at 5 x 10$^6$ cells per 225-cm$^2$ flask to generate a monolayer of 20% to 40% confluence when cells attach initially to the surface of the flask. An even density of cells over the entire substratum is essential for high yield and it can be achieved by moving the flask of newly plated cells gently in a crosswize pattern before cells become attached. The plate was placed in an incubator in 5% CO2 in air and allow cells to grow to 80% confluence (24 to 48 h.). One hour before transfection, half the medium was replaced in the flask with fresh medium. 23 µg of vector and of each helper plasmid were added to 4 ml of 300 mM CaCl$_2$. Gently this solution was added to 4 ml of 2x HBS and mixed immediately by gentle inversion three times. Immediately this mixture was transferred into the 225-cm$^2$ flask of 293 cells in 40 ml of DMEM/F12 medium plus 10% FCS and swirled to produce a homogeneous solution. Immediately the plate was returned to the incubator and incubated at 37°C for 4 to 6 hr. At the end of the incubation, the medium was replaced with pre-warmed DMEM/F12 culture medium containing 2% FCS. Three days after transfection, 1 ml of 0.5 M EDTA was added to the flask and incubate for 3 min at room temperature. The suspension of cells was collected and centrifuged at 300g for 10 min. The supernatant was removed and cells were re-suspended in 2 ml of TBS. Cells were lysed by three cycles of freezing and thawing by placing them alternately in a dry ice/ethanol bath and in a 37°C water bath. Tissue debris were removed by centrifugation at 10,000g for 10 min and the supernatant was collected.

*Purification of the AAV vector:*

11 ml of the previously prepared solution of 40% sucrose plus 0.01% BSA in TBS was placed in a sterile ultracentrifuge tube (Nalge Nunc, Rochester, NY). 48 ml of pooled supernatants was carefully placed on this solution. The crude viral particles were harvested by centrifugation at

100,000g for 16 hours at 4°C. Pellet was vigorously resuspended by agitation in 5 ml of DNase buffer. 1,000 units of DNase I was added and incubated it for 1 hr at 37°C. 250 µL of 0.5 M EDTA was added and then the debris was removed by centrifugation at 10,000g for 2 min. The supernatant was filtered through a low-protein-binding 5-µm syringe filter and loaded onto a two-tier CsCl gradient (1.25 g/ml and 1.50 g/ml) prepared in HNE buffer. The solution was spun at 35,000 rpm for 2 h at 16°C in an SW40 rotor (Beckman Instruments, Palo Alto, CA). The band of viral particles was collected and loaded on a second two-tier CsCl gradient (1.25 g/ml and 1.50 g/ml) prepared in HNE buffer. The gradient was spun at 65,000 rpm for 2 h at 16°C in a VTi65.2 rotor (Beckman Instruments). 0.5 ml fractions were collected and selected by semi-quantitative PCR analysis, Western blotting with antibodies against Cap, or quantitative DNA dot-blot hybridization to find the most virus-rich fraction. A dialysis cassette was used (Slide-A-Lyzer; Pierce, Rockford, IL) to desalt the virus-rich fraction by three cycles of dilution with 300 ml of HNE buffer. The concentration to 50 µL was carried out exploiting the Ultrafree-4 columns (Millipore, Bedford, MA) according to the manufacturer's instruction. The final titer usually ranged between $1 \times 10^{13}$ and $5 \times 10^{13}$ particles from $5 \times 10^8$ 293 cells, as determined by quantitative DNA dot-blot hybridization or Southern blotting.

## Hydrodynamic injection

*Nucleic Acid Preparation:*

1. Determine the required total injection volume by using the following formula:

Total volume needed per mouse (in ml) = mouse weight (g)/10 + 0.1 ml Delivery Solution (The addition of the 0.1 ml of Delivery Solution represents the void volume that remains in the syringe and needle after injection).

For example, a 20 g mouse would require a total volume of 2.1 ml.

NOTE: Optimal mouse weight is between 18-25 g, which requires 1.9-2.6 ml of injection volume per mouse.

2. Determine the volume of nucleic acid needed for the injection. It is recommended 1-50 µg as a starting range. Ten µg of DNA or 40 µg for siRNA are good starting points, but a titration may be beneficial for optimal delivery.

3. Subtract the volume of nucleic acid from the total injection volume (from Step 1) needed. The remainder represents the volume of Delivery Solution needed.

For example, to inject 10 µg of nucleic acid into a 20 g mouse:

Nucleic acid stock (100 µg/ml)

Total Volume:   2.1 ml

Volume of DNA required to inject 10 µg:   100µl

Volume of Delivery Solution:  2.1ml - 100 µl =  2ml

The nucleic acid and Delivery Solution can be scaled up for additional mice as needed for replicate injections.

4. Immediately prior to injection, add nucleic acid (from Step 2) to a sterile plastic tube.

5. Add the required volume of Delivery Solution (from Step 3) pre-warmed to 37°C to the tube containing the nucleic acid and mix well. Inject the nucleic acid/Delivery Solution within 30 minutes of mixing.

6. Connect the needle to the syringe and fill with the entire injection solution, ensuring that no air bubbles are present in the needle or syringe.

*Injection Protocol*

*Preparation of Animal for Injection*

Generally, younger laboratory mice (~5-6 weeks old) are optimal for gene delivery. Mice that are older or have more body fat may exhibit compromised gene delivery. It is recommend starting with mice that are 18-25 g each. Use of anesthesia is optional Anesthesia is generally not required when the provided restraint device is employed.

1. To facilitate tail vein visualization and ensure optimal injections, dilate the tail vessels immediately prior to injection by warming the tail of the mouse with a safe, effective heat source (e.g., warm water (~37°C) or heat lamp (120W bulb)) for 3-5 minutes. As the mouse tail warms up, the vein should dilate and become more visible. Do not overheat the mice with the heat lamp. Excessive movement and/or perspiration are indicators of overheated mice.

2. Use the provided restraint device to secure the mouse during the injection. The small opening at the bottom of the tube is designed to facilitate the animal's breathing during the injection procedure. The slit opening in the cap end of the tube is designed to allow tail exposure. Place the mouse head first into the tube, gently place the tail through the slit, and then screw the cap carefully onto the tube. The tail should now be easily accessible. For smaller mice, adding paper toweling or cheese cloth to the tube before screwing on the cap may provide more security. The restraint device can be taped to a table to maintain orientation.

*Injection*

1. While working under a light source, locate the dilated vein on the ventral side of the mouse tail, preferably near the distal end (tip) of the tail. Swab the area with an alcohol swab and allow it to air dry to further increase vein visibility and clean the injection site.

2. Place the syringe needle nearly parallel to the tail with the bevel down (toward the tail). Insert the needle into the tail vein. Check needle placement by injecting a small volume in the vein. If the needle is inserted correctly, the vein should begin to clear of blood. If there is significant resistance, the needle may not be properly inserted into the tail vein. Improper needle insertion into tail tissue is characterized by discoloration and local swelling. If this occurs, remove the needle and reposition it correctly moving further proximal on the tail.

3. Insert almost the full length of the needle into the vein (to prevent accidental removal of the needle while injecting). Dispense the complete injection volume into the mouse tail vein within 4-7 seconds at a constant rate. A good injection is characterized by a constant resistance that does not increase during the procedure.

# III. Aim

## Aim of the thesis

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

Current treatment for coagulation factor deficiencies is based on the intravenous administration of the missing proteins (replacement therapy), either plasma derived or produced by recombinant DNA technology (Pipe, High et al. 2008), in response to bleeding episodes or prophylactically (Carcao and Aledort 2004).

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

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

Enormous efforts have been pushed on the substitutive gene therapy that consists in the viral or non-viral mediated delivery of a copy of the defective gene (or better of the coding DNA sequence) into the patient's cells, thus triggering stable endogenous expression of the missing protein (Murphy and High 2008; Petrus, Chuah et al. 2010).

An emerging area of research is represented by the correction of the gene expression of the mutated gene by modulation of the messenger RNA (mRNA) processing by the small nuclear ribonucleoprotein U1 (U1snRNA), which has been successfully explored for the treatment of other human genetic disorders (Wood, Yin et al. 2007; Bonetta 2009; Wilton and Fletcher 2011). Notably, RNA targeting would permit restoration of gene expression while maintaining the gene promoter regulation in the cells belonging to the physiological tissue. Moreover, it has the potential to circumvent some limitation due to the large size of certain human disease genes, and could be also effective to address dominant-negative disease forms.

Aim of the research was to explore this approach to restore the expression of coagulation factor VII in the presence of a mutation impairing the pre-mRNA splicing.

As model to address this issue, we chose the F7 IVS7+5g/t mutation that, in several Italian FVII deficient patients, was found to be associated to undetectable plasma FVII levels and a severe bleeding symptomatology. The mutation occurs at position +5 of the donor splice site (5'ss) of intron 7 of F7 gene, which is located in the first of six highly homologous 37bp repeats, which

therefore possess a corresponding number of strong cryptic 5'ss, thus complicating the selection of the correct 5'ss by the spliceosome machinery. The molecular mechanisms underlying aberrant splicing, and the modulation by modified U1snRNA, were investigated by the creation of cellular models and animal models.

In the second part of the PhD, the study has been extended to FIX deficiency or Haemophila B, in which the number of splicing mutations causing severe defects is relatively high (>20%). As models, we chose to investigate mutations either at the donor splice site or at the acceptor site of exon 5, which are candidate to affect the exon definition and cause exon 5 skipping.

Altogether the data from these studies were aimed at providing insights into the ability of modified U1snRNA to restore the expression of coagulation factor VII and IX, and to propose this strategy as innovative therapeutical intervention in coagulation factors disease.

# IV. Results

# IV.1 Aberrant Splicing Mechanisms

## IV.1.1 F7 IVS7+5g/a mutation: Molecular mechanisms of aberrant splicing

As a model for this study in FVII deficiency, we chose the IVS7+5g/a mutation (F7 9726+5g/a) that, by altering a relatively conserved position of the donor splice site (Figure 23), is candidate to alter the splicing process of the nascent F7 pre-mRNA. The mutation IVS7+5g/a was identified in various (n=8) patients from villages located in the North-Eastern and Southern areas of the Sabini mountains in homozygous and heterozygous condition. All homozygous patients were clearly symptomatic (hemarthrosis, epistaxis, menorrhagia, gingival bleeding, muscle hematomas) with life-threatening hemorrhagic symptoms (gastro intestinal and intra-cranial bleeding) and required replacement therapy. The FVII antigen and activity levels, assessed by ELISA and coagulation assays, were virtually undetectable (<1%).



**Figure 23:** *Consensus sequence of the donor splice site.*
*The size of the letter indicates the average of conservation among all donor splice sites in human genes*

## IV.1.2 Computational analysis

To analyze the effect of a mutation on the splicing mechanisms, several tools are available on the Net. The Fruitfly tool available on the http://www.fruitfly.org/seq_tools/splice.html website is a Neural Network Splice site prediction tool that, given a sequence, shows the position and sequence of the predicted splice site (acceptor or donor splice site) with an evaluation (score) of the similarity between the submitted sequence and the conserved one.

When the normal *F7* IVS7 gene sequence was submitted, both natural donor splice site and cryptic ones were predicted correctly, even if with different scores. In presence of the mutation F7 IVS7+5g/a ( F7 9726g/a ), the cryptic donor splice sites were still present, whereas the natural donor splice site was no more detectable. The score of mutated donor splice site became very low, decreasing from 0,44 to below the threshold limit of the prediction tool (Figure 24).
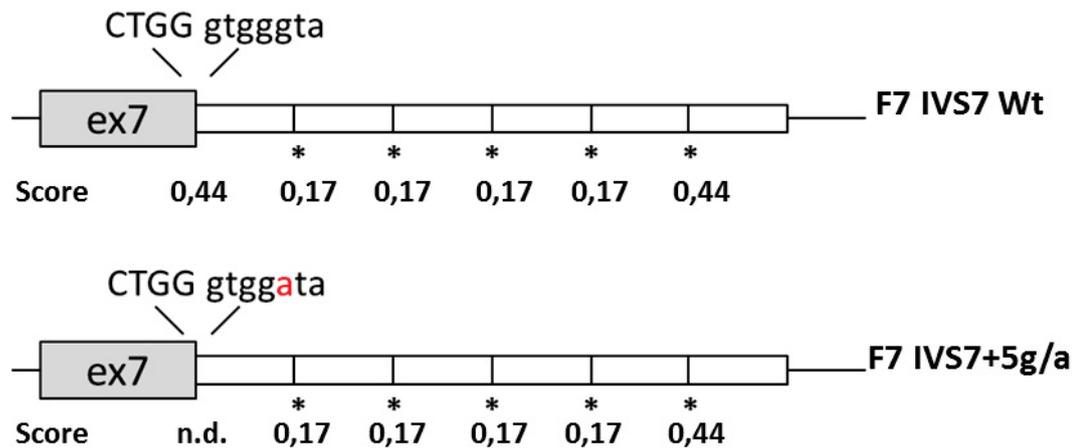
**Figure 24:** *Schematic representation of F7 IVS7 genomic contest.*
*Minisatellites 37bp long are represented by white boxes while exon 7 by grey boxes. Cryptic splice sites are indicated by asterisks. The value of each donor splice sites, calculated by a bioinformatics tool, is indicated below the asterisk. The exon/intron boundary sequence is shown above each construct; exon base-pairs are indicated in upper case. The point of mutation is shown in red.*

Considering the reduction of the score predicted for the donor splice site altered by the mutation, this mutated site should not be used during the splicing process, favoring the activation and utilization of the most proximal cryptic downstream site. The usage of this new donor splice site would generate a longer messenger, which include the first 37bp of the intron 7. When a donor splice site is abolished by a mutation, the most frequent mechanism observed is the exon skipping form, resulting from the utilization of the donor splice site of the upstream exon and the acceptor splice site of the downstream exon. In the presence of the F7 IVS7+5g/a mutation, the same mechanism should occur, with the skipping of exon 6. Both messengers, partial intron retention and the skipping of exon 6, are predicted to cause FVII mRNA frameshift and premature translation termination (Figure 25).
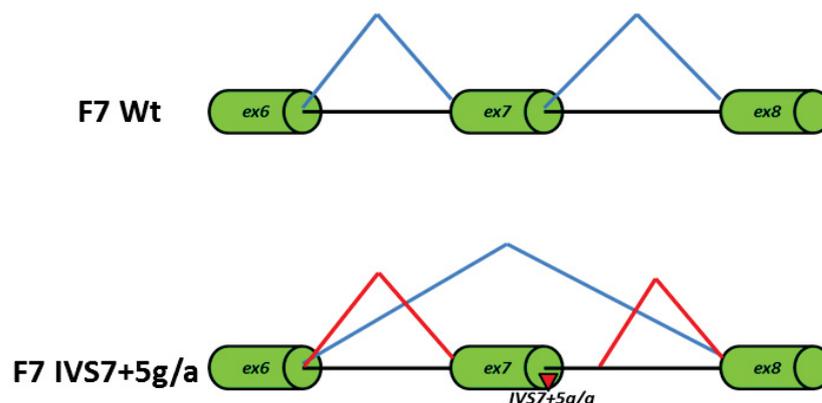


**Figure 25:** *Splicing prediction based on bioinformatics tool.*
*The splicing observed in vivo is shown in the upper part of the figure. Below it is shown the exon 7 skipping (blue line) and the partial intron 7 retention forms (red lines) deduced by bioinformatics analysis in presence of the F7 IVS7+5g/a mutation.*

The prediction – not foreseeing the usage of the correct 5'ss and thus the synthesis of a functional protein- appears in apparent contradiction with the opinion that the total absence of FVII is incompatible with life (due to the absence of subjects with large deletion of the FVII gene and the mortality of know-down mouse for FVII gene), opening an interesting issue on coagulation physiology of these patients.


## IV.1.3 Expression studies with extended minigenes

We are conscious that the investigation of FVII splicing patterns in patients' liver, the physiological site of FVII expression, would be the ideal approach to assess aberrant splicing mechanisms associated to the IVS7+5g/a mutation. However, this was not feasible.

To overcome this limitation, we investigated the effect of this mutation on the splicing mechanism through a minigene approach. This approach has been successfully exploited to study splicing mutations in F7 (Pinotti, Toso et al. 1998; Borensztajn, Sobrier et al. 2003) as well as in many other human disease genes (Susani, Pangrazio et al. 2004; Hartmann, Neveling et al. 2010; Tsuji-Wakisaka, Akao et al. 2011), and the splicing patterns observed *in vitro* mirrored those detected in patient's tissues, when available (Pinotti, Toso et al. 1998; Hartmann, Neveling et al. 2010; Tsuji-Wakisaka, Akao et al. 2011).

We introduced the mutation F7 IVS7+5g/a in an expression vector with the entire FVII genomic region spanning exon 6 through exon 8 through direct amplification of the entire region from a patient heterozygous for the mutation. We created a vector with the mutation, and a control minigene without the mutation, named pFVII IVS7+5g/a and pFVII wt respectively (Figure 26).
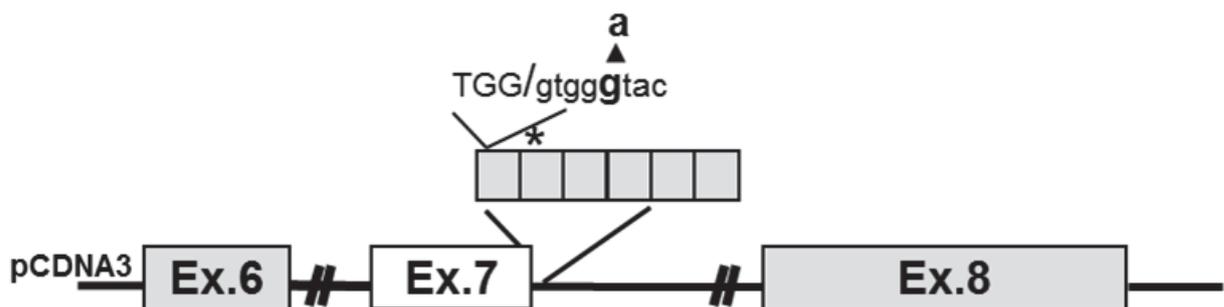


***Figure 26:*** *Schematic representation of pFVII plasmids*
*Schematic representation of the expression cassette cloned into pcDNA3 plasmid. The splicing cassette contains the F7 genomic region spanning exon 6 to exon 8. 37bp Minisatellite repeats are represented by grey squares. Asterisk represent the cryptic donor splice site predicted to be used by the bioinformatics tool. The exon 7-intron 7 boundary is shown in the upper part of the picture. 3' exon 7 end is represented in bold.*

Being the FVII expressed in liver, we chose an human liver cell line (Hep3B cell line) to properly evaluate the effect of the mutation F7 IVS7+5g/a on splicing. Investigation was conducted at RNA level, with the advantages of DNA/RNA amplification protocols. COS-1 cell was also exploited to explore major findings. This cell line, although of kidney origin, is able to properly express coagulation factors.

After the transfection of cells with 3ug of plasmid pFVII-wt or pFVII IVS7+5g/a, RNA was extracted after 48 hours of incubation and used to perform a RT-PCR with primers base pairing with exon 6 and exon 8.

In both cell lines, the Wt minigene drove the expression of a normal transcript (transcript 2 in the Figure 27). After transfection of cell with the pFVII IVS7+5g/a, RNA extraction, RT-PCR with primers base pairing with exon 6 and exon 8, and agarose electrophoresis of the products, two transcripts were detectable (transcript 1 and transcript 3). Transcript 1 came from the inclusion of the first 37 bp of the intron 7, due to the utilization of the first cryptic splice site; transcript 2 instead represented the skipping of the exon 7, with joining of exon 6 and exon 8 (Figure 27).

If translated in proteins, both transcripts would predict for FVII mRNA frameshift and premature translation termination. Exon 7 skipping accounted for the majority of transcripts (80%), as estimated through semiquantitative analysis of bands.
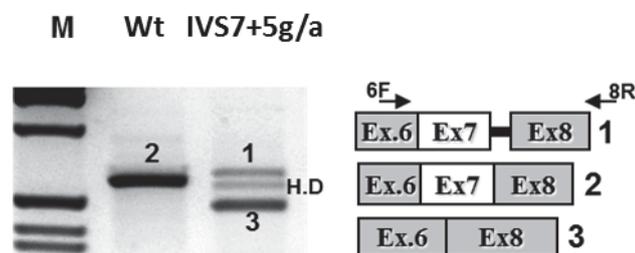


*Figure 27:* Electrophoretic separation on 3% agarose gel of RT-PCR products with primers 6F and 8R obtained from total RNA of cells transfected with the expression vectors for normal (Wt) or mutated (IVS7+5a) minigenes. The scheme of transcripts (1, 2, 3) and the primers used (arrows) are depicted on the right side. The additional band observed for the IVS7+5a mutant corresponded to heteroduplex (H.D). Fragment sizes were: 637 bp (1), 600 bp (2), and 480 bp (3).

The band marked by H.D is due to the high complementarity between the messenger 1 and 3, resulting in a heteroduplex form on agarose gel. In fact, when the band was excised from gel and re-amplified with the same couple of primer, transcripts 1 and 3 were obtained again, confirming its heteroduplex origin (Figure 28). Even more, if we heat-denatured the original RT-PCR to dissolve every heteroduplex form, and loaded it on gel, the middle transcript was no yet detectable (not shown).
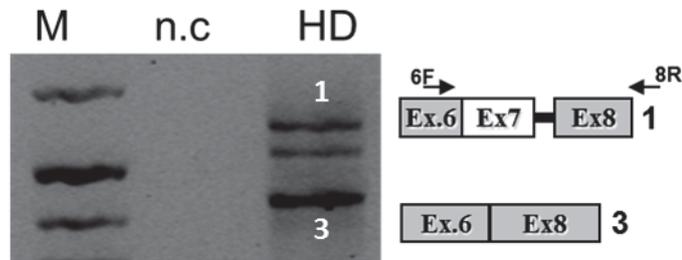
*Figure 28: Electrophoretic separation of PCR with primers 6F and 8R on purified heteroduplex form. The presence of forms 1 and 3 after its amplification confirms the nature of the band.*

To give reasons for the survival of patients, we sought for the presence of normal transcripts originated by the mutant minigene. To accomplish this task, we performed a new RT-PCR with inner primers, base pairing with exon 7 and exon 8. In this way, we amplified specifically only the wt messenger, if present, and the aberrant form including the first 37bp of intron 7. The RT-PCR with primers 7F and 8R' did not reveal the presence of normal transcripts, but only the aberrant one (Figure 29)
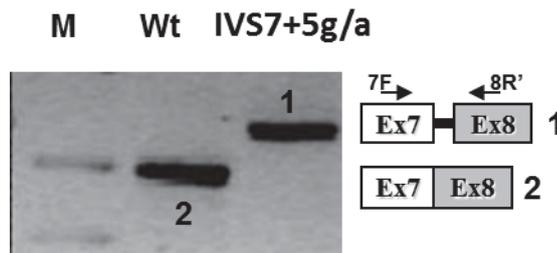


*Figure 29: Electrophoretic separation on 3% agarose gel of RT-PCR products with primers 7F and 8R obtained from total RNA of cells transfected with the expression vectors for normal (Wt) or mutated (IVS7+5g/a) minigenes. The scheme of transcripts (1, 2) and the primers used (arrows) are depicted on the right side. Fragment sizes were: 228 bp (1) and 191 bp (2).*

The hypothesis that the total absence of FVII is incompatible with life seems to be in contradiction with what observed. Due to the reduced detection sensitivity of the agarose gel electrophoresis (only bands with a concentration of 5ng are visible on an agarose gel), we moved towards a more sensible assay, using a capillary electrophoresis system (detection limits as low as $10^{-18}$ to $10^{-21}$ mol). To exploit this methodology we need fluorescent products of PCR, so we designed a FAM-labeled primer to specifically mark the normal transcript, plus the partial intron retention form. Due to the different amplification efficiency of the PCR for products that differ significantly in dimension, the primer that we designed base pairs with exon 7. In that way, the amplified products differ only for 37bp, so the amplification efficiency is almost the same. If even the exon skipping form was included also in the analysis, it would have been amplified with more efficiency, due to

its reduced length. On RNA previously extracted, we performed a new RT-PCR with 7F$^{Fam}$ and 8R. We loaded on capillary system the same amount of each sample, and also an overload of the mutant sample (5X) to try to detect residual traces of normal transcript. The capillary electrophoresis system revealed at low PCR cycle number (20) traces of normal transcript (0,2% of the aberrant form) in the mutant minigene (Figure 30).
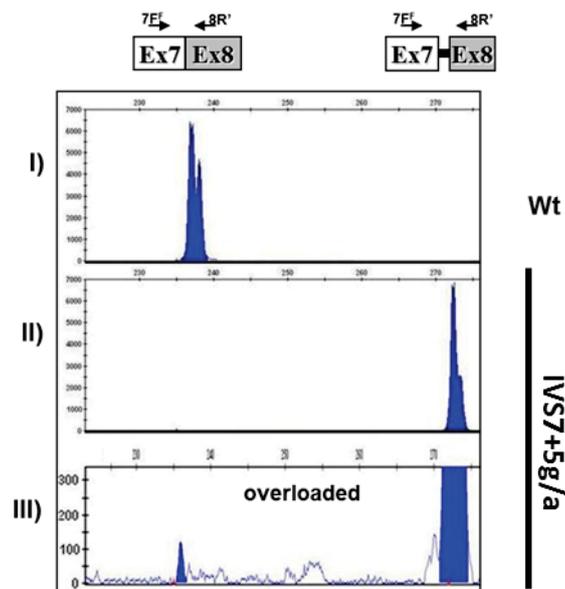


***Figure 30:*** *Traces of wt mRNA revealed by denaturing capillary approach.*
*Separation on a denaturing capillary system (automated ABI-3100) of fluorescently labeled RT-PCR products obtained from total RNA of cells expressing the Wt (i) or the IVS7+5g/a minigene (ii). A 5X overload of the sample in ii) is shown in the lower part of the figure (iii).*

# IV.2 Correction in cellular model

## IV.2.1 Rescue of FVII expression by engineered U1snRNA

The F7 IVS7+5g/a mutation drastically impair the donor splice site (5'ss) of *F7* IVS7 and is responsible for exon skipping and, to a less extent, to a 37bp inclusion (partial intron retention). The extremely reduced efficiency in using the correct exon-intron junction could be explained by the inefficient recognition of the mutated 5'ss by the U1-snRNA through base pairing with its 5'-tail. As a matter of fact, recognition of the 5ìss by the U1snRNA is the key event in the earliest steps of the splicing process.

The first idea to use a modified U1-snRNA, designed to recognize and bind the mutated donor splice site, was born from the observation of Alan M Zahler et al (Zahler, Tuttle et al. 2004) who observed a compensatory U1-snRNA in C.elegans able to partially suppress a splicing mutation at the invariable +1G position of the intron. From that, only few groups attempted to use modified U1s to suppress the effects of mutations affecting the donor splice sites .

Since the mutation F7 IVS7+5g/a impairs the recognition of exon 7 by the spliceosome, we thought to use a modified U1-snRNA based approach to restore exon 7 inclusion favoring the selection of the correct donor splice site.

In our laboratory, the gene coding for the U1-snRNA is available in a plasmid (pGem3 plasmid) under its own promoter. To engineer the 5' tail of the U1-snRNA, the sequence between BclI and BglII sites of the parental U1-snRNA gene was replaced with the desired sequence ( the sequence that U1 should recognize).

To evaluate the feasibility of modified U1 approach to restore exon 7 inclusion and selection of the correct 5'ss, we designed different U1-snRNAs. One U1-snRNA (U1+5a) was conceived to base pair with the mutated donor splice site, whereas other two U1-snRNA was designed to base pair with neighboring sequences (U1Mut1 and U1Mut2) to get insight on the causative mechanisms of the mutation. The U1+5a differs from the parental U1-snRNA gene for a single base pair substitution at +5 position, with a uracile nucleotide instead of a cytosine. The U1Mut1 and U1Mut2 variants base pairs with different sites in the IVS7 repeats, respectively the end of the exon 7 and the nucleotide positions 16 to 25 of the intron 7 (Figure 31).
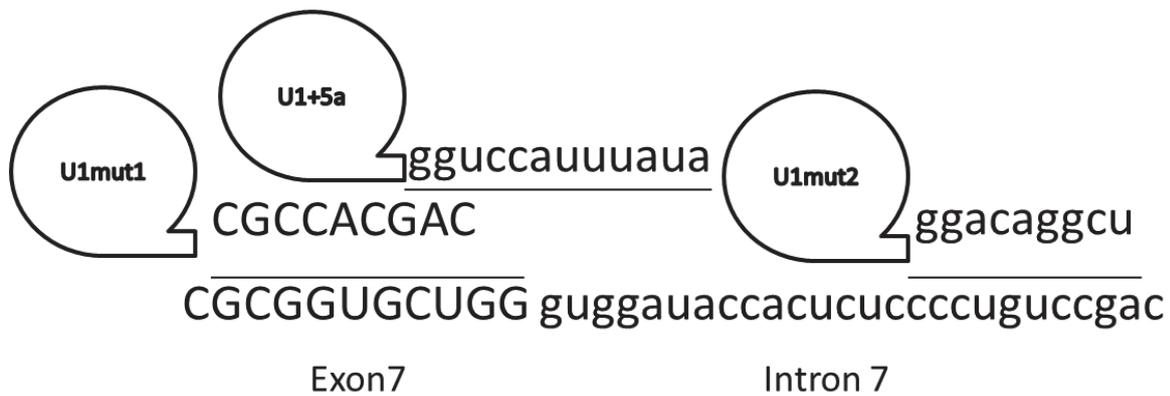
*Figure 31:Binding regions of U1snRNAs on nascent pre-mRNA F7 IVS7 region
Schematic representation of the positions where the U1snRNAs variants bind. Exonic sequence is
represented in capital letters and sequence in lowercase is part of the F7 IVS7intron.*

## IV.2.2 Rescue of FVII mRNA processing

Hep3B (human hepatoma) cells, seeded into 30mm dishes, were transiently transfected with 3ug of each minigene plasmid (pFVII-wt or pFVII IVS7+5g/a) and with equimolar concentrations of plasmids encoding for the modified U1-snRNA using Lipofectamine 2000 reagent. The retro-transcription of extracted RNA and amplification of FVII transcripts were conducted using primers 6F and 8R.

The gel separation of the bands resulting from the RT-PCR revealed that all mutated U1-snRNA reduced exon skipping form, with more than 50% of reduction based on band densitometry (Figure 32).
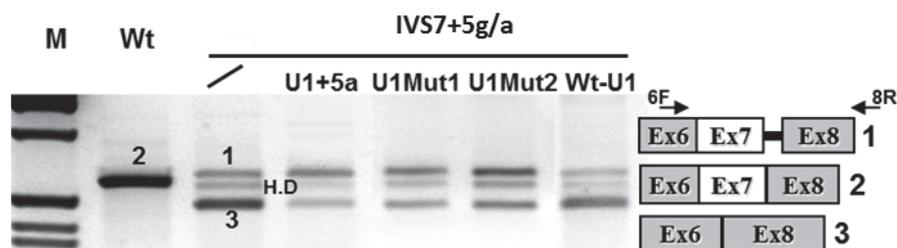


*Figure 32:Electrophoretic separation on 3% agarose gel of RT-PCR products obtained from total RNA of cells
transfected with the expression vectors for normal (Wt) or mutated (IVS7+5g/a) minigenes, and equimolar
concentrations of plasmids encoding for the modified U1-snRNAs. The scheme of transcripts (1, 2, 3) and the primers
used (arrows) are depicted on the right side. The name of modified U1-snRNAs used in transfection is indicated above.*

To evaluate the ability of engineered U1-snRNA to redirect the spliceosome on the right donor splice site in presence of the mutation F7 IVS7+5g/a, we performed a new RT-PCR with inner primers, which base pair with exon 7 and exon 8 ( 7F and 8R respectively). In that way, we can amplify selectively only transcript 3 (partial intron retention) and 2 (wt transcript). The new RT-PCR, with primers 7F and 8R, revealed that, among all engineered U1-snRNAs, only the U1+5a was able to redirect the spliceosome at the correct donor splice site and induce the generation of normal transcript (Figure 33).
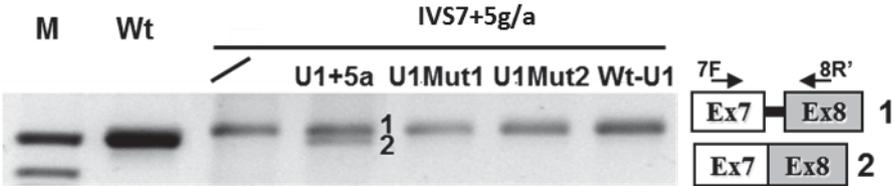


*Figure 33:Electrophoretic separation on 3% agarose gel of RT-PCR products conducted with primers 7F and 8R' on RNA of cells transfected with normal (Wt) o mutated (IVS7+5g/a) minigenes and equimolar concentrations of plasmids encoding for the modified U1-snRNAs. Only the engineered U1+5a show the presence of transcript 2.*

The nature of transcripts 1, 2 and 3 was further confirmed through sequencing.



*Sequencing of transcripts 1, 2 and 3*
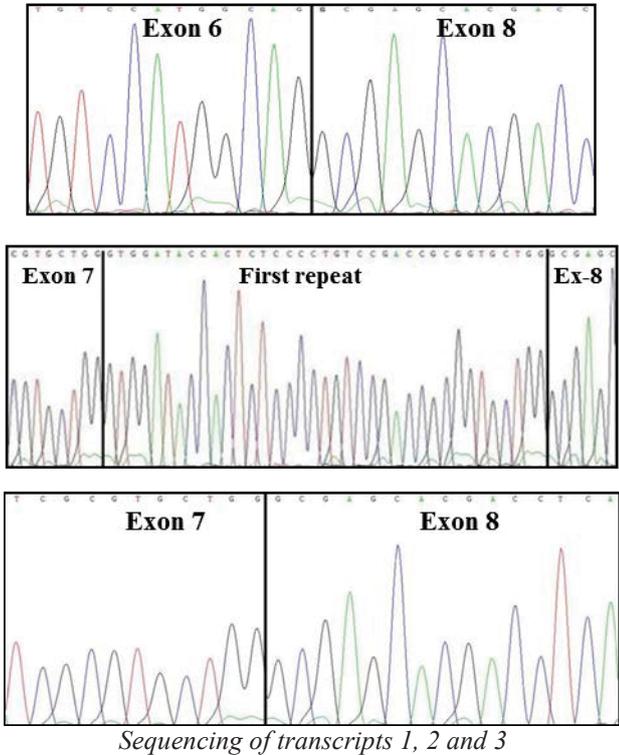
To better evaluate the rescue ability of the U1+5a on FVII mutation, we analyzed the effect of increasing molar amount of plasmid encoding the engineered U1snRNA (U1+5a). Cells were transiently transfected with the previous amount of each minigenes and with 0,5X, 1X and 1,5X molar excess of pU1+5a plasmid. The retro-transcription and amplification of the extracted RNA was conducted with primers 6F and 8R and it revealed that the exon skipping form (transcript 3 in the picture) decreased in a dose dependent manner (mainly at higher pU1+5a molar concentration) (Figure 34). Moreover, the cotransfection of the Wt minigene with U1+5a plasmid at 1,5X molar excess did not reveal any appreciable change in the expression profile of the Wt minigene.
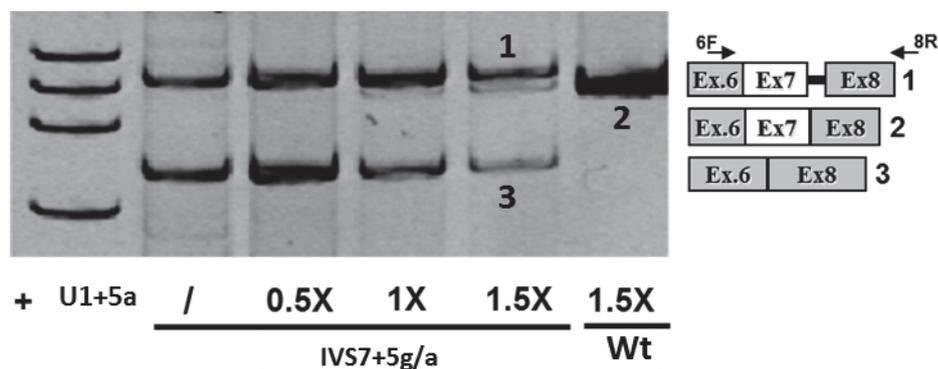


**Figure 34:** *Separation on 12% polyacrylamide gel of RT-PCR products conducted with primers 6F and 8R on RNA of cells transfected with normal (Wt) o mutated (IVS7+5g/a) minigenes and increasing molar concentrations of engineered U1-snRNA (U1+5a). The relative molar amount of transfected plasmids is indicated below.*

The specific quantification of the rescued form by real-time PCR was virtually impossible due to the high sequence complementarity between transcript 1 and 2. To better evaluate and quantify the rescue effect of the U1+5a, we moved toward the utilization of denaturing capillary electrophoresis system, known to have higher sensitivity and resolution than agarose electrophoresis. When the amplification efficiency of PCR products is comparable, it is possible evaluate their initial amount in the reaction if the analysis is performed during the linear amplification step of PCR protocol.

To exploit this methodology we need fluorescent products of PCR, so we designed a FAM-labeled primer to specifically mark the normal transcript, plus the partial intron retention form. Due to the different amplification efficiency of the PCR for products that differ significantly in dimension, the primer that we designed base pairs with exon 7. In that way, the amplified products differ only for 37bp, so the amplification efficiency is almost the same. If the exon skipping form was included also in the analysis, it would have been amplified with more efficiency, due to its reduced length.

On RNA previously extracted, we performed a new RT-PCR with 7F[Fam] and 8R. We loaded on capillary system the same amount of each sample, and also an overload of the mutant sample (5X) to try to detect residual traces of normal transcript. The capillary electrophoresis system, at low

PCR cycle number (20), further highlighted the U1+5a-mediated rescue (Figure 35). As suggested by agarose gel analysis, the correction resulted in a dose dependency manner. In particular, at equimolar concentrations of the pIVS7+5a and pU1+5a, the normal transcripts were 25±5% of the partial intron retention form and reached 53±7% with the further excess of pU1+5a. As observed on agarose gel, the increase of the rescued form (normal transcript) is not followed by a decrease of the aberrant form (partial intron retention form), being the rescue due to the higher exon definition (Figure 35).
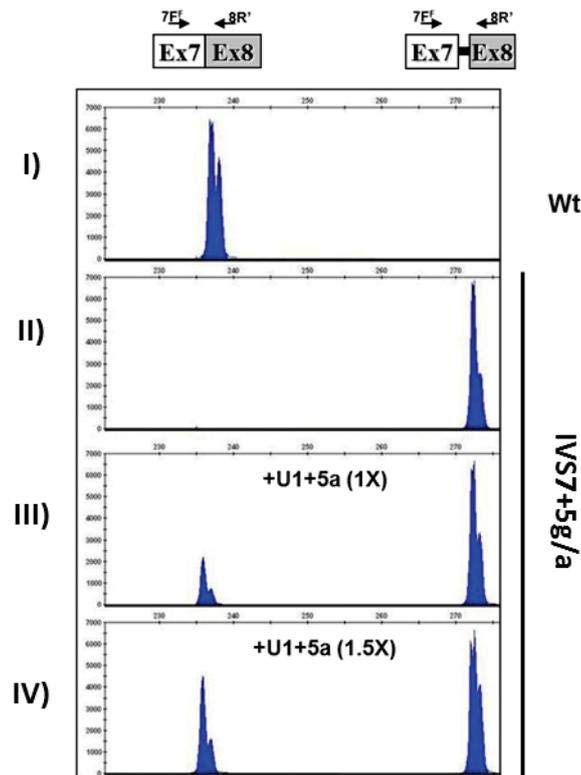


*Figure 35:Separation on a denaturing capillary system (automated ABI-3100) of fluorescently labeled RT-PCR products obtained from total RNA of cells expressing the Wt (i) or the IVS7+5g/a minigene without (ii) and on (iii and iv) overexpression of the U1+5a snRNA. The scheme of transcripts, and of primers used (arrows), is depicted on the top. Separation of 1 uL of 1:100 diluted RT-PCR reaction is shown. The peaks, which might suggest the presence of a doublet, were referred to as single bands by the automated ABI-3100 system. As expected, the fragment sizes of the normal and aberrant transcripts were 236 bp and 273 bp, respectively.*

The expression of U1+5A reduced from 80% to 40% the exon 7 skipping form (Figure 34) and increased the recognition of the correct 5'ss, resulting in an appreciable synthesis of normal transcripts (from hardly detectable to ~30%). If translated in protein, the FVII antigen should reach about 100ng/ml, a value that should ensure normal hemostasis.

## IV.2.3 Rescue of FVII protein biosynthesis and coagulant function

The data obtained with the minigene demonstrate that a modified U1snRNA loaded on the mutated donor splice site  is able to efficiently restore the splicing process. However, this experimental system has the limitation of not permitting the assement of the rescue of  protein biosynthesis and function, the key issue for the evaluation of the therapeutic potential.

For this reason, we designed and created a splicing competent minigene construct consisting of  the entire coding sequence of FVII (cDNA) harboring the introns 6 and 7 surrounding exon 7 of F7 gene. In particular, the *F7* region spanning exons 6 through 8 (nucleotides 8926-11157) from our previously prepared wild-type and mutated minigenes was subcloned through the XbaI site in exon 6, downstream the FVII coding region spanning exons 1 to 6 in pcDNA$_3$. New plasmids were named pSCFVII-wt and pSCFVII-IVS7+5g/a (Figure 36).
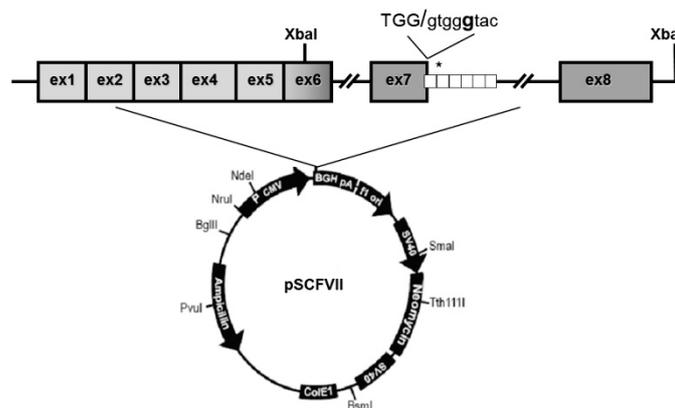


***Figure 36:****Schematic representation of the expression plasmid pSCFVII (wt or IVS7+5g/a). The 37-bp repeats in the IVS7 are indicated by white boxes, and the asterisk represents the 5'ss cryptic site in the second repeat. The 5'ss sequence and the nucleotide mutated (**g** in Wt  and **a** in IVS7+5) are shown at the top.*

To evaluate secreted protein levels and the underlying activity, either by fluorogenic or coagulant assays, we have selected a cellular system which does not naturally synthetize FVII, which would have a confounding effect. For this reason, we chose the COS-1 cells that are known for being able to properly synthesize FVII with its complex post-translational modifications (Hunault, Arbini et al. 1999).

## IV.2.3.1 Validation of the splicing competent construct

COS-1 cells were transiently transfected with the pSCFVII-wt and pSCFVII-IVS7+5g/a plasmids to assess, 72 hours post-transfection, the mRNA processing and secreted protein levels. To guarantee the proper γ-carboxylation of FVII, the medium was supplemented with Vitamin K, the cofactor of the γ-carboxylase.

The RT-PCR conducted with primers 6F and 8R on extracted RNA revealed the splicing patterns overlapping with those obtained with the previous minigene. In fact, the forms lacking exon 7 skipping or presenting the 37bp repeat (+37bp) were clearly detectable in cells transfected with the pSCFVII-IVS7+5a plasmid (Figure 37, left panel). Differently, the pSCFVII-Wt drove the synthesis of correct FVII transcripts. When the RT-PCR was carried out with primers 7F and 8R we observed the synthesis of properly processed FVII transcripts in the pSCFVII-wt expressing cells but only the +37bp form in the pSCFVII-IVS7+5a expressing cells (Figure 37, right panel).
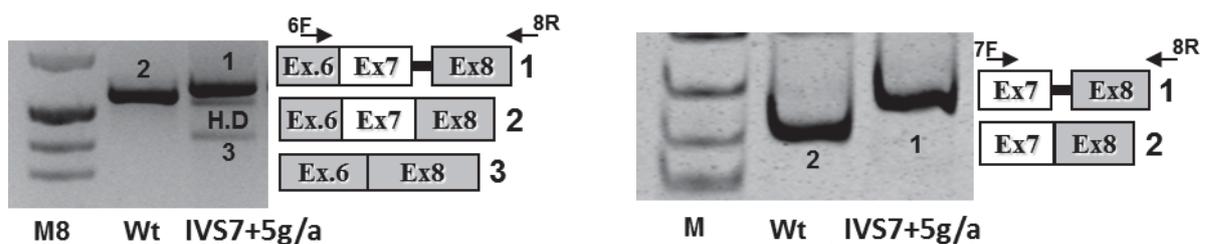


*Figure 37*:*Electrophoresis separation on 3% agarose gel of RT-PCR conducted with primers 6F-8R (left) and 7F-8R (right) from cells transfected with normal (Wt) or mutated (IVS7+5g/a) full-length splicing competent cassette. The schemes of transcripts and position of primers (arrows) are depicted beside the electrophoresis separation.*

At the protein level, the secreted FVII proteins were clearly measurable by ELISA (23±4ng/ml) in medium of cells transfected with pSCFVII-wt. The activity of these molecules (Figure 38) was appreciable both by fluorogenic FXa generation assays and by PT-based assays, the latter enabling the assessment of the procoagulant FVII activity.

At variance, no secreted molecules were detected in cells transfected with the pSCFVII-IVS7+5g/a. Consistently, the activity levels were undetectable and not distinguishable from those observed in medium from cells transfected with the empty pcDNA3 vector, used as negative control (Figure 38).
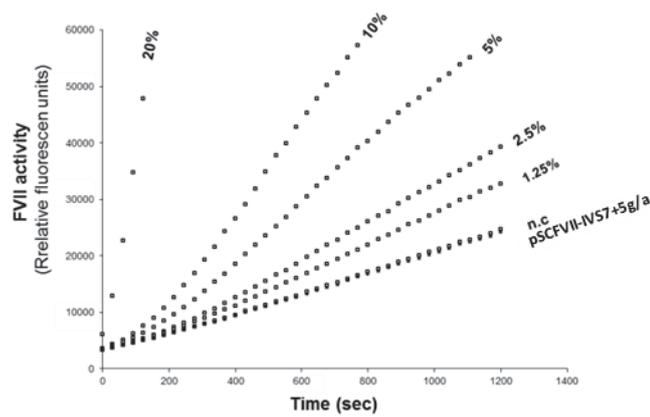
***Figure 38**:FVII activity revealed by fluorogenic FXa generation assay. Medium from cell transfected with pSCFVII-Wt was diluted in un-transfected cells media (n.c) to indicated dilution. The dilution is indicated above each curves. Medium from cells transfected with the mutated pSCFVII-IVS7+5g/a was tested un-diluted.*

These data were in accordance with the coagulation phenotype observed in the IVS7+5g/a homozygous patients (very low level of FVII antigen and activity), thus providing us an appropriate experimental model to evaluate, at protein level, the rescue effect of the engineered U1-snRNA (U1+5a).

## IV.2.3.2  Evaluation of the rescue by the U1+5a

As above reported, to assess the rescue by the u1+5a we cotransfected COS-1 cells with pSCFVII-IVS7+5g/a plasmid and increased molar concentrations of pU1+5a (1X and 1,5X molar excess of pU1+5a plasmid).

RT-PCR with primers 6F and 8R on RNA extracted from transfected cells showed that the expression of the U1+5a resulted in reduction of the exon skipping form in a dose dependence manner, in accordance with our previous findings (Figure 39, left panel). Moreover, RT-PCR conducted with inner primers (7F and 8R) revealed that the engineered U1 (U1+5a)- mediated rescue effect was dose dependent. Through this approach, no traces of correct transcripts were detected in cells transfected with the pSCFVII-IVS7+5g/a only (Figure 39, right panel).
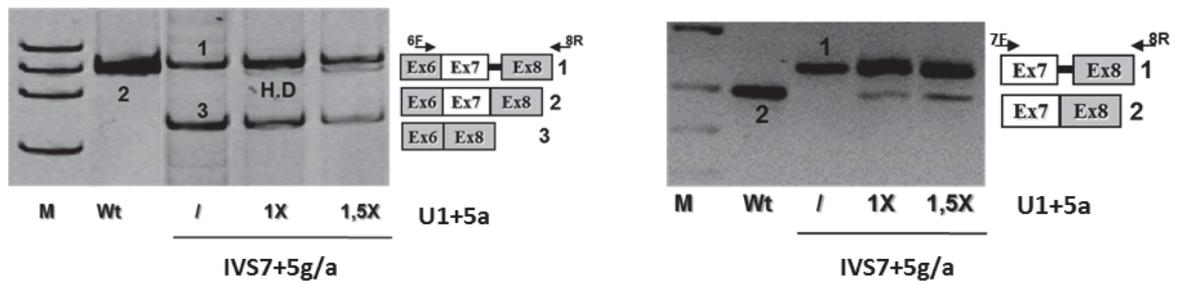
***Figure 39:*** *Separation on 12% polyacrylamide gel (left) or 3% agarose gel of RT-PCR products conducted with primers 6F-8R and 7F-8R, respectively on RNA from cells transfected with the normal (Wt) or mutated (IVS7+5g/a) full-length splicing competent cassette and increasing molar amount of pU1+5a plasmid. The scheme of transcripts and position of primers (arrows) are depicted beside the separation.*

To obtain a raw estimation of the real correction we need to evaluate the amount of the third mRNA form, caused by the exon 7 skipping. Densitometric analysis of RT-PCR bands obtained with primers 6F-8R indicated that the exon skipping form accounted for approximately 60% of all transcripts. Therefore we estimated that the correctly processed FVII form should represent approximately 20% of all FVII mRNAs. The nature of all transcripts was further confirmed through sequencing.

The extent of correction was further evaluated through fluorescent labeling of RT-PCR product with primer 7F$^{Fam}$ and 8R, and subsequent denaturing capillary electrophoresis, which permit the assessment of the relative amount of the correct form over that including the additional 37bp system. In the presence of a molar excess (1,5X) of pU1+5a, the levels of correct transcripts accounted for 48±4% of the +37bp aberrant form, in accordance with our previous findings (Figure 40).
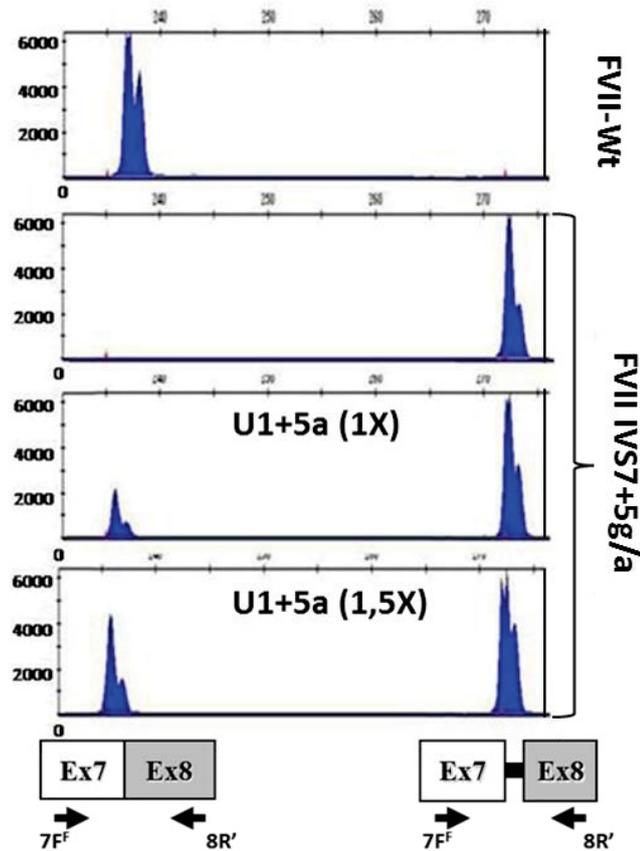
*Figure 40: Separation on a denaturing capillary system (automated ABI-3100) of fluorescently labeled RT-PCR products obtained from total RNA of cells transfected with the normal (Wt) or mutated (IVS7+5g/a) full-length splicing competent cassette alone or in combination with the indicated molar amount of pU1+5a plasmid. The scheme of transcripts, and of primers used (7F^F, fluorescently labeled; 8R), is depicted at the bottom. Separation of 1 uL of 1:100 diluted RT-PCR reaction is shown. As expected, the fragment sizes of the normal and aberrant transcripts were 236 bp and 273 bp, respectively*

The key of a successful genetic treatment is demonstrate at protein level that the innovative approach is able to produce a functional protein at significant level. Therefore, in media withdrew from cells cotransfected with increasing amount of pU1+5a we performed a FVII-specific ELISA and functional assay aimed at assessing secreted FVII protein and its activity.

In medium from cells cotransfected with an excess of pU1+5a, the FVII antigen levels were $5.0 \pm 2.8$ ng/mL, with un-detectable levels in cells transfected only with the mutant minigene.

The FVII activity was assessed through two functional assays: fluorogenic FXa generation assays and PT-based assays. The FXa generation assay is designed to evaluate, in a completely reconstituted system, the activity of FVII toward its physiologic substrate, the Factor X. Since the amount of activated FX (FXa) is measured by means of a FXa-specific fluorogenic substrate, the assay is very sensitive even to very low FVII activity.

The co-expression of the U1+5a was responsible for a significant and dose-dependent increase of FVII activity in medium (Figure 41). In particular, the FVII activity levels from cells co-transfected

with equimolar (1X) or higher (1,5X) concentrations of pU1+5a, were 3,1±1,4% and 8,2±2,6 of those obtained with the pSCFVII-wt, respectively.
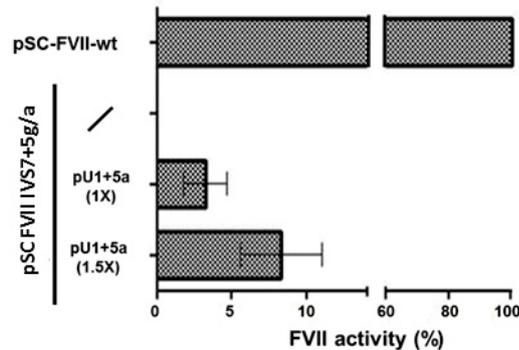


*Figure 41: FVII-mediated FXa generation in conditioned medium from COS-1 cells transfected with pSCFVII-wt or the pSCFVII-IVS7+5g/a without and with equimolar concentration (1X) or an excess (1.5X) of pU1+5a. Experiments are made in triplicate. Mean and SD are shown.*

To further corroborate these findings, we subsequently exploited the PT-based assay that allows measuring the FVII procoagulant activity. Basically, the conditioned medium containing FVII is added to FVII-depleted plasma, which provides with all components of the coagulation cascade. Then, the coagulation is triggered by the addition of reagent containing the cofactor of FVII, Tissue factor and the coagulation time recorded.

Assays were standardized using a standard curve made with serial dilution of FVII-wt in medium from cells transfected with empty pcDNA3, used as negative control. The medium from cells cotransfected with pSCFVII-IVS7+5g/a and a 1X molar excess of pU1+5a did not result in any appreciable change in PTs with respect to the negative control, which is explained by the low sensitivity threshold of the assays. At variance, the coagulant activity in medium from cells cotransfected with the pSCFVII-IVS7+5g/a and a 1,5X molar excess of pU1+5a resulted in coagulation times significantly shorter (100±6 seconds) than those produced by medium from cell transfected with the pSCFVII-IVS7+5g/a alone (119,1-119,6 seconds). Compared to the standard curve, the activity upon rescue corresponded to 9,5±3,2% of that FVII-wt, having a concentration of 25ng/ml (Figure 42).
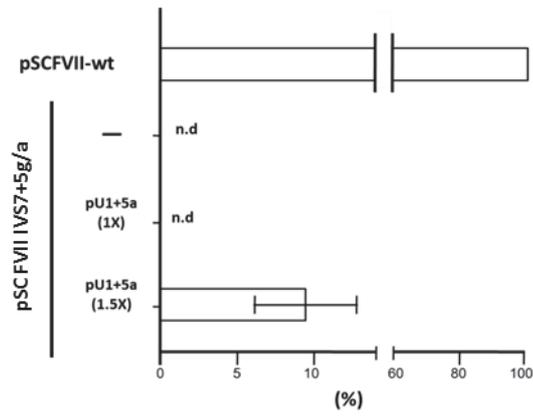
*Figure 42:* *FVII coagulant activity in conditioned medium from COS-1 cells transfected with pSCFVII-wt or the pSCFVII-IVS7+5g/a without and with equimolar concentration (1X) or an excess (1.5X) of pU1+5a. Mean and SD of experiments made in triplicate are shown.*

These promising results prompted us to further investigate the U1-mediated rescue effect in a mouse model of FVII deficiency, that has not been exploited yet.

# IV.3 Rescue of FVII biosynthesis *in vivo*

## IV.3.1 U1+5a-mediated rescue of FVII mRNA processing and protein biosynthesis in the mouse model

The data obtained in the cellular models of FVII deficiency caused by the IVS7+5g/a provided experimental evidence that the modified U1+5a is able to re-direct splicing to the mutated donor splice site and in turn to rescue FVII protein biosynthesis and coagulant function. This is the first successful use of the U1snRNA-mediated correction approach in coagulation factor deficiencies, in which even a tiny increase in circulating levels of the functional protein might have a significant impact on the clinical phenotype. However, the U1snRNA-mediated correction approach has been tested only in the cellular models and there are no data *in vivo*, neither for coagulation deficiencies nor for the other human disease. Even though U1snRNAs have been exploited in animal models of Duchenne Muscular Distrophy (Denti, Rosa et al. 2006; Incitti, De Angelis et al. 2010), they were modified to deliver antisense sequences and induce exon skipping, and not to re-direct the spliceosome and restore correct splicing.

Indeed, we created an *in vivo* model of FVII deficiency caused by the FVII IVS7+5g/a mutation. Since no mouse model is available for this mutation, and because the creation of *ad-hoc* transgenic mouse would have been very laborious, we proceeded through the expression of the human FVII splicing competent minigene in a normal mouse strain (C57BL-6J). The rational was based on the possibility of assessing i) the human FVII (hFVII) mRNA by designing human FVII specific RT-PCR and ii) the hFVII protein through the use of specific anti-hFVII antibodies.

The human FVII shares an homology of 69% with murine FVII at protein level, such as 76% at RNA level. So, the approach to evaluate the effects of modified U1 *in vivo* was based on various commercial antibodies, able to recognize human FVII and with very reduced inter species reactivity, particularly toward mouse, so it should be possible only recognize human FVII in mouse contest (ELISA and IHC assay). The differences between human and mouse FVII RNA should allow the design of human specific primers to just detect human from mouse FVII RNA in mouse liver contest.

Therefore we initially set up and optimized the assays at the mRNA (RT-PCR) and protein (ELISA, immune-hystochemistry) level to properly evaluate the human FVII expression into the mouse context.

To express the splicing competent minigene in the C57BL-6J mice, and to evaluate the efficiency of the U1+5a-mediated rescue, we exploited the hydrodynamic approach and then the use of adenoassociated viral (AAV) vectors, the latter allowing the evaluation of the prolonged restoration of FVII expression.

The scheme of the *in vivo* study was as reported below:
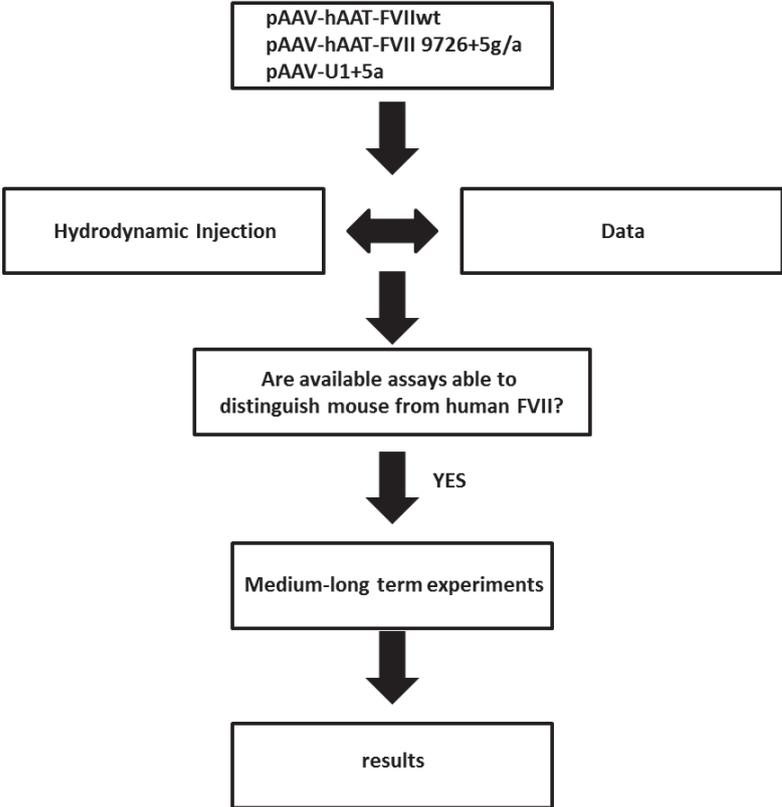


*Scheme of the in vivo approach to evaluate U+5a mediated rescue of FVII expression.*

The *in vivo* experiments have been conducted at the Children's Hospital of Philadelphia, where Dott. Valder Arruda has an established and longstanding experience in the gene therapy of hemophilia B with AAV vectors.

## IV.3.2 Creation of plasmids

The hFVII splicing competent cassette (pSCFVII-wt or pSCFVII-IVS7+5g/a) used in the previous experiments (Part II) was cloned into the AAV backbone to create the pAAV-hAAT-FVII wt and pAAV-hAAT-FVII IVS7+5g/a plasmids. To ensure the liver-specific expression of the transgene, hFVII splicing competent cassette was inserted, through the ClaI-XhoI sites, downstream of a liver specific promoter (hAAT), with consistes of the human α1 antitrypsin promoter and four copies of human apoE enhancer.

To create the pAAV-U1 plasmid, the U1 gene (either U1wt or U1+5a) was cloned into the pAAV plasmid using BamHI sites but keeping its own promoter, known as very "strong" and constitutive promoter.

These vectors are characterized by the presence of two Inverted Terminal Repeats (ITRs), surrounding the cloned gene, to permit the creation of AAV viruses (Figure 43).
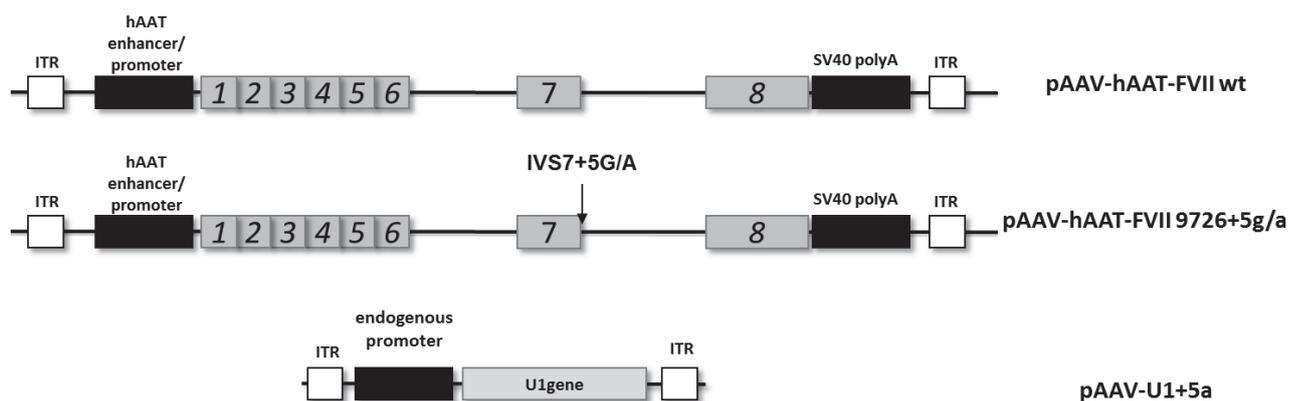


***Figure 43:*** *Scheme of plasmids created to evaluate the feasibility of engineered U1 approach in vivo. The hAAT promoter contains the human α1 antitrypsin (hAAT) promoters and four copies of human apo E enhancer to restrict the expression of FVII cassette in mouse liver. The expression of U1 cassette is driven by its own endogenous promoter.*

## IV.3.3 Studies of the U1+5a-mediated rescue of hFVII expression through Hydrodynamic Injection of vectors

As a preliminary study to assess the U1+5a mediated-rescue of the hFVII expression impaired by the F7 IVS7+5g/a mutation in a short time window we exploited the hydrodynamic injection approach, in which the naked DNA is injected at constant rate into the mouse blood stream. This study phase was aimed at evaluating the relationship between the amount of injected DNA and of the rescued hFVII mRNa/protein  and at validating the sensibility of assays to measure the levels of hFVII in the mouse contest. Altogether these data would provide key feasibility elements for the subsequent experiments with the AAV.

The scheme of the experiment was the following:

1) Group of mice injected with pAAV-FVII wt (positive control)

2) Group of mice injected with pAAV-FVII-IVS7+5g/a (they should mimic the patient's phenotype)

3) Group of mice injected with pAAV-FVII-IVS7+5g/a + pAAV-U1+5a (to test correction efficacy)

4) Group of mice injected with pAAV-U1+5a (to test off-target effects of the U1+5a)

5) Group of mice injected with saline (negative control)

| Group | Number of mice | Plasmid | Dose I (µg/gr body mass) | Dose II (µg/gr body mass) |
|---|---|---|---|---|
| 1 | 4 | pAAV-FVII wt | 1 | 2 |
| 2 | 4 | pAAV-FVII-IVS7+5g/a | 1 | 2 |
| 3 | 4 | pAAV-FVII-IVS7+5g/a | 1 | 2 |
|  |  | pAAV-U1+5a | 1,5 | 3 |
| 4 | 4 | pAAV-U1+5a | 1,5 | 3 |
| 5 | 4 | Saline | / | / |

*Table showing the number of mice and relative amount of plasmid for each group*

Two different series of injections were done: one with 1ug of plasmids per gram of mouse mass, and one with double quantity of plasmids per gram of mouse body mass. The molar ratio between the effector, namely the pAAV-U1+5a, and the substrate, namely the pAAV-hAAT-FVII IVS7+5g/a, was maintained 1,5:1 in both experiments. The plasmid DNA solution was injected in 5-8sec through a 36g needle into the lateral tail vein and 48 hours later the plasma samples, withdrawn by retro-orbital bleeding procedure, were collected to assess hFVII by ELISA. in 3,8%

sodium citrate. Mice were then sacrificed to collect livers (to isolate mRNA or to perform immunohistochemistry).


## IV.3.3.1 Studies at the mRNA level


Being a mutation affecting FVII splcing we initially investigated the hFVII mRNA in mouse liver. Due to the high homology between human and mouse FVII, the position of primers used to amplify the cDNA is crucial to selectively amplify only the hFVII. To overcome this problem, we focused our attention on the region spanning exon 6 trough exon 8 in these genes. We aligned these regions from human and murine FVII, considering only the region indicated above (exon 6 – exon 8) (Figure 44).

```
hF7    TTGAATATCCATGTGGAAAAATACCTATTCTAGAAAAAAGAAATGCCAGCAAACCCCAAG
mF7    TTGAGTACCCGTGTGGGAGAATACCTGTTGTAGAAAAAAGAAACTCCAGCAGCCGCCAAG
       ****.** **.*****.*.*******.** ************* ******..* *****

hF7    GCCGAATTGTGGGGGGCAAGGTGTGCCCCAAAGGGGAGTGTCCATGGCAGTCCTGTTGT
mF7    GCCGCATTGTGGGAGGCAACGTGTGCCCCAAAGGGGAGTGTCCATGGCAGGCTGTGCTGA
       ****.********.***** **************************** ** **:

hF7    TGGTGAATGGAGCTCAGTTGTGTGGGGGGGACCCTGATCAACACCATCTGGGTGGTCTCCG
mF7    AAATCAATGGGTTATTGCTGTGTGGGGCCGTCCTGCTGGACGCCAGATGGATAGTGACCG
       :..* *****.   : :* *******    ****.* .**.*** .***.*.** :***

hF7    CGGCCCACTGTTTCGACAAAATCAAGAACTGGAGGAACCTGATCGCGGTGCTGGGCGAGC
mF7    CAGCCCACTGCTTCGATAATATCCGCTACTGGGGAAACATCACAGTGGTGATGGGTGAAC
       *.******** ***** **:***.. :*****.*.***.* * .* ****.**** **.*

hF7    ACGACCTCAGCGAGCACGACGGGGATGAGCAGAGCCGGCGGGTGGCGCAGGTCATCATCC
mF7    ATGACTTCAGTGAGAAGGATGGGGATGAGCAAGTACGACGGGTGACACAGGTCATCATGC
       * *** **** ***.* ** **********.. .**.*****.*.********** *

hF7    CCAGCACGTACGTCCCGGGCACCACCAACCACGACATCGCGCTGCTCCGCCTGCACCAGC
mF7    CCGACAAGTACATCCGCGGCAAGATCAACCACGACATTGCCCTGCTCCGCCTTCACCGGC
       **..**.****.***  ****. * ************ ** *********** ****.**

hF7    CCGTGGTCCTCACTGACCATGTGGTGCCCCTCTGCCTGCCCGAACGGACGTTCTCTGAGA
mF7    CTGTGACCTTCACTGACTACGTGGTGCCCCTGTGTCTGCCTGAAAAGTCCTTCTCCGAGA
       * ***. * ******** * ********** ** ***** ***..*:* ***** ****

hF7    GGACGCTGGCCTTCGTGCGCTTCTCATTGGTCAGCGGCTGGGGCCAGCTGCTGGACCGTG
mF7    ACACCCTAGCCAGAATCCGCTTCTCAAGGGTCAGTGGCTGGGGCCAGCTACTGGACCGTG
       . ** **.***:  .* *********: ****** **************.*********

hF7    GCGCCACGGCCCTGGAGCTCATGGTCCTCAACGTGCCCCGGCTGATGACCCAGGACTGCC
mF7    GTGCCACAGCCCTGGAACTCATGTCCATCGAGGTGCCCCGGCTGATGACCCAGGACTGTC
       * *****.********.****** *.**.* ************************** *

hF7    TGCAGCAGTCACGGAAGGTGGGAGACTCCCCAAATATCACGGAGTACATGTTCTGTGCCG
mF7    TGGAGCACGCCAAGCACAGCTCTAACACCCCCAAAATCACAGAGAACATGTTCTGCGCTG
       ** ****  *...*.* .    :.**:****.**:***** ***:********** ** *

hF7    GCTACTCGGATGGCAGCAAGGACTCCTGCAAGGGGGACAGTGGAGGCCCACATGCCACCC
mF7    GCTACATGGATGGTACCAAGGACGCCTGCAAGGGTGACAGCGGTGGCCCACATGCCACGC
       *****: ****** * ******* ********** ***** ** *************** *

hF7    ACTACCGGGGCACGTGGTACCTGACGGGCATCGTCAGCTGGGGCCAGGGCTGCGCAACCG
mF7    ACTACCATGGCACATGGTATCTGACAGGTGTGGTCAGCTGGGGGGAGGGCTGTGCAGCTA
       ******. *****.***** *****.** . .* ********** ******* ***.* .

hF7    TGGGCCACTTTGGGGTGTACACCAGGGTCTCCCAGTACATCGAGTGGCTGCAAAAGCTCA
mF7    TTGGTCACATTGGGGTGTACACCAGGGTCTCCCAGTACATAGACTGGCTGGTCAGACACA
       * ** ***:*******************************.** ****** ::.*..*:**

hF7    TGCGCTCAGAGAGCCACGCCCAGGAGTCCTCCTGCGAGCCCCATTTCCCTAG
mF7    TGGACTCCAAGCTCCAGGTTGGGGTTTTCCGAC---TCCCACTACTGTAG
       ** .***..*** .*.  :**.** *** .*    **** *:* ***
```

***Figure 44:*** *Human (hF7) and mouse (mF7) FVII alignment using only the genomic region spanning exon 6 to exon 8. Exons are color coded: human exon 6 in dark grey, human exon 7 in dark green, mouse exon 6 in bright grey, mouse exon 7 in bright green, human and mouse exon 8 in white. At the beginning the extremity of human and mouse exon 5 is indicated.*

To design specie-specific primers, we sought in hFVII mRNA for sequences with lower complementarity with mouse FVII, indicated by the number of gap in the alignment (indicated by a white space). As shown in Figure 44, the homology between human and mouse FVII mRNA is very high, reaching 76% for the indicated sequence. To reduce aspecific amplification during the PCR reaction, we designed primers in order to base pair with the human sequence with the lower homology versus the mouse FVII; moreover, the 3' of the primer, essential for the specific amplification of nascent fragment, was located in the region with the lowest homology between human and mouse FVII mRNA (Figure 45). In that way, we should get the maximal capability to selectively amplify the human FVII in mouse liver context.

***Figure 45:*** *Human (hF7) and mouse (mF7) FVII alignment with indicated the primers designed to selectively amplify the human FVII mRNA. The 3' end of each primers lands on a nucleotide (indicated by an empty space in the third line) not conserved between human and mouse mRNA.*

The RNA was isolated from random sections of livers derived from each group of mice injected with the higher amount of DNA. After the quantification of the RNA extracted, we performed the retro-transcription to obtain total cDNA using random primer and 1ug of total RNA. On this cDNA, we performed two PCRs with two different couple of primers, base pairing with human exon 6 and exon 7 or exon 8 respectively. Based on our previous works, through PCR with primers on exon 6 and exon 8 (primers hFVII6F and hFVIIR8 respectively) we should be able to detect every event of splicing, including the exon skipping, the partial intron retention (+37bp) and eventually traces of normal transcript (generally masked by the heteroduplex isoform generated by base pairing between

the exon skipping and the partial intron retention forms). With the inner PCR instead, with primers located in human exon 7 and exon 8 (primers hFVII7F and hFVIIR8 respectively), we should evaluate again the presence of the partial intron retention form, but more intriguingly the presence of normal transcripts. In fact, with these couple of primers we amplified the partial intron retention form and the normal transcript, but not the exon skipping form, so avoiding the heteroduplex isoform formation. In that way, we should be able to evaluate the rescue effect of the engineered U1 at RNA level. We analyzed each group of the hydrodynamic injection experiment and the results are reported in Figure 46.
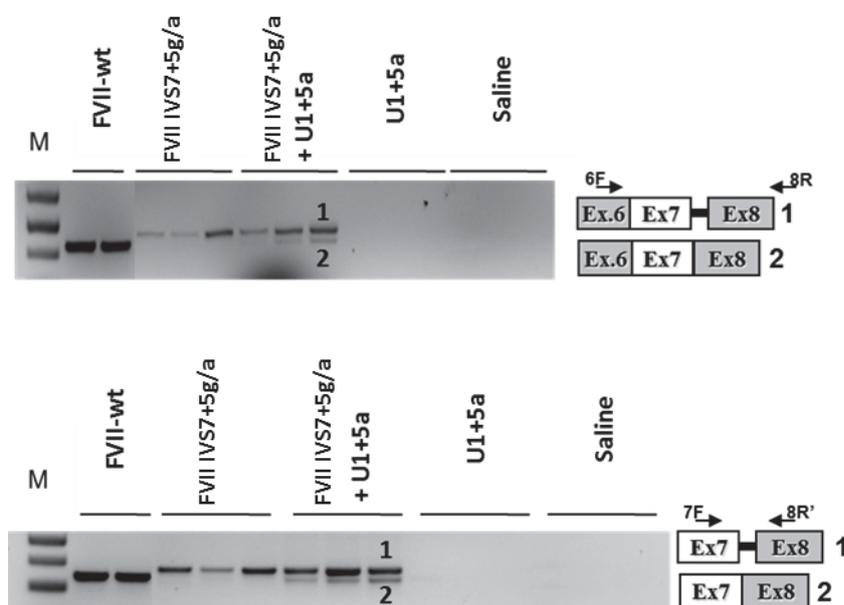


*Figure 46:* *Electrophoresis separation on 3% agarose gel of RT-PCR conducted on RNA from mouse livers after the hydrodynamic injection experiment with primers hFVII6F-hFVIIR8 (upper part) and hFVII7F-hFVIIR8 (lower part). Schemes of transcripts and primers used (arrows) are indicated in the right part of the figure.*

In mice receiving only the pAAV-U1A5 (groups 4) or the saline solution (group 5), and representing the negative controls, no traces of human FVII transcripts were detectable after PCR amplification, indicating that the couples of primers used were specie-specific. In mice receiving the pAAV-hAAT-FVII-wt plasmid (group 1), we observed in both PCRs only a band corresponding to a correctly spliced transcript, as expected. Instead, in mice injected with pAAV-hAAT-FVII IVS7+5g/a plasmid (group 2), we noticed no traces of normal transcript, but only the presence of aberrant transcripts. Surprisingly, we observed only the partial intron retention form, but no traces of exon skipping form.

Considering these results, both PCRs showed the same splicing pattern, with differences only in fragment length. Intriguingly, in group 3, receiving either the plasmid bearing the mutation and the engineered U1 plasmid, we observed with both PCRs the appearance of normal transcripts, originated by a correctly spliced mRNA, indicating that the engineered U1 is able to redirect the spliceosome to the correct splicing junction, favoring the exon definition.

To quantify the rescue obtained upon the utilization of the U1+5a  and verify the presence of normal transcripts in mice receiving the mutated plasmid, we used the denaturing capillary system to analyze fluorescently labeled PCR products. We labeled the previous PCR (performed with primers hFVII6F and hFVIIR8) with an aspecific fluorescent dye and loaded it onto a Experion chip. This denaturing capillary system returns a series of peaks, corresponding to the bands observed on the gel, where their time course is proportional to their length in base pairs. The length is provided comparing the time course of the peak with an internal fluorescent ladder. The area below the peak is proportional to the loaded amount of the fragment, allowing the rough quantification of each band.

Through this system, the amount of the normal transcript in mice co-injected with the mutant FVII construct and the engineered U1 was 26±10% compared to the +37bp form. Noticeably, even in mice injected with the mutated plasmid, some traces of normal transcripts were detectable, reaching 2±1,5% of the aberrant form (Figure 47). This data was in accordance with our previous findings observed *in vitro*, and further demonstrated the association of the mutation with traces of correct splicing and thus the conclusion that the mutation produces a defective 5'ss.
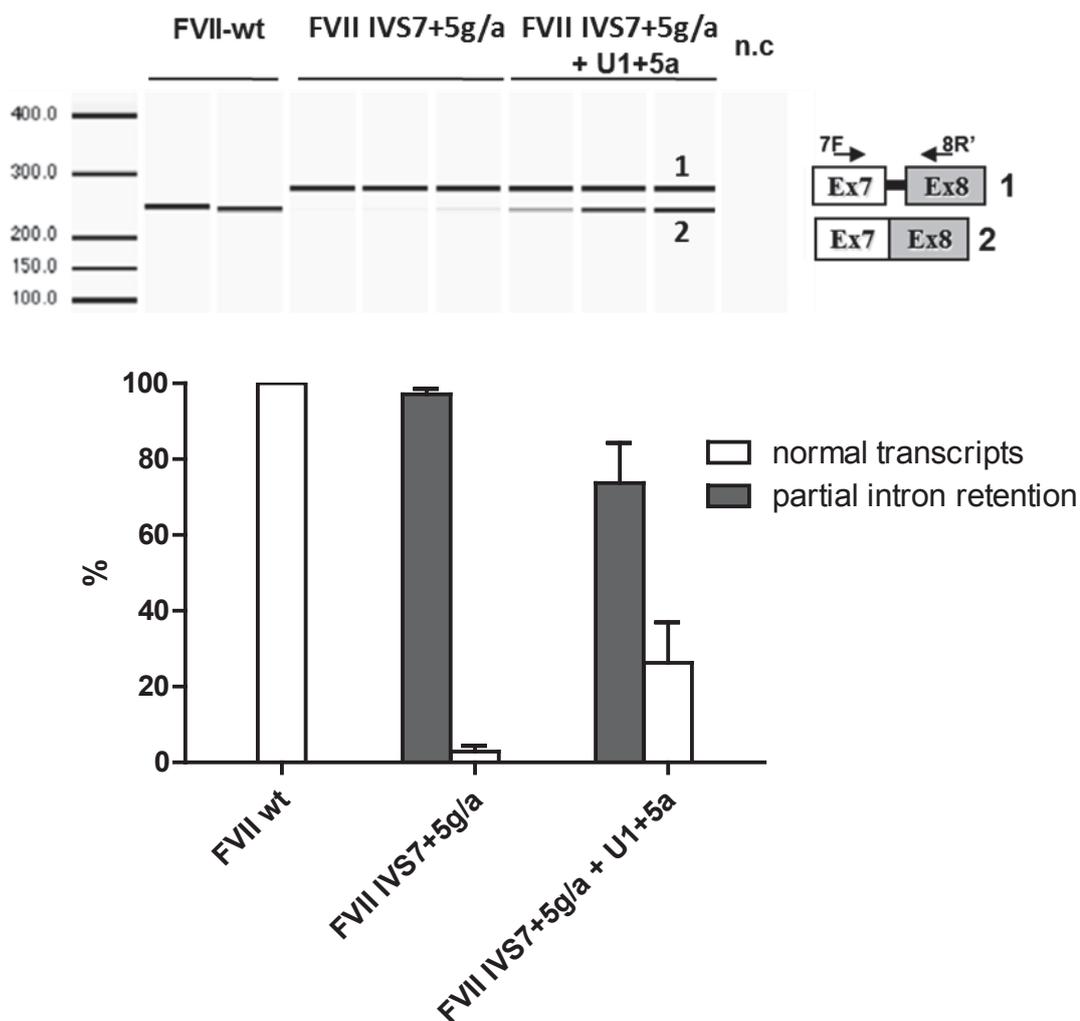
***Figure 47:*** *Separation on a denaturing capillary system (Experion, Automated Electrophoresis Station, BIO-RAD) ) of fluorescently labeled RT-PCR products obtained from total RNA of mice injected with normal (FVII-Wt), mutated alone (FVII IVS7+5g/a) or in combination with the modified U1snRNA. The pherogram of separation, likely an agarose gel separation, is shown on the top of the picture. The scheme of transcripts, and of primers used (7F and 8R), is depicted in the right side of the pherogram. The fragment sizes of the normal and aberrant transcripts were 236 bp and 273 bp, respectively. The percentile of correction, calculated on ratio between the normal transcript or the aberrant form (partial intron retention) over total RNA, is shown in the lower part of the picture.*

## IV.3.3.2 Studies at the FVII protein level in plasma

Upon the injection of the pAAV-FVII wt plasmid, we detected remarkable levels of circulating hFVII in mouse plasma, reaching levels of at least 2µg/ml (2,2±0,4 µg/ml) with the highest dose of plasmid injected, and 1±0,3 µg/ml in mice injected with the lowest dose of DNA (1 µg/gr). At variance, the hFVII levels in plasma from mice injected with the mutant plasmid pAAV-hAAT-FVII IVS7+5g/a (group 2) were undetectable, which is consistent with the circulating FVII levels measured in the homozygous patients. Worth noting that the hFVII was undetectable in mice injected with the pU1+5a only (group 4) or the saline solution (group 5), thus indicating the ability of the ELISA to properly discriminate the hFVII from mouse FVII (Figure 48).

Intriguingly, upon co-injection of the pAAV-hAAT-FVII IVS7+5g/a with the pAAV-U1+5a (group 3), the hFVII protein was clearly appreciable in mouse plasma. In particular, upon injection of the highest dose (2 µg/gr) the hFVII levels were above 100ng/ml in all mice (179±126 ng/ml). Compared to the levels obtained in mice injected with the FVII wt plasmid, the correction obtained after the injection of the U1+5a was 4,9% (1µg/gr of plasmids) and 8,4% (2 µg/gr of plasmids) respectively, thus indicating a robust and dose-dependent rescue of FVII expression (Figure 48). It must be notices that, if obtained in FVII deficient patients, the correction observed would be well above the therapeutic threshold (2%).
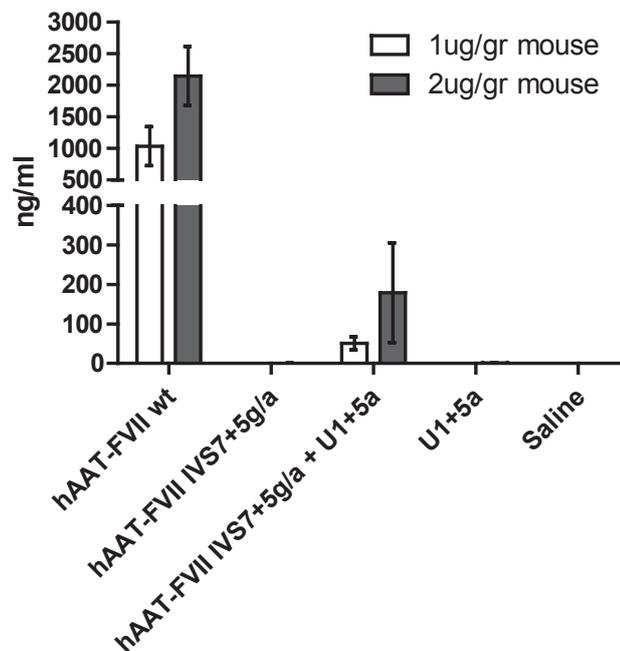


**Figure 48:** *Histogram showing the amount of human FVII (ng/ml) detected in mouse plasma during the hydrodynamic experiment.*

To evaluate the activity of hFVII expressed in mouse plasma, we also tried to detect differences in FVII coagulant function through Protrombin Time assay. To this purpose, mouse plasma samples, rich in mouse FVII but with variable levels of human FVII (when present), was added to human FVII deficient plasma. Tromboplastin (phospholipids, Calcium ions and membranes) was then used to trigger coagulation and record times to clot formation (Figure 49).
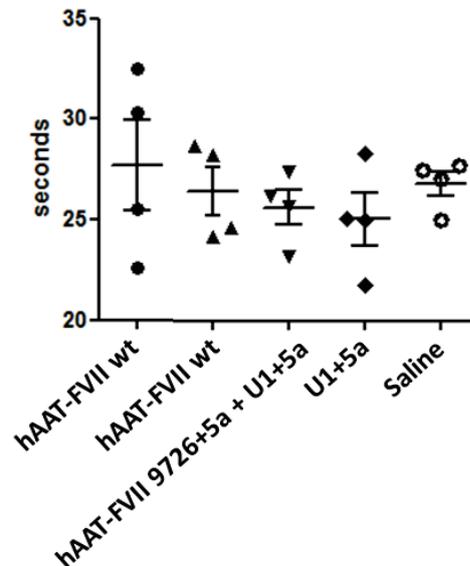


**Figure 49:** *FVII coagulant activity in mouse plasma evaluated by Protrombin Time assay.*

As shown in Figure 49, no significant differences were appreciable upon addition of the plasma samples from mice treated with the pU1+5a. This was expected since the amount of the endogeneous murine FVII masks the contribution of hFVII upon correction. Based on this result, the PT-based assay was no longer used to evaluate the impact of the U1+5a.

## IV.3.3.3 Studies at the intracellular FVII protein level in mouse liver

Another way to detect the rescue of human FVII expression by the modified U1+5a is to directly assess the hFVII protein into mouse liver sections. For this purpose, we fixed livers withdrawn from mice in 4% paraformaldehyde overnight at 4°C. After inclusion into OCT compound we prepared various slices 8nm thick from each livers. The detection of human protein in mouse liver is a big issue due the high homology among proteins of these species. For this reason, we optimized a standard Immunohistochemistry protocol to selectively detect hFVII in mouse liver. After incubation with anti-mouse IgG fragment H+L and Fetal sheep serum to bind and mask endogenous mouse IgG, accountable for non-specificity signal, a standard direct Immunohistochemistry

protocol was applied on liver sections. As antibody we used an FITC-labelled antibody specific for human FVII and raised in sheep. The cells were labelled also with DAPI, to detect nuclei.

Once the immunohistochemistry protocol was optimized, we were able to distinguish the expression of the hFVII in mouse hepatocytes. In fact, cells expressing hFVII were labelled in green, while non-expressing cells were colourless (with nuclei only marked in blue by DAPI). Green cells were visible only in mice injected with the pAAV-hAAT FVII-wt plasmid (group 1) or in mice belonging to group 3 (injected with the pAAV-hAAT FVII IVS7+5g/a and the pAAV-U1+5a plasmids). Only blue cells were detectable in sections from mice injected with the mutated plasmid alone, corroborating the very low levels of hFVII driven by this construct (Figure 50).
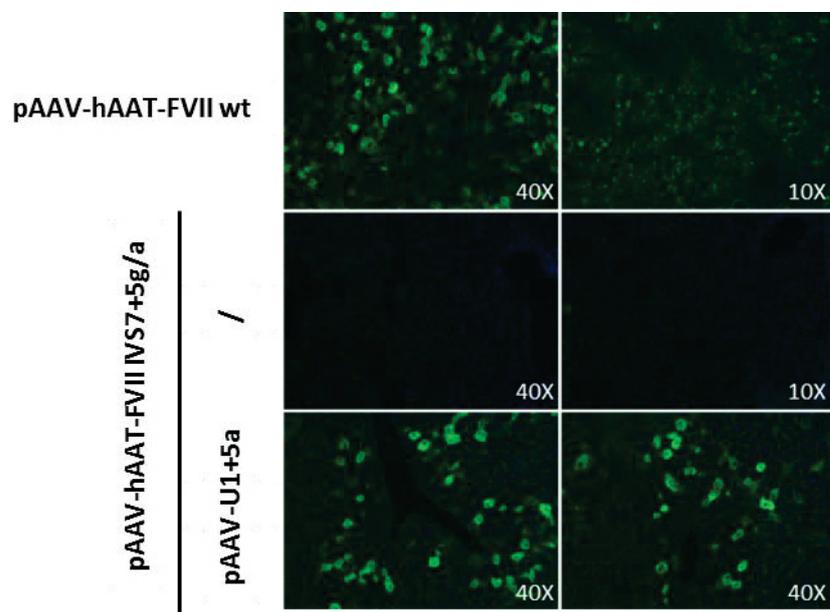


**Figure 50:** *Detection of human FVII protein in mouse live sections through Immunohistochemistry approach. Two liver sections are shown for each group of mice injected. Positive cells are marked in green (blue dye denotes cell nuclei). Magnification index in indicated in the low right side of each pictures.*

It must be noticed that not all sections were positive for hFVII, even in mice injected with the pAAV-hAAT FVII-wt. This is due to the fact that only restricted regions of the liver acquired the plasmid. Due to this bias, we cautiously exploited the data from immunohistochemistry to assess the rescue of hFVII expression by the engineered U1.

## IV.3.4 Study of the U1+5a-mediated rescue of hFVII expression by AAV vectors

These results with the hydrodynamic injection prompted us to further investigate the U1+5a-mediated rescue of FVII expression in the mouse model by exploiting adeno-associated viral (AAV) vectors that, at variance from the previous approach, allow the assessment of the rescue in a prolonged time window.

These vectors are emerging as powerful tools for the gene therapy of human diseases (High 2011; Asokan, Schaffer et al. 2012; Hu and Lipshutz 2012; Ortolano, Spuch et al. 2012) for a number of reasons. Among them, their ability to reach long term expression in large immune-competent animals due to the minimal activation of innate immune pathways and to efficiently transduce non dividing cells including muscle and liver. A drawback of this approach is the limited capacity of the capsid to accommodate large genome (no more than 4.7kbp), which limits the size of the transgene. However, the U1 coding sequence is very small thus rendering the AAV an ideal delivery system.

To assess the U1+5a-mediated rescue of hFVII expression impaired by the FVII IVS7+5g/a *in vivo* and overcome the problem related to risk of the development of neutralizing antibodies, we planned to exploit two different AAV serotypes for the delivery of the hFVII mutated construct, the "substrate", and of the engineered U1+5a, the "effector". Since the target organ is liver, we decided to use AAV2 and AAV8 serotypes to carry the FVII mutation and the engineered U1 plasmids, respectively. The AAV8 is very efficient in transducing hepatocytes. On the other hand, the AAV2 presents natural tropism towards hepatocytes but also skeletal muscles, neurons and vascular smooth muscle cells. However, the use of a strong and liver specific promoter driving the expression of the hFVII transgene makes the off-target hFVII expression very unlikely.

The creation of the AAV viruses was done using the triple transfection protocol, as indicated in the material and methods section. Briefly, HEK293 cells were transfected with a solution of three plasmids, two of them harboring the structural genes of AAV and responsible for the serotype, and the third being the plasmids exploited in the Hydrodynamic Injection protocol. The latters carry the hFVII or U1+5a expression cassette within the ITS sequences that are needed for packaging of plasmids into the nascent capsid. After amplification and purification of the viral stock, viruses were stored at -80°C in PBS/10% glucose solution.

## IV.3.4.1 Co-injection of AAV2-hFVII and the AAV8-U1+5a

The initial strategy chosen to assess the U1+5a-mediated rescue of hFVII expression was to co-infect C57BL/6J mice with the AAV2-hFVII and the AAV8-U1+5a.

We set the initial amount of viruses to be injected into the mouse to $1,2*10^{12}$vg/mouse, a dose successfully used for other gene therapy experiments by prof. Arruda and co-workers (Margaritis, Arruda et al. 2004). In the initial experiments, three cohorts of mice were treated as follow:

| Group | Number of mice | Vector | Dose (vg/mouse) |
|-------|----------------|--------|-----------------|
| 1 | 7 | AAV2-hAAT-FVII IVS7+5g/a | $1,2*10^{12}$ |
| 2 | 8 | AAV2-hAAT-FVII IVS7+5g/a | $1,2*10^{12}$ |
|   |   | AAV8-U1+5a | $1,2*10^{12}$ |
| 3 | 7 | AAV2-hAAT-FVII IVS7+5g/a | $2,4*10^{12}$ |
|   |   | AAV8-U1+5a | $2,4*10^{12}$ |

*Table showing the number of mice and relative amount of plasmid for each group*

The AAV was injected in the lateral tail vein in 5-8 seconds using a 36g needle. We carefully examined the animals throughout the period post-injection and collected blood from the retro-orbital sinus each week post-injection.

One week post-injection all mice looked normal, but 14 days later some mice in group 2 and 3 started to look feeble and sick. Between the second and third week post-injection, all mice in group 2 and 3 died in a dose-dependent manner (Figure 51).
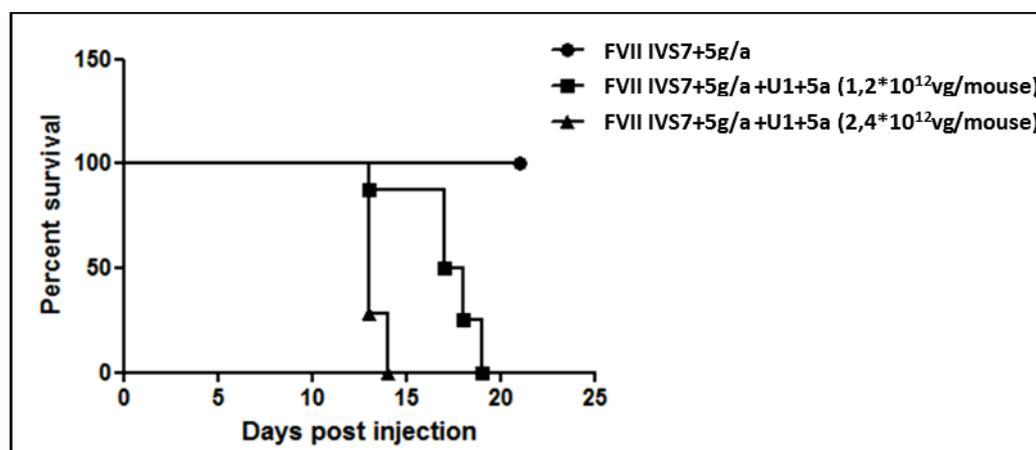


**Figure 51:** *Survival curve of mice injected with $1,2*10^{12}$ vg/mouse (square dots) and $2,4*10^{12}$ vg/mouse (triangle dots) of AAV8-U1+5a compared to mice receiving only AAV2-FVII IVS7+5g/a vectors (circle dot).*

Animal autopsy revealed the all organs looked normal and fine, but livers were notably altered, with considerably signs of necrosis (Figure 52). Another indication of liver damage was the icteric sera withdrawn from mice, probably due to increased level of circulating bilirubin.



**Figure 52:** *Pictures of mice livers withdrawn from a mouse injected with 1,2\*10$^{12}$ vg/mouse of AAV2-FVII IVS7+5g/a (left) or AAV8-U1+5a (right). It is evident the different status between two livers. Mouse receiving AAV8-U1+5a from which the liver is shown looked sick and feeble just before its sacrifice.*

It is worth noting that mice started dying around two weeks post-injection, when the expression of the modified U1 reached the maximum. In fact, the expression of a gene delivered through AAV2 serotype is expected to reach the maximum after four weeks post injection, while the peak of expression occurs two weeks post-injection with AAV8 serotype.

ELISA assay for hFVII detection in mouse plasma was negative in all samples. The undetectable hFVII levels in mice belonging to group 1 indicated that the F7 IVS7+5g/a mutation impair hFVII expression, in accordance with the data from homozygous patients.

The undetectable hFVII levels upon co-injection with the AAV8-U1+5a did not permit the demonstration of the U1-mediated correction observed with hydrodynamic injection studies. A number of reasons might explain this results. As a matter of fact, the shift in the expression of the FVII transgene in AAV2 (peak at 4 weeks) as compared to that of the U1+5a in AAV8 (peak at two weeks) might have vanished the U1+5a effect being too low the expression of the substrate. Moreover, the detrimental effect of the U1+5a on physiology of mouse liver, the target tissue for out therapeutic approach.

## IV.3.4.2 Sincronization of AAV2-hFVII and AAV8-U1+5a expression

The above reported data led us to modify the experimental protocol and to seek for the best dose and timing. The rationale was that, like other drugs, the U1+5a could be toxic at high doses but not at low doses, and be still effective. Moreover, we tried to avoid the shift in the expression of our transgenes and to synchronize their expression peaks. Therefore, the mice expressing the hFVII-IVS7+5g/a minigene from the previous experiment (group 1), injected 4 weeks before and expected to have reached the maximum expression of the substrate, were split in two groups (as indicated in the table below). These two sub-groups were infected with two different AAV8-U1+5a doses, half ($6*10^{11}$vg/mouse) and one-tenth ($1,2*10^{11}$ vg/mouse) of the previous amount that was found to be lethal.

| Group | Number of mice | Injected | Dose (vg/mouse) |
|---|---|---|---|
| 1-a | 4 | AAV8-U1+5a | $6*10^{11}$ |
| 1-b | 3 | AAV8-U1+5a | $1,2*10^{11}$ |

*Table showing the number of mice and relative amount of plasmid for each group*

In these conditions, all mice looked fine until three weeks post-injection. After that, mice receiving $6*10^{11}$vg/mouse of AAV8-U1+5a started to died and livers (Figure 53), upon autopsy, were likely the previous ones, with evident traces of necrosis.
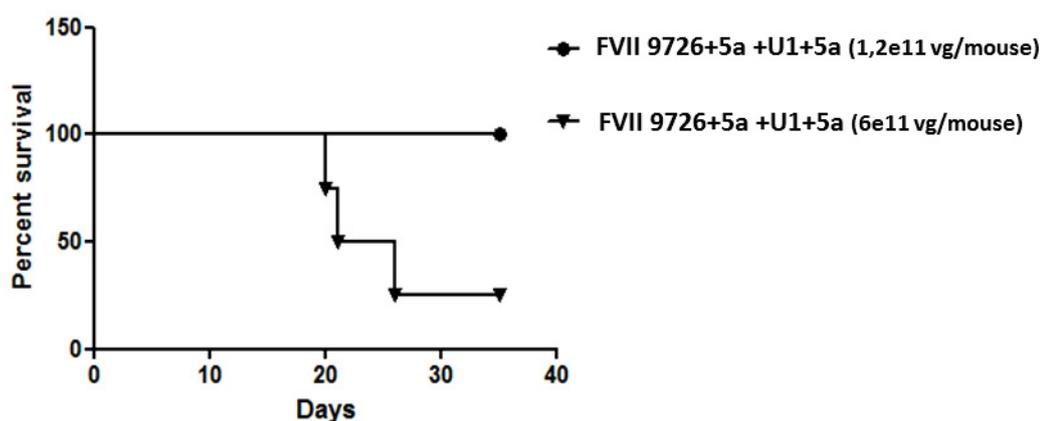


**Figure 53:** *Survival curve of mice injected with $1,2*10^{11}$ vg/mouse (circle dots) and $6*10^{11}$ vg/mouse (triangle dots)of AAV8-U1+5a.*

When the hFVII specific ELISA was performed in mouse plasma we detected in all mice the presence of hFVII (Figure 54), a finding that, taking into account the undetectable hFVII levels in mice injected with the AAV2-hFVII alone, provide experimental evidence for the U1+5a mediated rescue of hFVII expression. The highest levels were observed in mice receiving the highest dose of AAV8-U1+5a ($6*10^{11}$vg/mouse), reaching 30ng/ml of hFVII. One mouse showed a decrease in FVII antigen levels because its blood was withdrawn the day before its death, meaning that its liver was highly damaged. Notably, even mice injected with the lowest dose of AAV8-U1+5a ($1,2*10^{11}$ vg/mouse) showed a correction, with the FVII antigen levels growing up to ~5ng/ml. This U1+5a mediated correction was sustained up to 6 weeks after injection without any evidence of toxic effect on mice health.
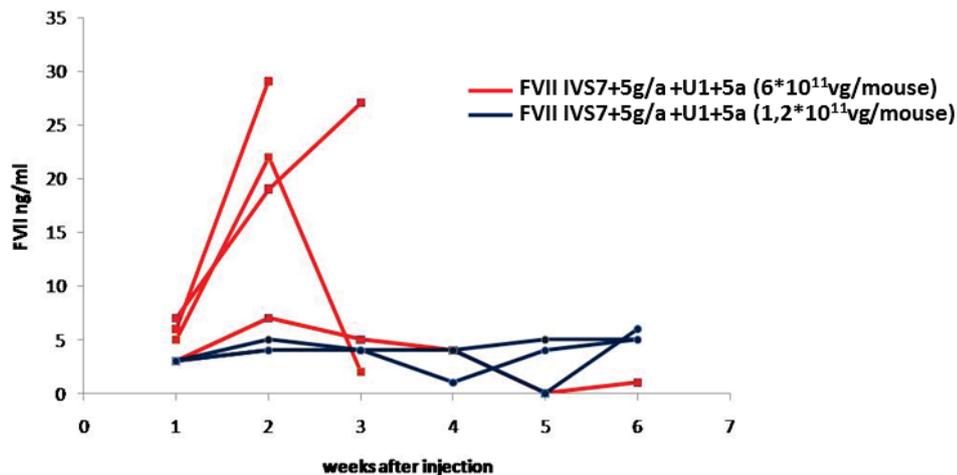


***Figure 54:*** *Human FVII antigen levels evaluated through a selective human FVII ELISA.*
*Mice receiving $6*10^{11}$ vg/mouse and $1,2*10^{11}$ vg/mouse of AAV8-U1+5a are indicated by the red and blue line, respectively.*

To assess the effect on hFVII splicing pattern, we sacrificed all mice to isolate livers. RNA was isolated from random portions from entire livers each livers and retro-transcribed to cDNA using random primer. The retrotranscription and amplification of hFVII through hFVIIex6F-hFVIIR8 primers followed by denaturing capillary electrophoresis revealed the appearance of the correct transcript in all treated mice thus witnessing the rescue at the mRNA level.

In mice injected with the highest dose of AAV8-U1+5a ($6*10^{11}$ vg/mouse) the correct transcripts reached 16±3% of the total transcripts while with the lowest dose ($1,2*10^{11}$ vg/mouse) the correct transcripts represented the 4±0,5% (Figure 55). However, for reason not clarified yet, when each

single mouse was considered, we did not observe a clear relationship between the rescued levels of correct hFVII mRNA forms and protein in plasma.

Taking into account that the levels of correct hFVII transcripts associated to the IVS7+5g/a mutation are barely detectable with this experimental system, this observation demonstrated that the correction was dose-dependent at the mRNA level too.
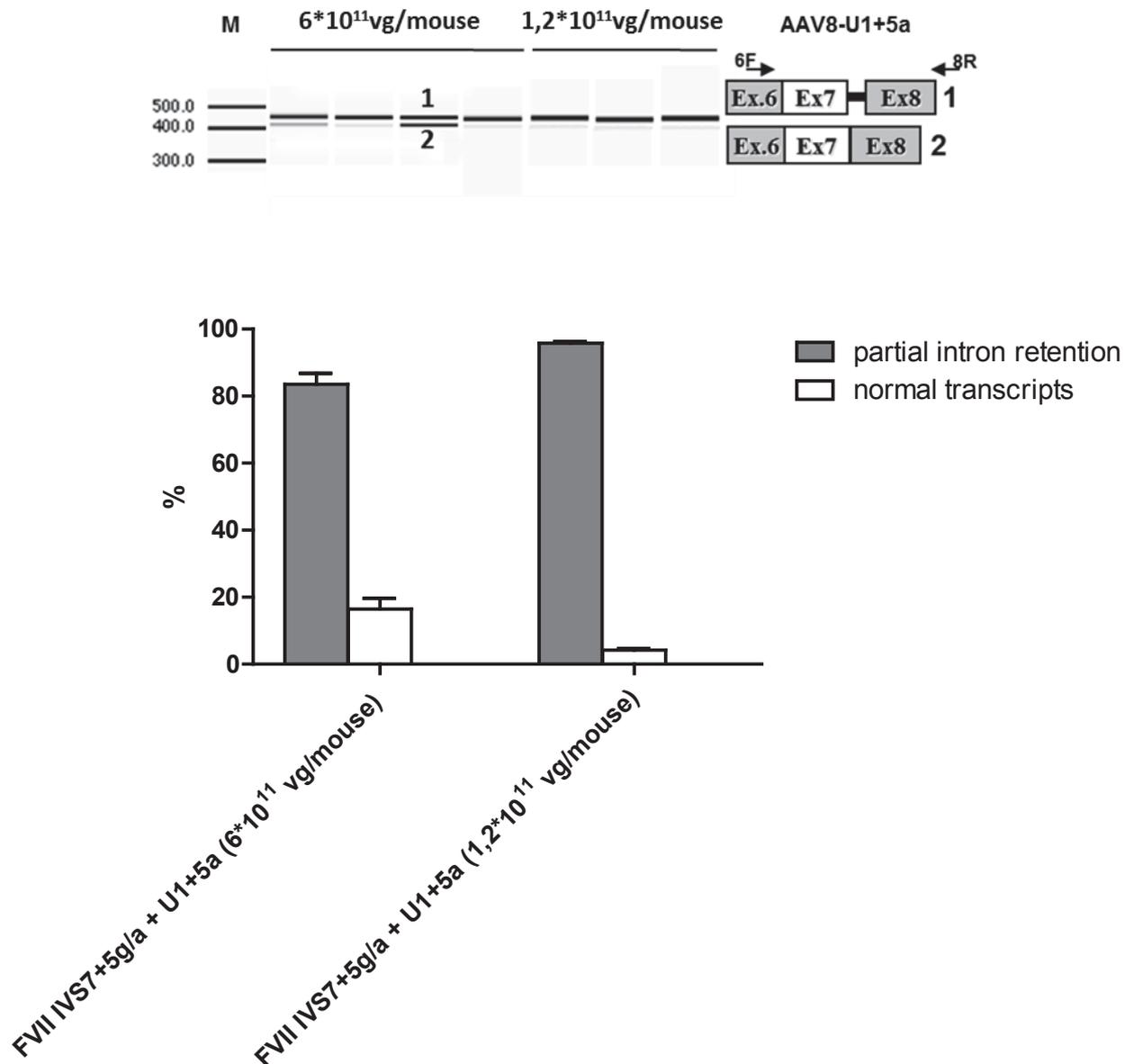


***Figure 55:*** *Separation on a denaturing capillary system (Experion, Automoted Electrophoresis Station, BIO-RAD) ) of fluorescently labeled RT-PCR products obtained from total RNA extracted from mice livers injected with different amount of AAV8-U1+5a vectors . The pherogram of separation, likely an agarose gel separation, is shown on the top of the picture. The scheme of transcripts, and of primers used (6F and 8R), is depicted in the right side of the pherogram. The observed correction is shown below the pherogram as percentile between the normal transcript and the aberrant form (partial intron retention).*

On livers sections we also performed the IHC protocol to assess the expression of hFVII in hepatocytes. Since all mice have been injected and treated with the U1+5a vector, we used a liver from an un-injected mouse as negative control. After the staining with an anti-hFVII FITC labeled antibody, we observed some positive cells ( green cells ) for human FVII expression in sections obtained by treated mice (Figure 56). As observed before, not all sections were positive and, in those positive, only some regions showed the expression of hFVII.
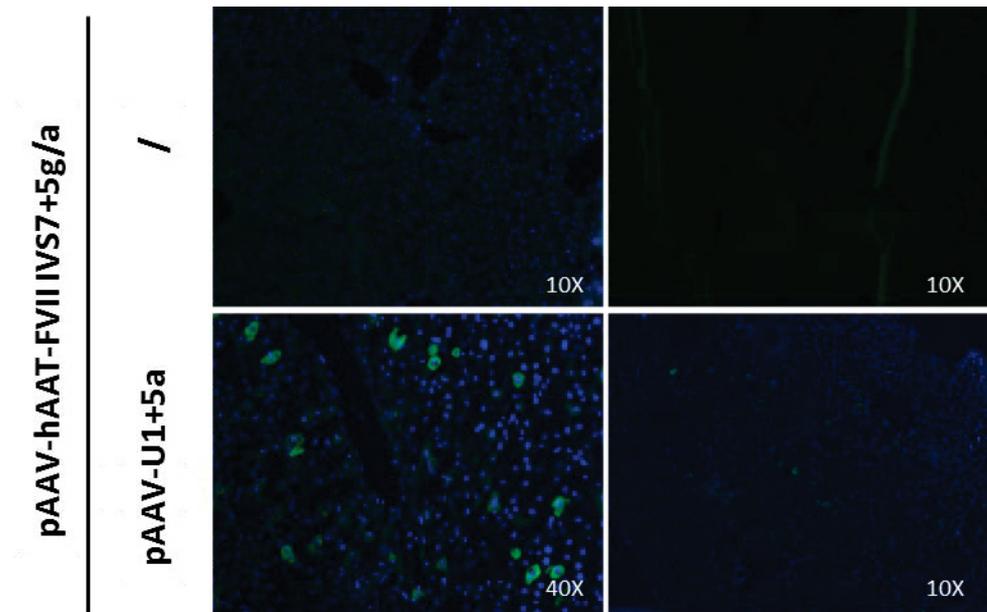


*Figure 56: Detection of hFVII in mouse liver sections through Immunohistochemistry approach. Two liver sections are shown for each group of mice injected. Positive cells are marked in green (blue dye denotes cell nuclei). Magnification index in indicated in the low right side of each pictures.*

Usually, expression of exogenous proteins in immunocompetent animals (in this case human protein in mouse contest) stimulates a strong immune response, through the development of specific antibodies which impair protein function or antigen level. We sought for the presence of these anti-hFVII antibodies in mice treated with the AAV8-U1+5a and showing the presence of human FVII in plasma. For this purpose, we developed a specific ELISA able to detect mouse anti human FVII IgG. Basically, we coated the plate with purified human FVII and incubated it with mouse plasma samples. Addition of a secondary anti-mouse Ig antibody conjugated with horseradish peroxidase allowed the detection of minimal amount of mouse anti human FVII IgG antibodies (ELISA threshold: ~ 30 ng/ml).

We analyzed results in two different ways:

- all mice were considered as single group to detect global changes.

- each group of injection is considered to distinguish mice receiving the high or low dose of AAV8-U1+5a.

In both analysis we noticed a little, but not significative, increase in the presence of mouse anti hFVII IgG, with levels that were always below the threshold (500ng/ml) cosidered to be consistent with a immune response (Figure 57). This result could be due, at least in part, to the high homology between human and mouse FVII protein (76%).

Thaken toghether, these results shown that, when human FVII is expressed in mouse, no strong immune response against that immunogen is detectable up to four weeks (Figure 57).
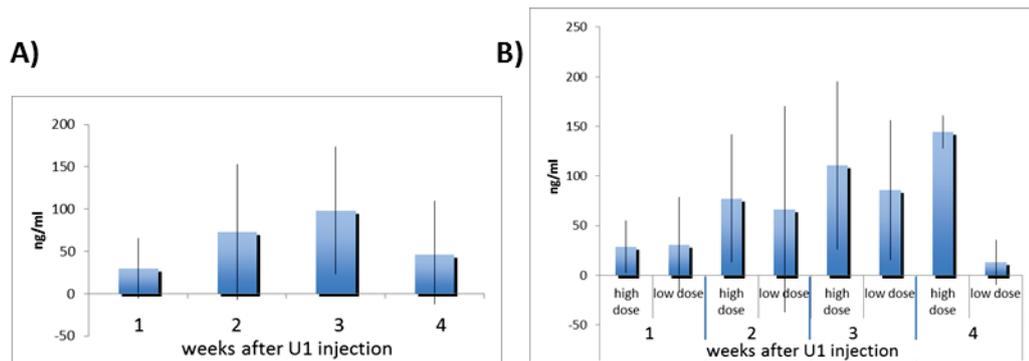


*Figure 57:* *Evaluation of development of mouse anti human FVII IgG in mice showing a reduced levels of human FVII antigen than those observed previously in mice injected with equal amount of AAV8-U1+5a. Two kinds of analysis are shown: in A) all mice were grouped as one single group; in B) mice are considered following the amount of AAV8-U1+5a injected.*

## IV.3.4.3 Effects of consecutive administrations of AAV8-U1+5a.

The possibility of re-administering the U1+5a to substain the correction in a long term prospective is an attractive approach. To verify this hypothesis, we planned to inject new mice with a very low dose of AAV8-U1+5a and after one week to re-administer it to reach the dose that showed the best rescue ($6*10^{11}$ and $1,2*10^{11}$ vg/mouse respectively).

| Group | Number of mice | Vector | I Dose (vg/mouse) | II Dose (vg/mouse) | Final dose (vg/mouse) |
|---|---|---|---|---|---|
| 2-a | 7 | AAV2-hAAT-FVII IVS7+5g/a | $1,2*10^{12}$ | / | $1,2*10^{12}$ |
| | | AAV8-U1+5a | $6*10^{10}$ | $5,4*10^{11}$ | $6*10^{11}$ |
| 2-b | 7 | AAV2-hAAT-FVII IVS7+5g/a | $1,2*10^{12}$ | / | $1,2*10^{12}$ |
| | | AAV8-U1+5a | $2,4*10^{10}$ | $9,6*10^{10}$ | $1,2*10^{12}$ |

*Table showing the number of mice and relative amount of plasmid for each group*

During the first injection, we injected AAV2-FVII IVS7+5g/a ($1,2*10^{12}$ vg/mouse) and AAV8-U1+5a at $6*10^{10}$ (group 2-a) and $2,4*10^{10}$ vg/mouse (group 2-b) respectively. One week post injection, we re-administered only the AAV8-U1+5a to reach the final concentration of $6*10^{11}$ and $1,2*10^{11}$ vg/mouse respectively, the same levels that showed a correction at protein level. We observed that no mice died during the observation period (4 weeks), differently to what observed with mice injected directly with $6*10^{11}$ vg/mouse (Group 1.a). At the protein level, we observed a correction in all mice injected. No significant differences were obtained with $1.2*10^{11}$ vg/mouse either through single dose or consecutive injections (4-5ng/ml). Differently, the peak level obtained with $6*10^{11}$ vg/mouse through consecutive injections (3-4ng/ml) was remarkably lower than those previously observed with a single dose injection (30ng/ml) (Figure 58).
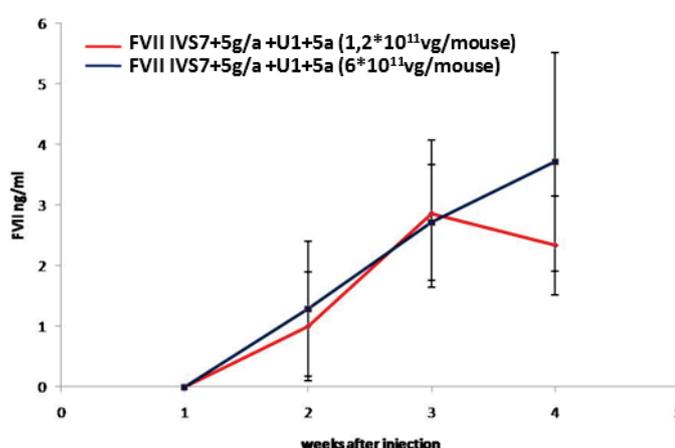


***Figure 58:*** *Human FVII antigen levels evaluated through a selective human FVII ELISA.*
*Mice receiving $6*10^{11}$ vg/mouse and $1,2*10^{11}$ vg/mouse of AAV8-U1+5a in two subsequent administrations are indicated by the red and blue line, respectively.*

The absence of death in mice and the reduction of hFVII level in plasma after treatment with the U1+5a at the highest dose is probably due to the development of antibodies against capsid protein of vector AAV8, with a neutralizing effect towards the re-injected dose of viruses. This should promote the clearance of injected viruses from blood stream and prevent the liver transduction. Since the majority of viruses injected has been administered with the second injection, the effect on correction of neutralizing antibodies is maximal, reducing the overall correction observed.

It must be noticed that the mice with group 2.b produced appreciable amounts of hFVII even upon injection of the first dose of $2,4*10^{10}$ vg/mouse, which corresponds to one fifth of the non-lethal but effective AAV dose ($1.2*10^{11}$vg/mouse) (Figure 58). This further indicate the ability of the U1+5a to restore hFVII expression *in vivo*.

At the RNA level, we isolated livers from mice after four weeks post-injection and extracted total RNA. After the retro-transcription of RNA to cDNA through random examer primers, we performed a PCR with primer hFVIIex6F and hFVIIR1ex8 on the five mice receiving the higher dose of AAV8-U1+5a. We observed the presence of the +37bp aberrant form and, to a less extent, of normal processed messenger, resulting from the redirection of spliceosome on the mutated donor splice site (Figure 59). The quantification of the wt form showed that the correction obtained through consecutive injection of AAV8-U1+5a, and reaching an AAV dose of $6*10^{11}$ vg/mouse, was lower than that observed with a single administration (4,7±0,7% of correction compared to 16%), due to the development of AAV capsid neutralizing antibodies.
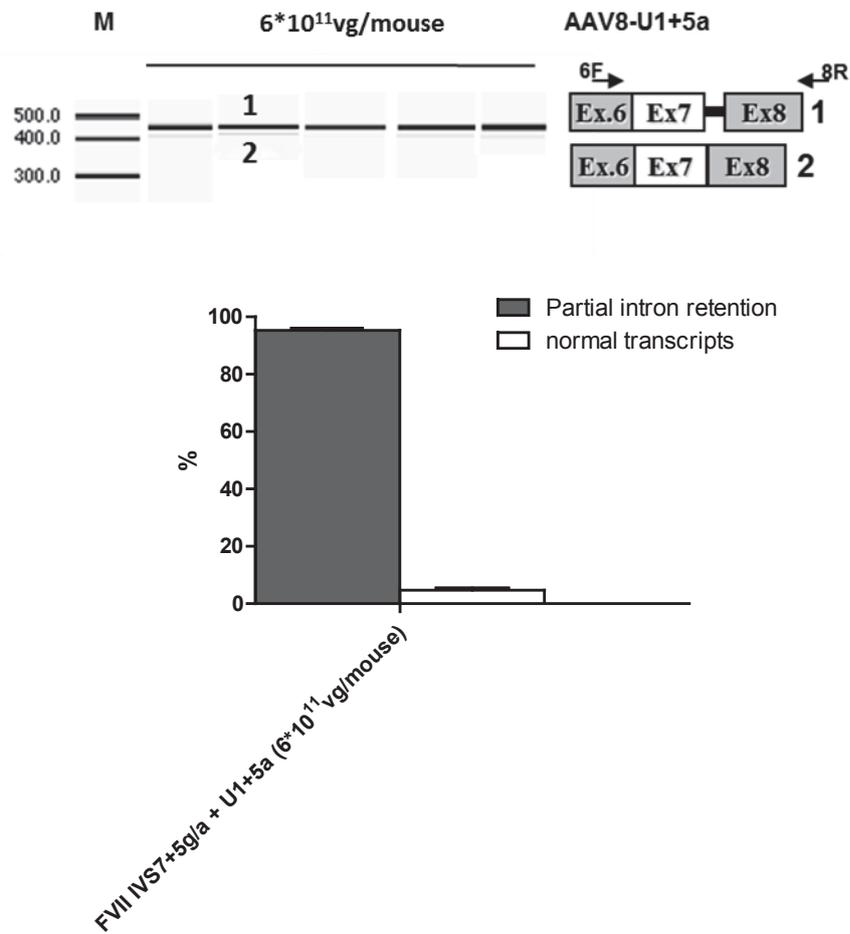
*Figure 59:* *Separation on a denaturing capillary system (Experion, Automoted Electrophoresis Station, BIO-RAD) ) of fluorescently labeled RT-PCR products obtained from total RNA extracted from mice livers injected with different amount of AAV8-U1+5a vectors . The pherogram of separation, likely an agarose gel separation, is shown on the top of the picture. The scheme of transcripts, and of primers used (6F and 8R), is depicted in the right side of the pherogram. The observed correction is shown below the pherogram as percentile between the normal transcript and the aberrant form (partial intron retention).*

## IV.3.4.4 Effect of "substrate" dose:

The rescue effect of modified U1 occurs only when the "substrate", or so called the messenger RNA affected by splicing mutation, and the "effector", in this case the engineered U1snRNA, are located in the same cell at the same time. Through our two vector system this is not guarantee, thus leading to an underestimate of the correction efficiency.

Being the FVII IVS7+5g/a and the U1+5a packaged in two different serotype, with different liver tropism and liver transduction efficacy, the rescue effect is directly determined by the number of cells transduced by the AAV2-FVII IVS7+5g/a. When the number of cells transduced by the substrate is high, the rescue event mediated by the engineered U1 seems likely to occur. With this as background, we designed an experiment in which 14 mice were injected with a 5 fold increased amount of the AAV2-FVII IVS7+5g/a ($6*10^{12}$ vg/mouse) and treated with a non-lethal dose of U1+5a ($1,2*10^{11}$ vg/mouse).

| Group | Number of mice | Vector | I Dose (vg/mouse) |
|---|---|---|---|
| 3-a | 14 | AAV2-hAAT-FVII IVS7+5g/a | $6*10^{12}$ |
| | | AAV8-U1+5a | $1,2*10^{11}$ |

*Table showing the number of mice and relative amount of plasmid for each group*

On plasma withdrew at second and fourth weeks post injection, we performed the specific ELISA for hFVII expression. We noticed that some mice showed high level of human FVII protein (13ng/ml of human FVII) at 2 weeks post-injection and that correction was almost maintained up to four weeks post injection, with a peak of expression up to 26ng/ml of hFVII protein (Figure 60).
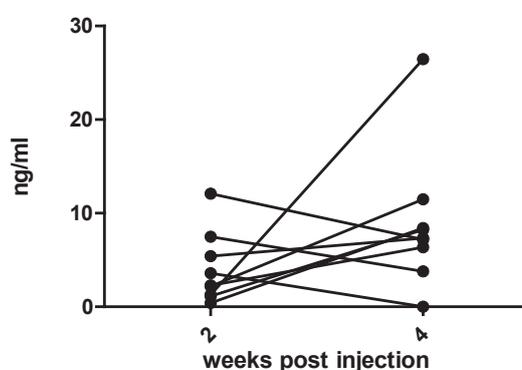


***Figure 60:*** *Determination of human FVII antigen level at 2 and 4 weeks post injection in mice injected with an increased amount of AAV2-FVII IVS7+5g/a vectors.*

We compared this data with those obtained by the injection of mice with the same amount of AAV8-U1+5a ($1,2*10^{11}$ vg/mouse), but with a lower dose of AAV2-hAAT-FVII IVS7+5g/a ($1,2*10^{11}$ vg/mouse). Although obtained in a limited cohort of animals, and thus not reaching the statistical power, the rescue effect appeared more relevant when mice were injected with a higher dose of "substrate" (mutated construct) (8ng/ml compared to 3ng/ml ). The difference became more marked four weeks post-injection, where both transgene have the peak of expression (Figure 61).
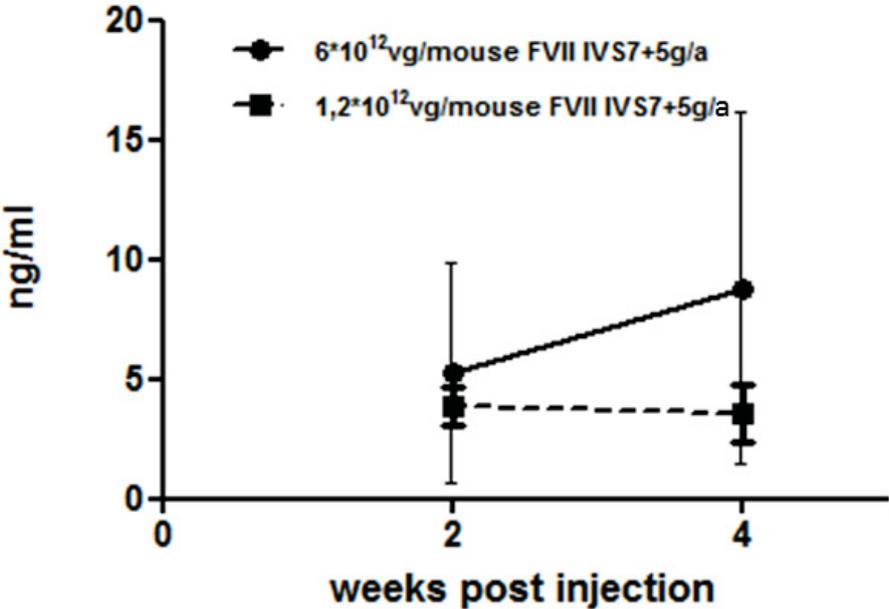


**Figure 61:** *Comparative analysis of human FVII antigen level in mice injected with the same amount of AAV8-U1+5a (1,2e11vg/mouse) but receiving high (6e12vg/mouse, circle dots) o low (1,2e12 vg/mouse, square dots) of AAV2-hAAT-FVII IVS7+5g/a.*

# IV.4 Toxicity

## IV.4.1 Effects of the U1+5a on cell viability and growth

Although the ability of modified U1snRNA to rescue correct splicing has been previous investigated in other human disease model, no attempts have been so far conducted *in vivo*. Our data with the AAV-U1+5a showed that the injection of doses higher than 1,2e11 vg/mouse are lethal, probably because of liver failure.

To provide elements to explain this observation, we went back to cellular models that in our case were represented by cultured mouse hepatocytes (mouse hepatoma cell line, Hepa 1-6). These cells were therefore transfected with the U1+5a, to mimic the treatment conditions, and with the U1-wt or an empty plasmid, to provide negative controls. The effects were subsequently assessed by different approaches to evaluate i) cell grow rate, ii) apoptosis and/or necrosis and iii) cell-cycle.

## IV.4.2 Analysis of cellular grow rate

The MTT assay is a colorimetric assays for measuring the activity of enzymes that reduce MTT to formazan dyes, giving a purple color. It has been developed to assess the viability (cell counting) and the proliferation of cells (cell culture assays). We used it to determine cytotoxicity of modified U1 snRNA as indirect evaluation of cell viability, since engineered U1snRNAs could stimulate or inhibit cell growth through mechanisms not yet defined. However it is important keep in mind that this assay has a drawback : the formation of reduced MTT depends on the metabolic activity of cell, so sometimes different results can be obtain even if number of viable cells is constant.  To overcome this limitation, we planned to compared results obtained from cells transfected with the molar excess (1,5X) of U1+5a with those obtained from transfected cells with the same amount of U1-wt or pUC18, both as negative control. Moreover, we monitored cell viability at 24 and 48 hours after transfection, to obtain a time-response curve.

As shown in Figure 62, we did not detect significant changes in cell viability in the different conditions, even if a higher cell number (cell viability) is noticed (due to replication of cells ) at 48 hours after transfection.
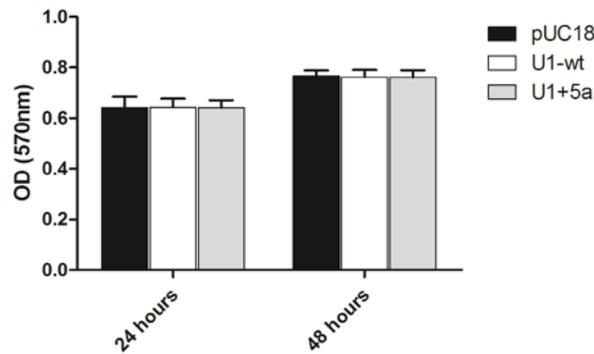
**Figure 62:** *Cellular grow rate evaluated through MTT assay. The adsorption at 570nm is proportional to the number of cells present into the well. Cells were transfected with pUC18 (black bars), pU1-Wt (white bars) or pU1+5a (grey bars) respectively. Results of experiment are expressed as means ± SD, based on at least three independent experiments.*

## IV.4.3 Evaluation of Apoptosis and Necrosis

Apoptosis, or programmed cell death, plays a fundamental role in many normal biological processes. It can be induced by various stimuli that all produce the same result: systematic and deliberate cell death. One method for studying apoptosis is based on detection of changes in the position of phosphatidylserine (PS) in the cell membrane. In non-apoptotic cells, most PS molecules are localized at the inner layer of the plasma membrane, but soon after inducing apoptosis, PS redistributes to the outer layer of the membrane, and becomes exposed to the extracellular environment. PS translocation precedes other apoptotic events, thus allowing early detection of apoptosis, discerning early apoptosis from late apoptosis. Exposed PS can be easily detected with annexin V, a 35.8-kDa protein that has a strong affinity for PS. Moreover, the kit allows the determination of the number of cells in necrosis. In fact the kit contains Propidium Iodide, a fluorescent vital dye that stains DNA. It does not cross the plasma membrane of cells that are viable or in the early stages of apoptosis because they maintain plasma membrane integrity. In contrast those cells in the late stages of apoptosis or already dead have lost plasma membrane integrity and are permeable to Propidium Iodide. Hepa 1-6 cells transfected with 1,5X molar excess of U1+5a, and with comparable amount of U1-wt and pUC18 (both as negative control), upon incubation with FITC labeled Annexin V antibody and Propidium Iodide, were evaluated through flow cytometry. Data obtained have been analyzed for event-type and for time point (Figure 63).
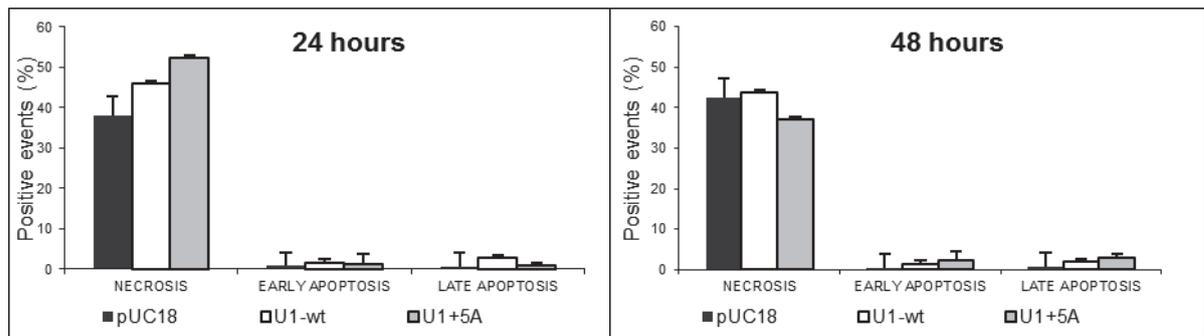
*Figure 63:* *Apoptosis and necrosis evaluation at 24 and 48 hours post transfection with pUC18 (black bars), pU1-wt (white bars) or pU1+5a (grey bars). Data are considered as percentile of gated events. Results of experiment are expressed as means ± SD, based on at least three independent experiments.*

As far as early or late apoptosis is concerned, none of the conditions were found to trigger major apoptosis, thus ruling out differential effects of the U1snRNA.

On the other hand, we observed induction of necrosis in all experimental conditions, which could be related to the cytotoxic effect of the reagent used for transfection.

As compared to cells transfected with the empty plasmid (38±4,8%), those transfected with U1snRNA appeared to undergo necrosis to slightly higher extent, but no remarkable difference were appreciable between cells transfected with the U1+5a (53±3%) or the U1-wt (46±3%) (Figure 63).

## IV.4.4 Analysis of effects on cell cycle

Differently to the above approach, the cell cycle analysis uses the Propidium Iodide DNA staining in combination with a permeabilizing agent to allow the staining of nuclear DNA. The fluorescence intensity of the stained cells at certain wavelengths will therefore correlate with the amount of DNA they contain. As the DNA content of cells duplicates during the S phase of the cell cycle, the relative amount of cells in the G0 phase and G1 phase (before S phase), in the S phase, and in the G2 phase and M phase (after S phase) can be determined, as the fluorescence of cells in the G2/M phase will be twice as high as that of cells in the G0/G1 phase. As shown in Figure 64 and 65, the analysis of cells transfected transiently with pU1+5a or pU1-wt or pUC18 did not reveal significant changes among groups.
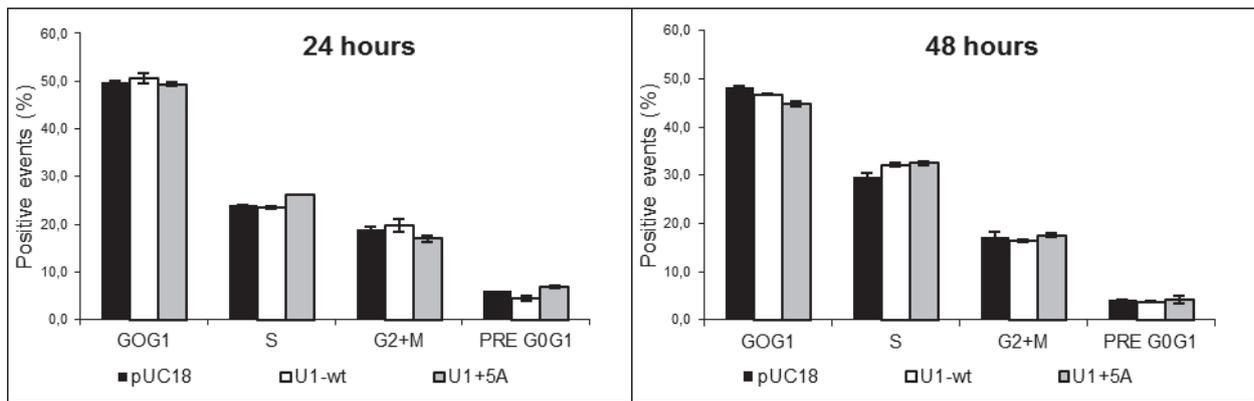
***Figure 64:*** *Cell cycle phases evaluation at 24 and 48 hours post transfection with pUC18 (black bars), pU1-wt (white bars) or pU1+5a (grey bars). Data are considered for type event and as percentile of gated events. Results of experiment are expressed as means ± SD, based on at least three independent experiments.*
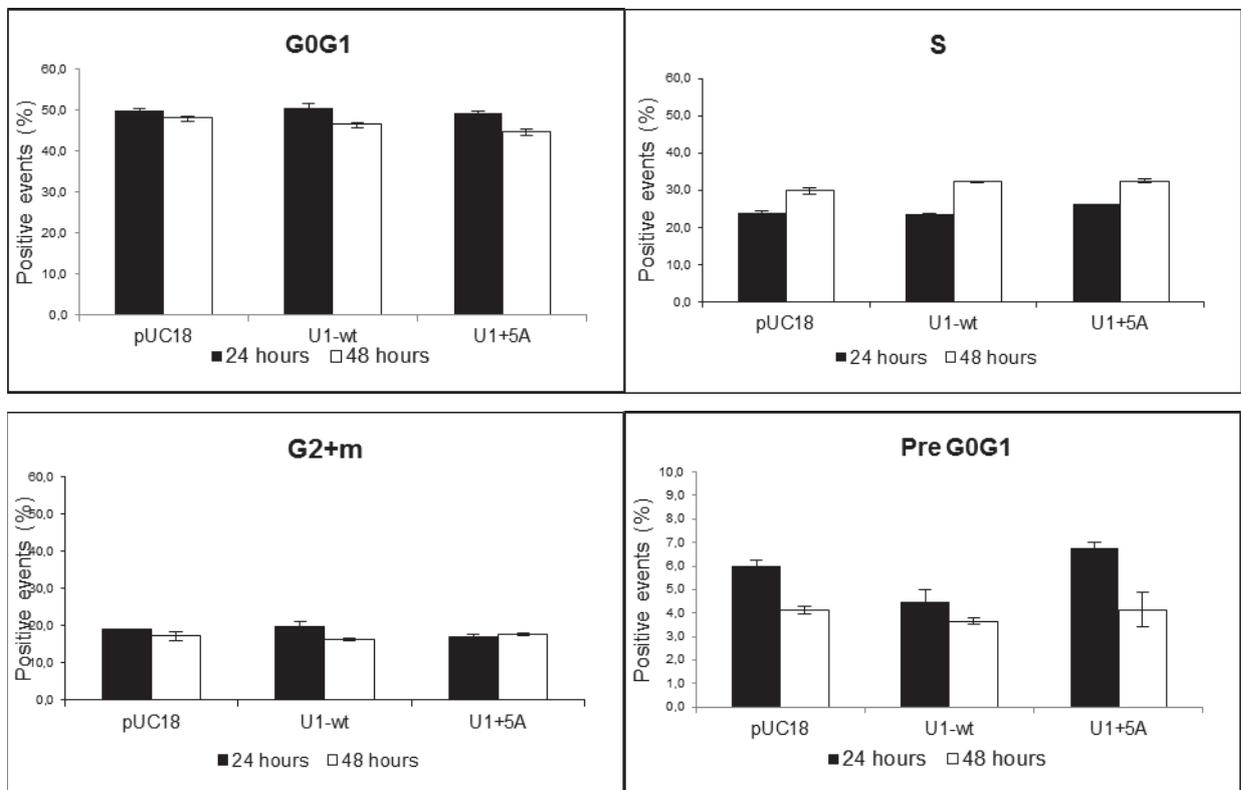


***Figure 65:*** *Cell cycle phases evaluation at 24 (black bars) and 48 (white bars) hours post transfection with pUC18, pU1-wt or pU1+5a. Data are analyzed by time point and as percentile of gated events. Results of experiment are expressed as means ± SD, based on at least three independent experiments.*

132

# IV.5 Rescue of FIX expression by ExSpeU1

## IV.5.1 Rescue of FIX expression by ExSpeU1snRNA

The study on the IVS7+5g/a (9726+5g/a) mutation in *F7* gene provided evidence both in cellular and animal models that a U1snRNA adapted to the mutated donor splice site is able to re-direct splicing to correct junction and to promote synthesis of correct mRNA and functional protein. The extent of the rescue obtained, if translated into patients, would be above the therapeutic threshold (2% or 10ng/ml), thus encouraging further studies aimed at proposing this approach for the therapy of coagulation factor defects caused by mutations producing defective donor splice sites.

However, even if the U1+5a at the proper dose will be demonstrated safe, one would imagine one different modified U1snRNA for each mutation, which make the drug development very laborious and expensive.

To investigate whether this strategy could be improved and allow to target different mutation with the same modified U1snRNA, we chose as model coagulation FIX. The rationale behind this choice was the observation that the deficiency FIX (hemophilia B) is much more frequent (1/40.000) than that of FVII (1/500.000) and thus that hemophilia B offers many models of splicing mutations to be investigated. Moreover, at variance from *F7* gene, *F9* gene is X-linked and indeed in male patients with splicing mutations the aberrant splicing is the sole cause of the deficiency.

In this study, from the many splicing mutations so far described in severe hemophilia patients (www.HGMD.org; http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html.) we selected as models those occurring either at the donor or acceptor splice sites of F9 exon 5, which are all candidate to affect the exon 5 definition by the spliceosome.

## IV.5.2 Characterization of a panel of *F9* splicing mutations surrounding exon 5.

In this study, we analyzed (Table 3):

-three synonymous point mutations (A>G, A>C and A>T) at position c.17896 in exon 5, which correspond to position -2 of the donor splice site junction;

-the c.17897G>T in exon 5, which correspond to position -1 of the donor splice site junction;

-the c.17897+1G>A, c.17897+2T>C, c.17897+4A>G and c.17897+13A>G in the donor splice site. 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 hemophilia 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).

Moreover, we focused the attention to other two mutations located in the poly-pyrimidine tract: the c.17660T>G and c.17661T>G mutations, located upstream to the acceptor site of FIX 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 moderate haemophilia B (Montejo, Magallon et al. 1999),

| Donor splice site mutations | | | | |
|---|---|---|---|---|
| *Name* | *Sequence* | *Aminoacid change* | *Phenotype* | |
| *wt* | CAGgtcata | | | |
| -2T | CTGgtcata | 127, Ala to Ala | Severe | |
| -2C | CCGgtcata | 127, Ala to Ala | Severe | |
| -2G | CGGgtcata | 127, Ala to Ala | Severe | |
| -1T | CATgtcata | 128, Val to Leu | Severe | Hemophilia B Mutation Database |
| +1A | CAGatcata | none | Severe | |
| +2C | CAGgccata | none | Severe | |
| +4G | CAGgtcgta | none | Moderate | |
| +13G | tgagtaaga | none | Mild-moderate | (Koeberl, Bottema et al. 1990; Ketterling, Drost et al. 1999) |
| Polypyrimidine tract mutations | | | | |
| *wt* | tgcttcttttagATG | | | |
| -8G | tgctgcttttagATG | none | Moderate | Hemophilia B Mutation Database |
| -9G | tgcgtcttttagATG | none | Moderate | (Montejo, Magallon et al. 1999) |

*Table 3: Hemophilia B database is available at*
[http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html](http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html).*Polypyrimidine tract and donor site mutations are named with their position respect to 5'ss. Uppercase letters represent exon sequences while lowercase letters indicate the intronic sequences. Natural mutations are red coloured.*

To understand the effect of these mutations on exon processing, we analyzed the strength of normal and mutated splice sites in terms of consensus values by the bioinformatics tool present to the site http://www.fruitfly.org/seq_tools/splice.html. Analysis with in silico program showed that the score of normal FIX exon 5 donor splice site was significantly reduced, and in presence of mutations, the strength of mutated sites were not detected.

In order to assess whether the mutations impair pre-mRNA processing, we exploited a splicing functional assay. For this purpose we used a modified version of the pTB hybrid minigene that has been reported to reproduce the *in vivo* splicing pattern of several gene systems (Pagani, Buratti et al. 2000; Pagani, Buratti et al. 2002; Baralle, Baralle et al. 2003; Pagani, Stuani et al. 2003). The minigene we designed was 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, Stuani et al. 2003). FIX exon 5 along with part of its intronic flanking region (IVS4 and IVS5) was cloned using a unique NdeI restriction site, which is located in the large fibronectin intron.

First, we compared the pFIX ex5 wt minigene splicing pattern in transiently transfected Hela cells with the splicing pattern in liver tissue. Exon 5 was amplified with primers base pairing to exon 4 and exon 6 and the splicing pattern was compared. Exon 5, in both the minigene and the liver, is not completely included, showing some degree of exon skipping in the mature mRNAs. These results indicated that FIX exon 5 is alternatively spliced, in accordance with the prediction of the in silico tool, and that the hybrid minigene (pFIX) successfully mimics the *in vivo* splicing pattern (Figure 66).
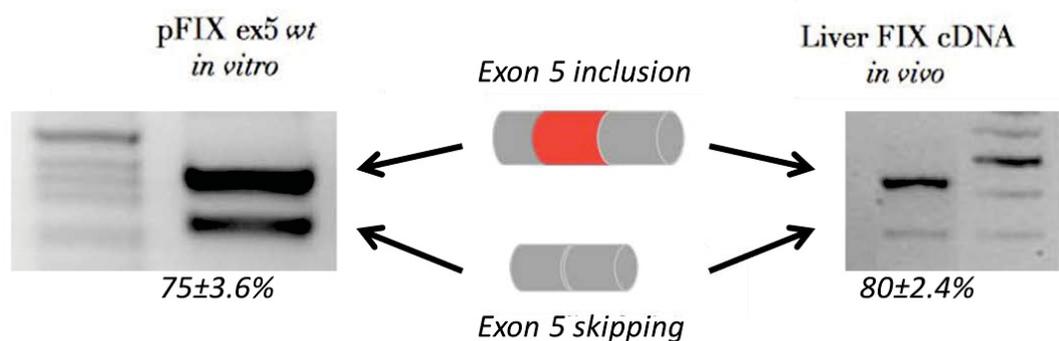


*Figure 66: Comparison between the levels of F9 exon 5 splicing in vivo and using a hybrid-minigene system: Hela cells were transfected with 0.5 µg of pFIX ex5 wt minigene and splicing pattern evaluated by RT- PCR with alpha2,3 and bra2 primers (left side of the picture). Amplified products were run 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. Human liver FIX cDNA in vivo amplified using FIXex4F and FIXex6R primers (right side of the picture).*

The splicing pattern observed in HeLa cells transfected with the mutant plasmids indicated that all mutations, with the exception of those at position +4 and +13, impaired the  definition of exon 5 and thus induced exon 5 skipping. F9 exon 5 PPT -8G and -9G and donor splice site mutation at positions -2, -1, +1, +2 triggered exon 5 skipping (Figure 67).

In particular the mutations at positions -1, +1 and +2 of the donor splice site virtually abolished correct splicing, as demonstrated by the undetectable levels of correct transcript even by radioactive labeling of RT-PCR product and capillary denaturing electrophoresis (not shown). At variance, the three mutations at position -2 mutants caused exon skipping but were compatible with very low amount of correct transcript (less than 3%), thus indicating that they produced a defective donor splice site.  Surprisingly, and in contrast to the prediction  of the in silico program, 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 (Figure 67). 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, we 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 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 (Figure 67).
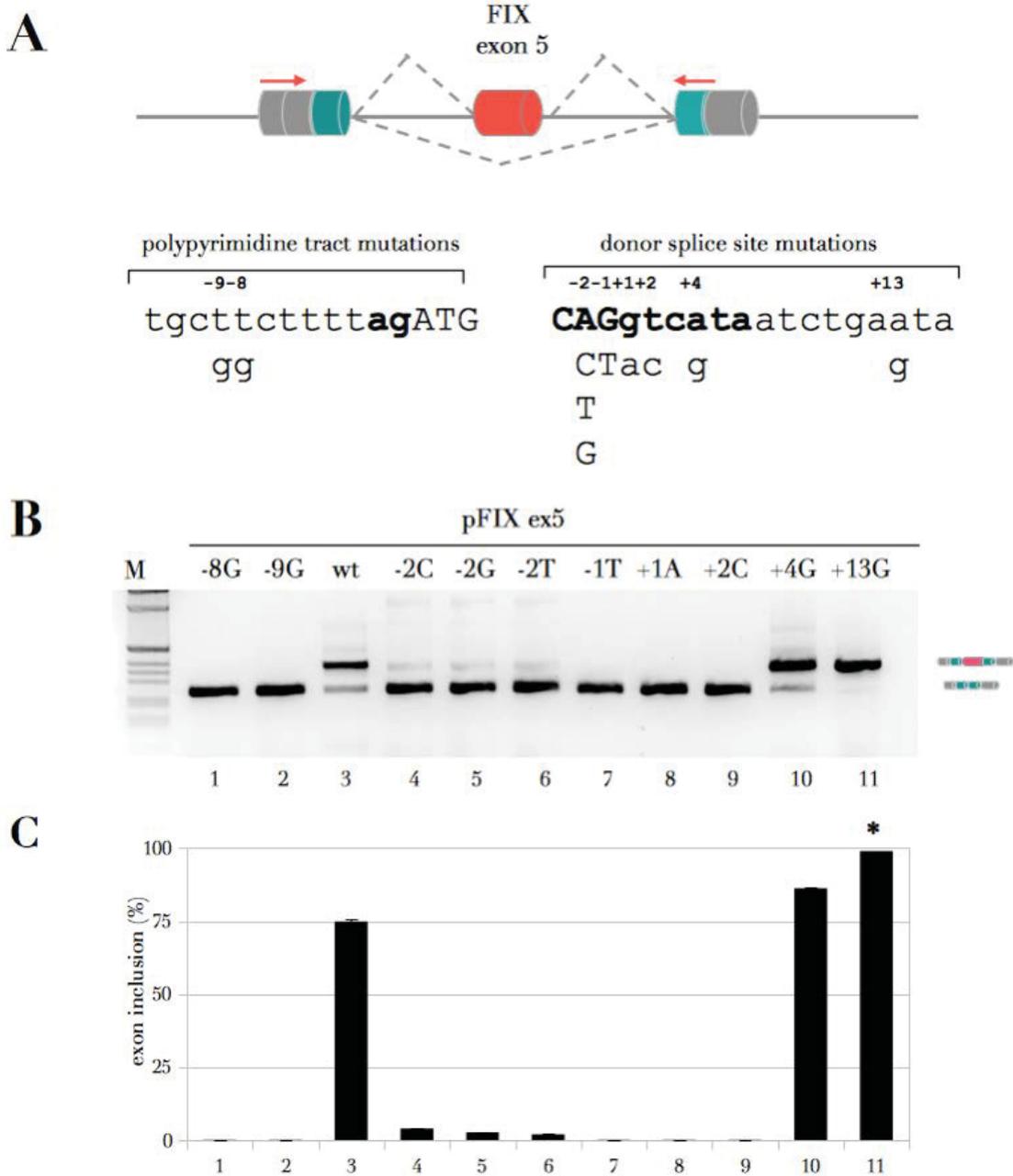
**Figure 67: Effect of polypyrimidine tract and donor splice site mutations on splicing of FIX exon 5.**

*A, Schematic representation of pFIX ex5 minigene. Dashed lines (- - -) illustrate possible splicing outcome. Arrows represent primers (alpha2,3 and bra2) 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 (indicated in the panel A). 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 based on gel densitometry analysis and is expressed as means ± SD, based on at least three independent experiments done in duplicate. *, cryptic site usage.*

137

## IV.5.3 Effect on splicing of modified U1snRNAs complementary to donor splice site mutations.

The canonical donor splice site of FIX exon 5 shows several mismatches with the 5' tail of U1 snRNA, attested even by the low score predicted by the bioinformatics tool. In fact FIX exon 5 donor splice site presents three mismatches: a C, a U and an A at positions +3, +5 and +6, respectively. To understand the correlation about the interaction between the U1snRNA and FIX exon 5 donor splice site, we introduced compensatory changes within the 5'-tail of U1 snRNA. So we created plasmids for engineered U1snRNAs with perfect complementarity to the wt or each mutated 5' ss sequences (Figure 68).
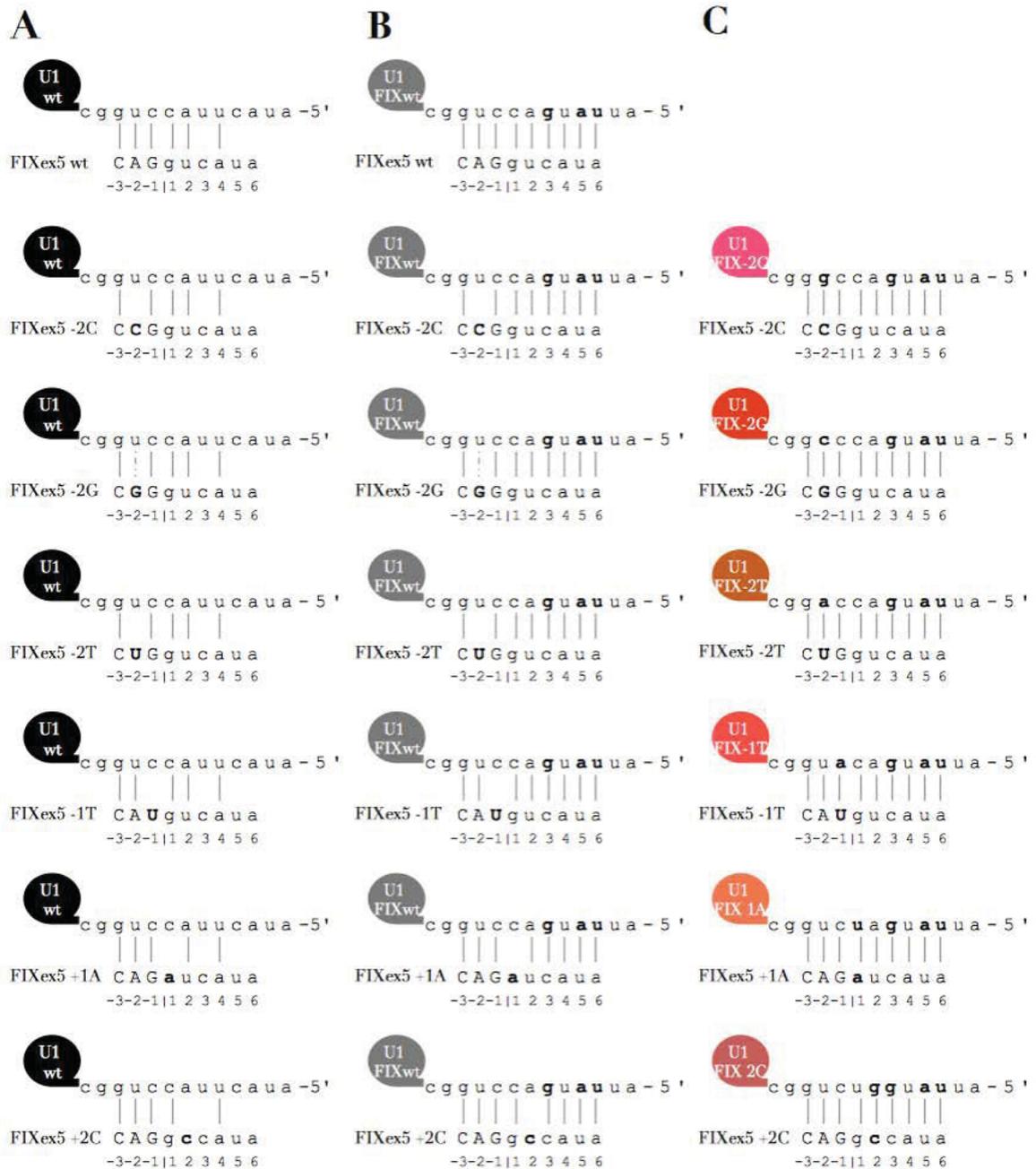
**Figure 68: Complementary between modified U1 snRNAs and FIX exon 5 donor site mutants.**
*A, Base pairing of FIX exon 5 wt and mutant donor sites with normal U1 snRNA (U1 wt).*
*B, Base pairing of FIX exon 5 wt and mutant donor sites with U1FIXwt. U1FIXwt was created by introducing three substitutions (a cytosine, an adenine and a thymidine; bold, U1 5'-tail) within normal U1 snRNA to increase base pairing to FIX exon 5 wt donor site. These substitutions allow the formation of Watson-Crick bonds.*
*C, Complementarity of FIX 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 FIX 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.*

139

Upon the co-transfection of the modified U1snRNA fully complementary to the FIX exon 5 wt donor splice site with the pFIX exon 5 wt minigene, or bearing the mutations under investigation ( -8, -9 in the poly-pyrimidine tract and -2, -1, +1, +2, +4 and +13 in the donor splice site), the splicing pattern of exon 5 was evaluated using primers alpha2,3 and bra2 to avoid the amplification of aspecific products. Co-transfection of U1FIXwt with FIX exon 5 wt minigene improved exon inclusion, from 75% to nearly complete inclusion (Figure 69). Over-expression of the U1-wt did not change FIX exon 5 wt splicing pattern. Co-transfection of U1FIXwt with -2 mutant minigenes also led to complete exon inclusion, regardless of the base change under analysis. In contrast, U1FIXwt did not revert exon skipping caused by mutations -1T, +1A and +2C. When we explored the effect of U1FIXwt on exon 5 skipping caused by the poly-pyrimidine tract mutations -8G and -9G, U1FIXwt improved exon inclusion greatly, reaching about ~75% for the poly-pyrimidine tract mutant minigenes (Figure 69).
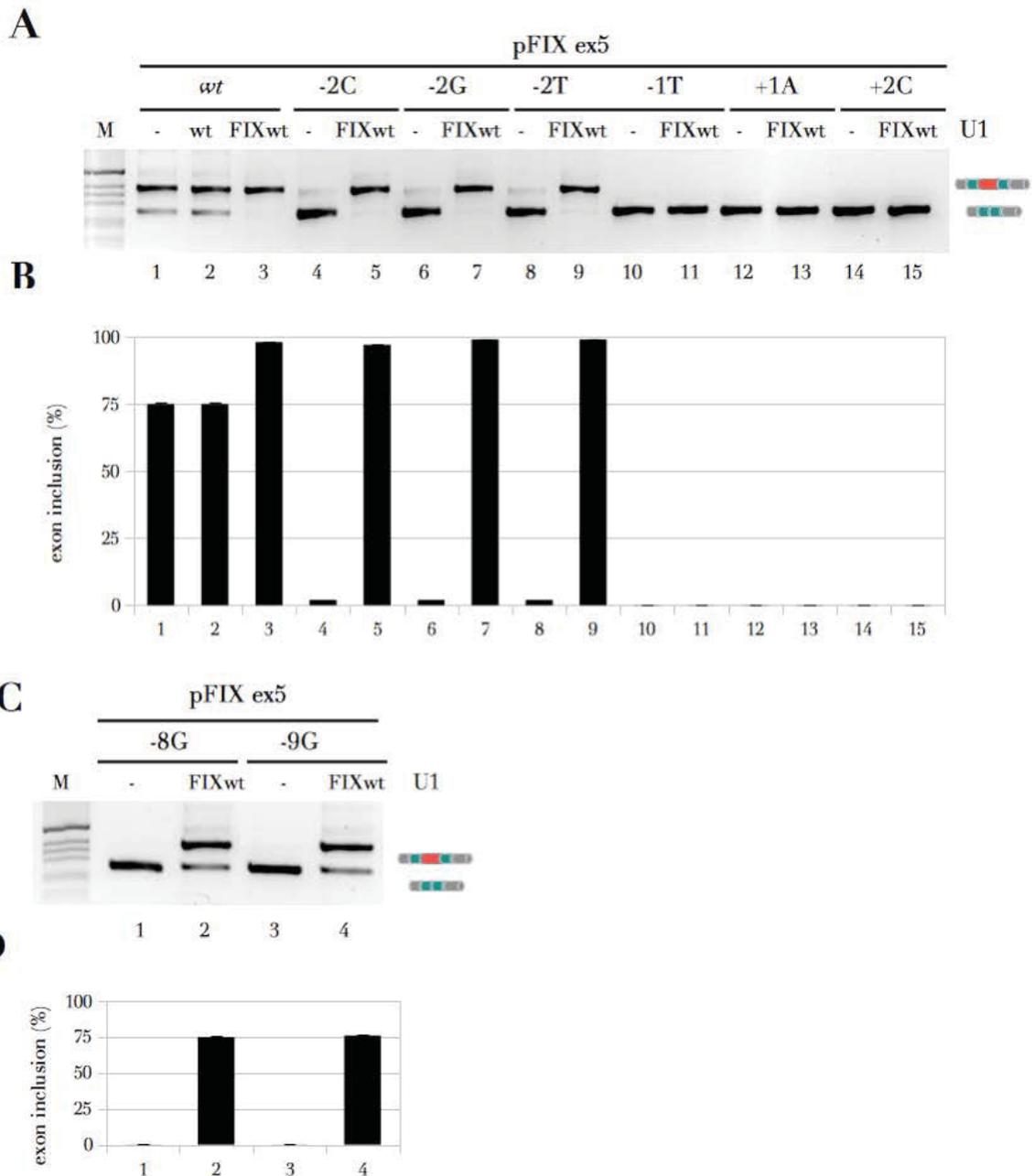
***Figure 69: FIX exon 5 mutant minigenes co‑transfected with U1FIXwt***

*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 FIX 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).*

*B, Quantification of splicing pattern analysis in A. Exon 5 inclusion percentage was estimated using gel densitometry 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 FIX exon 5 poly-pyrimidine tract mutant minigenes transfected either alone or with U1FIXwt, as in B.*

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

After that we evaluated the effect of FIX mutation specific-U1snRNAs by co-transfecting each mutated FIX minigene with its corresponding modified U1snRNA (Figure 70). U1FIX -2C, -2G and -2T restored exon 5 inclusion of their corresponding mutant minigenes, whereas U1FIX -1T, +1A and +2C failed to improve the splicing pattern. 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 (Figure 70).

Taken together, these results revealed that loading of U1 snRNAs complementary to the mutated donor site can recover several exon skipping-causing mutations. In the case of donor site mutants, the rescue efficiency depends on the type of mutation. In fact, exon skipping caused by mutations -2C, -2G and -2T within FIX exon 5 were corrected (although with variable efficacy) while FIX exon 5 -1T, +1A and +2C failed to be rescued.
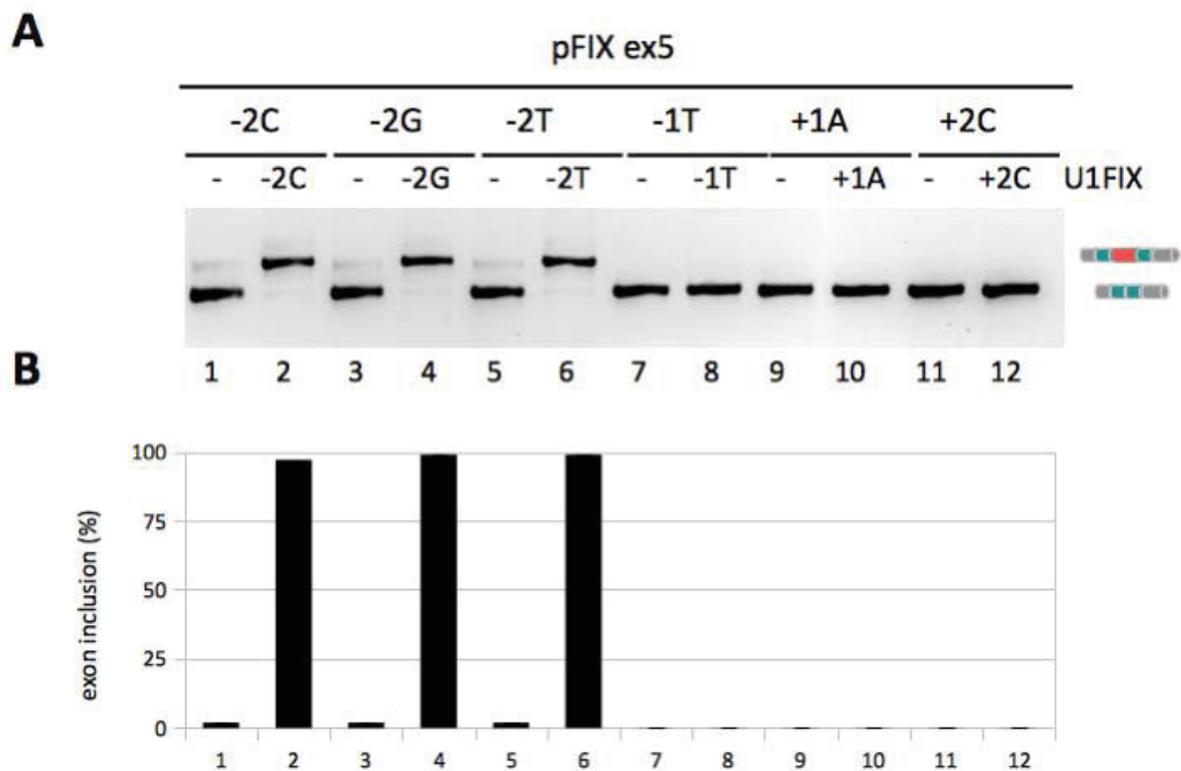


**Figure 70: FIX exon 5 donor site mutant minigenes co-transfected with compensatory U1FIX snRNAs.**
*A, Splicing pattern analysis of the RT-PCR products derived from RNA of transfected HeLa cells, separated on 2% agarose gel. Donor splice site mutant FIX exon 5 minigenes were transfected alone (0.5 μg) or with their corresponding compensatory U1FIX (0.5 μg) and the splicing pattern evaluated with alpha2,3 and bra2 primer.*
*B, Quantification of splicing pattern analysis in A based on densitrometric analysis of agarose gel bands. Exon 5 inclusion is expressed as means ± SD, based on at least three independent experiments done in duplicate.*

## IV.5.4 Effect of Exon-Specific U1 snRNAs on FIX exon 5 mutations

The approach so far exploited to rescue FVII and FIX expression impaired by splicing mutations was based on the modification of the 5' tail of the U1nsRNA enabling it to target by complementarity the donor splice site from position -3 to position +6 bp. So, the recognition is based on the base pairing of only 9 bp in total. Only the GT dinucleotide is almost universally conserved. The remaining nucleotides positions display variable conservation, with some positions being more conserved than others. Therefore, the base pairing between the donor splice site and the U1snRNA 5' tail is not perfect, with differences depending on the sequence of the donor splice site. The relatively degenerated sequence of the donor splice sites and the observation that even those significantly diverging  from the consensus sequence are functional open the possibility that the engineered U1snRNA might bind to sequences other than the target one and alter the splicing pattern of other genes. It must be noticed that a sequence with the feature of donor splice site do not necessarily trigger the assembly of the spliceosome machinery, being necessary other neighboring splicing regulatory elements (i.e. ESE,ISE,ESS,ISS) recruiting auxillary splcing factors.  This would make the off-target binding of modified U1snRNA to uneffective in most cases. However, we cannot rule out the activation of cryptic sites having features and sequence context prone to recruit the spliceosome.

One strategy to overcome this potential risk could be to increase the specificity of the modified U1snRNA for the target gene region. Since the intron sequences are less conserved than the boundaries between exons and introns, one way to achieve the required higher specificity of the engineered U1 is to design the 5' tail of the U1snRNA to base-pair with an intron sequence in the proximity of the mutated donor splice site. To test this hypothesis, we proceeded though the design and creation of the so called "shifted U1snRNAs", base-pairing in the surrounding of the mutated donor splice site.

This approach is however hardly applicable to the donor splice site of F7 exon 7, where the 9726+5g/a occurs, since it is characterized by the presence of a repeated minisatellite generating multiple cryptic donor splice sites with sequence equal to the donor splice site. For this reason, we used as model the F9 exon 5.

The aim was to evaluate the ability of a shifted U1, landing in an intron sequence located downstream the mutated donor splice site of exon 5, to rescue the panel of splcing mutations investigated above. We then created a series of U1snRNA with the 5'ss complementary to intronic sequences downstream of the donor splice site (Table 4) (Figure 71).

| U1 name | Intron sequence targeted | Sequence lenght |
|---------|--------------------------|-----------------|
| 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 4: Table showing the name, intron sequence bound and its length*
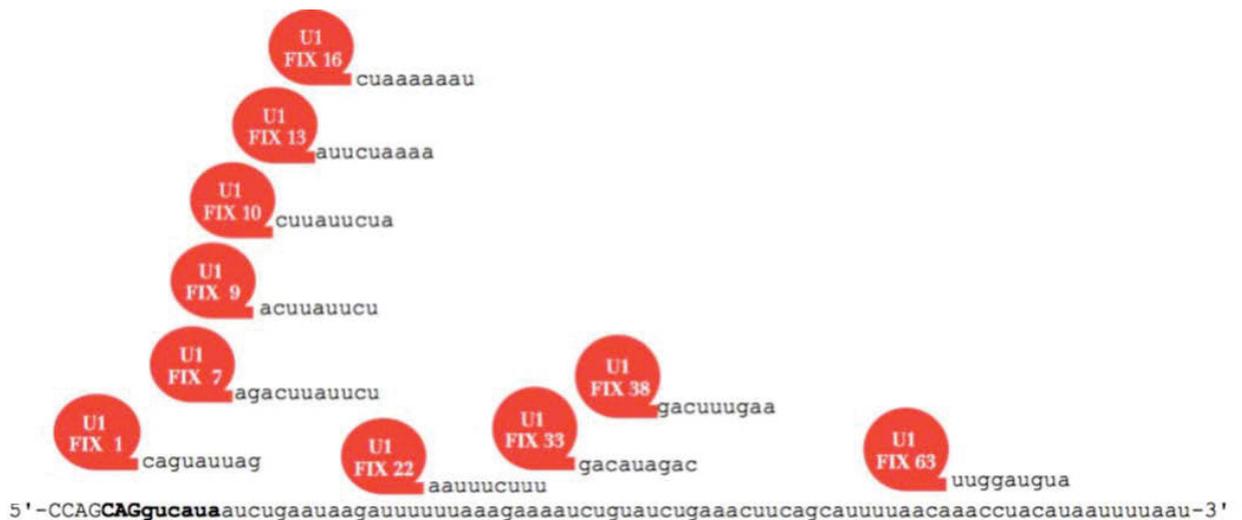


***Figure 71: Binding regions of exon-specific U1FIX snRNAs***
*Schematic representation of the positions where U1FIXs bind. Exonic sequences is represented in uppercase letters and part of the IVS5 flanking region in lowercase letters. FIX exon 5 donor splice site is bolded.*

New U1sRNAs were initially tested only on the pFIX ex5 -2C minigene. The plasmid has been transfected alone or in combination with different U1snRNAs and the splicing pattern was evaluated for exon 5 inclusion (Figure 72). Analysis of the splicing pattern revealed that the majority of U1FIXs were able, although with different efficiency, to correct exon 5 skipping. In particular, U1FIX 1, 9, and 10 rescued very efficiently the aberrant splicing of the -2C mutation, leading to a percentage of inclusion of about 95-100% (Figure 72). Generally, the percentage of

exon 5 inclusion progressively decreased as binding of the U1FIXs moved downstream (U1FIX 13, 16, 22, 33, 38 and 63). Base pairing of U1FIXs between position +1 and +18 had the strongest stimulatory effect on -2C mutant exon 5 inclusion whereas downstream binding progressively decreased the efficacy (Figure 72).
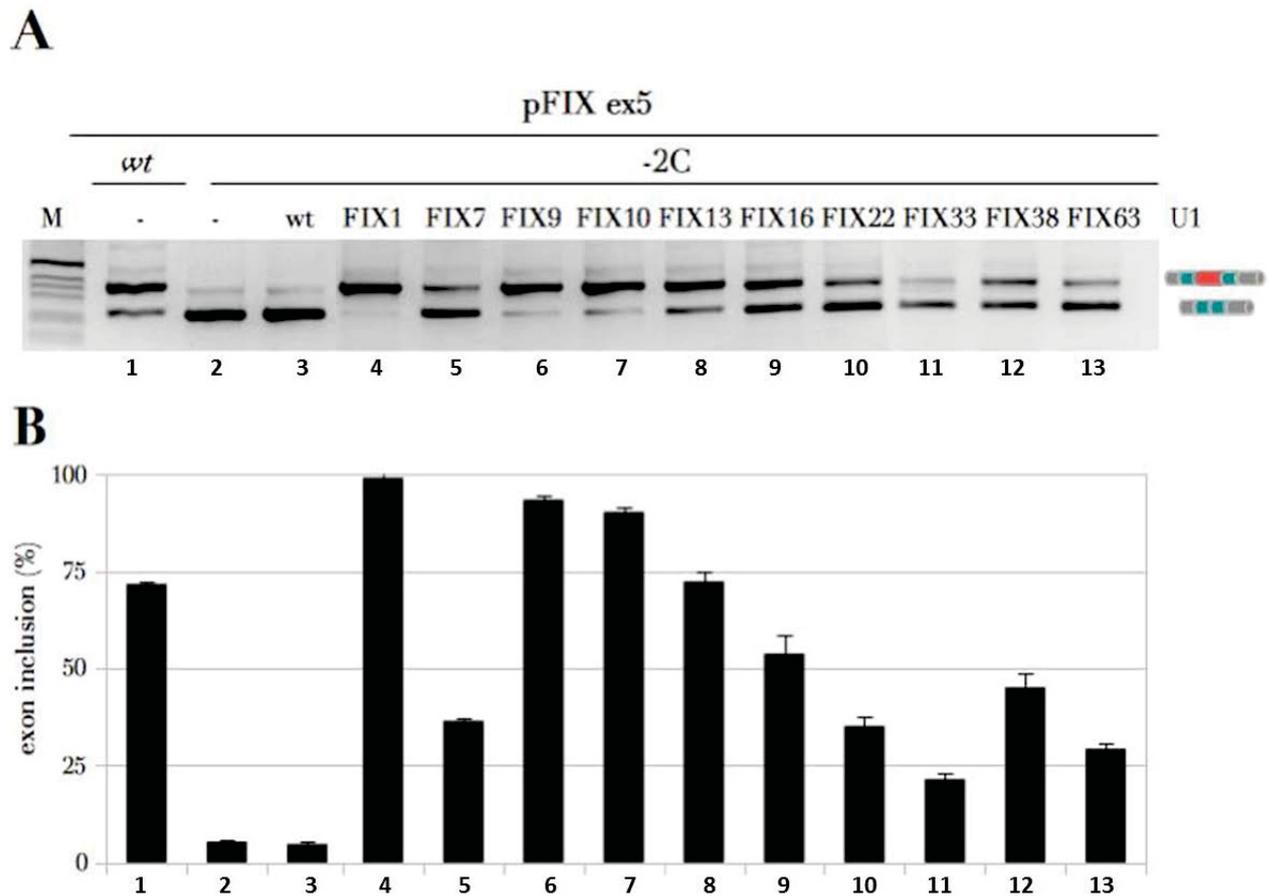


***Figure 72: Effects of exon‑specific U1FIX snRNAs on splicing pattern of pFIX ex5 -2A>C minigene***
*A, Splicing pattern of FIX exon 5 -2C minigene co‑transfected with plasmids encoding for U1FIX snRNAs (U1FIX 1, 7, 9, 10, 13, 16, 22, 33, 38, 63) is visualized on a 2% agarose gel after RT‑PCR.*
*U1FIX 1, 9 and 10 exhibited the highest stimulatory activity, which is clearly evidenced from the histogram in B.*
*B, Quantification of exon 5 inclusion by densitrometic analysis of RT-PCR in A. Exon inclusion is expressed as percentage and as means ± SD, based on at least three independent experiments done in duplicate.*

Of the three most active U1FIXs (1, 9 and 10), we chose to address our attention on the U1FIX 9 to further evaluate the efficacy on the other FIX mutations, the two additional synonymous mutations at position -2, the mutations -1T, +1A and +2C and the poly-pyrimidine tract (PPT) mutations -8G and -9G. Upon the co-transfection of U1FIX 9, the exon 5 skipping caused by -8G and -9G PPT mutations and by -2 (-2G and -2T) mutations were completely rescued (Figure 73). As expected, the splicing defects present in the -1T, +1A and +2C minigenes were not corrected by co-transfection of U1FIX 9. Interestingly, the FIX exon 5 mutants rescued by U1FIX 9 were also similarly rescued by U1FIXwt (Figure 73).
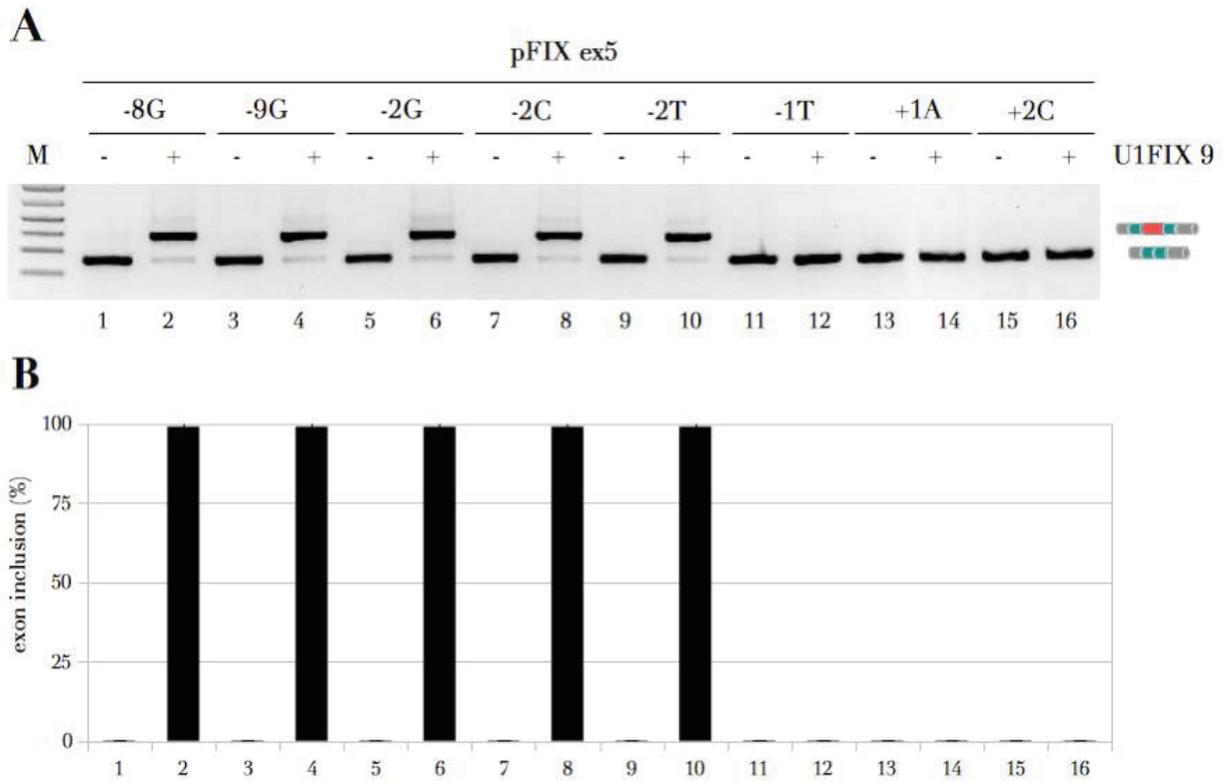
***Figure 73: Effect of ExSpeU1FIX 9 on splicing pattern of polypyrimidine tract***

***and donor site mutations in FIX 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 U1FIX 9 is visualized on a 2% agarose gel after RT-PCR.*
***B,*** *Quantification of exon 5 inclusion by densitrometic analysis of RT-PCR in A. Exon inclusion is expressed as percentage*
*and as means ± SD, based on at least three independent experiments done in duplicate.*

## IV.5.5 Effect of the ExSpeU1FIX 9 on FIX biosynthesis and function.

As for the study on the 9726+5g/a mutation in F7 gene, we investigated THE ABILITY OF THE ExSpeU1 to rescue the synthesis of functional FIX in the presence of a panel of splicing mutations either in the donor or in the acceptor splice site. This is the key of a successfully therapeutical strategy.

To evaluate whether U1FIXs can also rescue protein synthesis (antigen and activity level), a full-length splicing-competent FIX hybrid cassette (pSC-FIX exon 5) has to be designed. This expression plasmid includes the FIX exon 5 and its flanking intronic sequences inserted in the factor IX cDNA. In this system, a unique NdeI restriction site is located in the surrounding intronic sequences to allow fast and efficient subcloning of exon 5 wt or mutant cassettes into the new plasmid. The correct splicing processing of the upstream and downstream introns flanking exon 5 give rise to full-length FIX mRNA coding for normal coagulation factor IX (Figure 74).
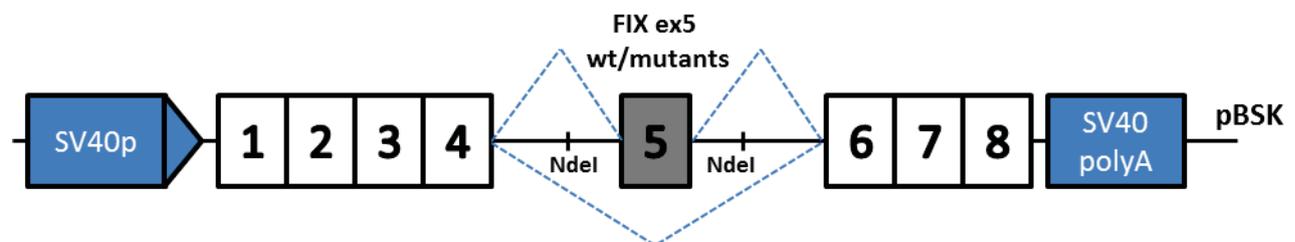


***Figure 74:*** *Scheme of the pSC-FIX exon 5 plasmid created to evaluate the mediated-U1 rescue of FIX exon 5 mutations at protein level*

We analyzed the effect of the FIX fully complementary U1FIX wt or of the shifted U1FIX 9 on five different mutations: two at the poly-pyrimidine tract (PPT) -8G and -9G and the three variants located in position -2 in the donor splice site -2G, -2C, -2T. So we created five mutant FIX exon 5 splicing competent expression plasmids: 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 5PPT -8G and exon 5PPT -9G (carrying PPT mutations -8G and -9G, respectively). These five constructs were co-transfected with each modified U1 snRNAs (U1FIX wt or U1FIX 9) in BHK cells, and splicing pattern analysis was carried out by RT-PCR using primers base-pairing to exon 4 and 6 (FIXex4F and FIXex6R respectively). Media was also collected to further analyze the rescue at protein level (antigen and activity). BHK cells were chosen because known to properly synthesize the functional FIX protein. We have had to switch from CMV promoter, used in all previously experiments in FVII or FIX experiments, to SV40 promoter due to the presence in the first one of various NdeI restriction sites, avoiding the possibility to sub-clone easily our exon 5 cassette (wt or bearing mutations) into the new plasmid. Amplification of pSC-FIX exon 5 wt transcripts using FIXex4F and FIXex6R primers showed the presence of two bands, corresponding to exon 5 inclusion (604 bp) and exon 5 skipping

(475 bp). The quantification of the bands intensity by densitometric analysis revealed that exon inclusion accounted for ~15% of exon in pSC-FIX exon 5 wt (Figure 75). This percentage is significantly lower than the previously observed in normal liver (~80%) or by the pFIX ex5 wt minigene (~75%). 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. When the five mutants were analyzed (pSC-FIX exon 5 PPT -8G, exon 5 PPT -9G, exon 5 -2C, exon 5 -2G and exon 5 -2T), the complete exon 5 skipping occurred in each mutant as expected. The co-transfection of either U1FIXwt or U1FIX 9 significantly with all five mutants ameliorated splicing, with the exon inclusion form that accounted for ~70-80% of the total transcripts (Figure 75).
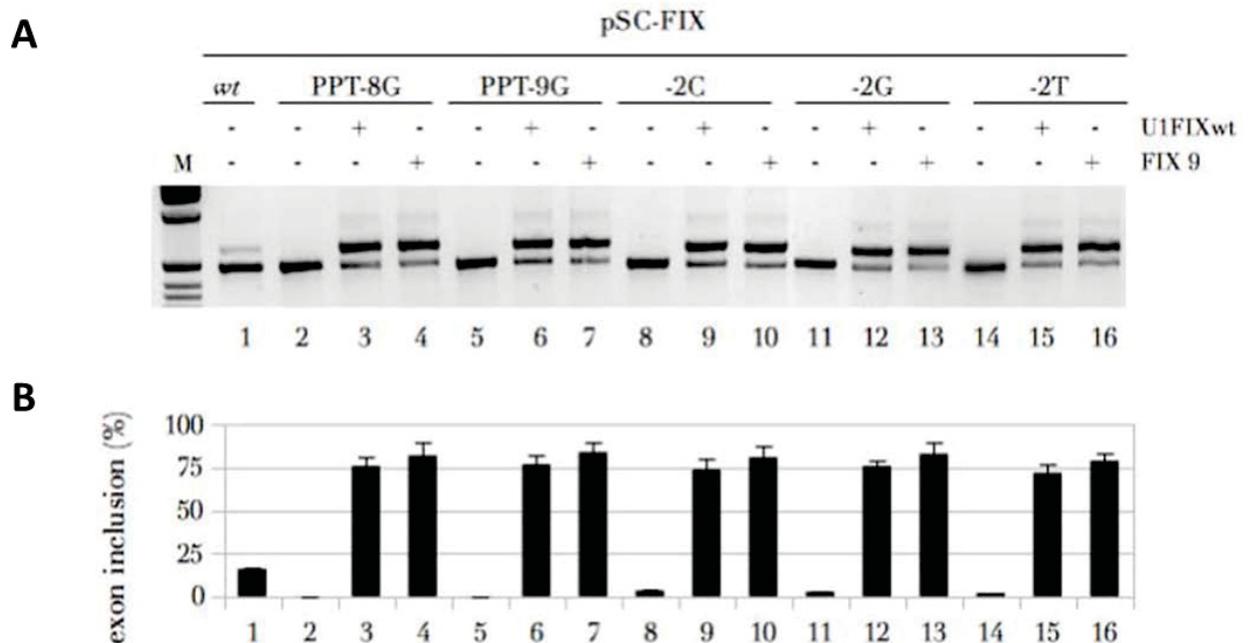


*Figure 75: A, Analysis of spliced transcripts. BHK cells were transfected with 0.5 µg of pSC-FIX exon 5wt, pSC-FIX exon 5PPT -8G, pSC-FIX exon 5PPT -9G, pSC-FIX exon 5 -2G, pSC-FIX exon 5 -2C or pSC-FIX exon 5 -2T either alone or with U1FIXwt or U1FIX 9. Splicing pattern was evaluated by RT-PCR with FIXex4R and FIXex6F primers and amplified products were resolved on a 2% agarose gel. B, Quantification of exon 5 inclusion by densitrometic analysis of RT-PCR in A. Exon inclusion is expressed as percentage and as means ± SD, based on at least three independent experiments done in duplicate.*

Since the Factor IX protein is normally secreted in the blood stream, it can be recovered from the culture supernatants and quantified or analyzed for its activity. 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, pooled normal plasma (PNP) from healthy patients was used, which contains normal levels of factor IX protein (5ug/ml). Western blot analysis was performed on media previously collected exploiting a polyclonal anti FIX antibody. pSC-FIX exon5 wt generated both factor IX protein isoforms: normal factor IX protein (lane 2, upper band) and truncated factor IX protein (lane 2, lower band). The normal one was more abundant than the truncated one, in contrast to the observed FIX splicing pattern (with exon skipping form more abundant than the exon inclusion one) (Figure 76). A possible explanation of the discrepancy observed between the splicing pattern and the protein secreted could reside in a cytoplasmatic degradation or inefficient secretion of the truncated protein, due to the aberrant folding of the shorter protein. On the other hand, transfection of pSC-FIX exon 5PPT -8G, exon 5PPT -9G, exon 5 -2C, exon 5 -2G and exon 5 -2T minigenes led to the production of only the ~51 kDa form (Figure 76). Strikingly, for all pSC-FIX exon 5 mutants the presence of either U1FIXwt or U1FIX 9 completely restored the full-length factor IX protein.
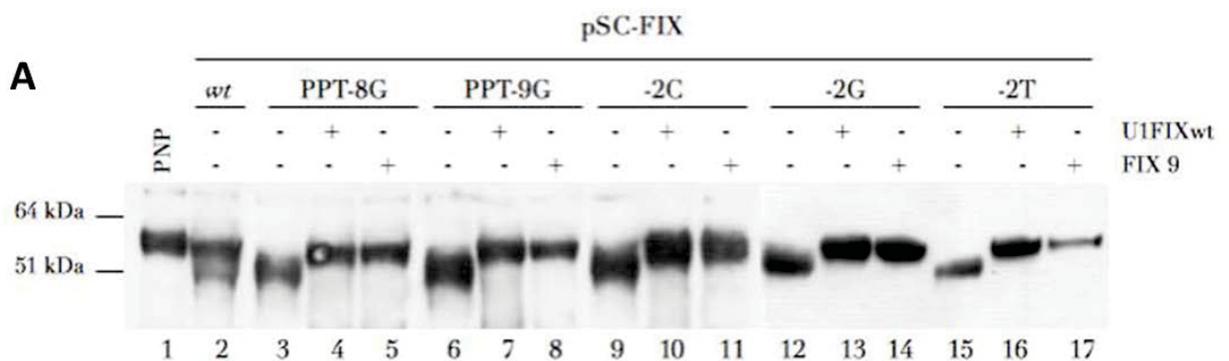


**Figure 76:** *Western Blotting against factor IX protein on supernatant of BHK cells transfected with pSC-FIX exon 5 wt ,pSC-FIX exon 5PPT -8G, pSC-FIX exon 5PPT -9G, pSC-FIX exon 5 -2G, pSC-FIX exon 5 -2C or pSC-FIX exon 5 -2T either alone (lanes 2, 3, 6 and 9) or with U1FIXwt (lanes 4, 7, 10) or U1FIX 9 (lanes 5, 8, 11). As a control, pooled normal pooled (PNP, lane 1) from healthy patients was used for WB.*

To quantify the U1-mediated rescue of FIX protein, we performed a sandwich ELISA assay with a polyclonal anti FIX capture antibody. The assay revealed that even in media from cells transfected with only the mutated plasmids an appreciable amount of FIX antigen was detected (Figure 77, panel A). This suggests that the primary antibody used cannot discriminate between the full-length active isoform and the shorter inactive form, capturing both of them. Anyway in media from cells transfected with pSC-FIX exon 5 mutant plasmids and U1FIXwt or U1FIX 9 a significant ~2-3 fold increase of factor IX antigen was detected (Figure 77, panel A).

Finally, the coagulant activity of secreted FIX was evaluated through an aPTT assay. As expected, no factor IX activity was detected when pSC-FIX exon 5PPT -8G, exon 5PPT -9G, exon 5 -2C, exon 5 -2G and exon 5 -2T were transfected alone, since it is expected that the truncated factor IX

protein does not possess any activity, due to the lack of a domain crucial for its function. Intriguingly, upon co-expression of the modified U1snRNAs, either U1FIXwt or U1FIX 9, the FIX coagulant activity in cells expressing the splicing FIX variants resulted to be comparable to that of the wt contruct, thus indicating the effective rescue. (Figure 77, panel B). Taken together these results demonstrated that the exon-specific U1FIXwt and the shifted U1FIX 9 are able to correct a panel of different mutations at the PPT and donor splice site mutations at positions -2 in a splicing competent expression vector system, re-establishing protein synthesis, secretion and activity (Figure 77, panel A and B).
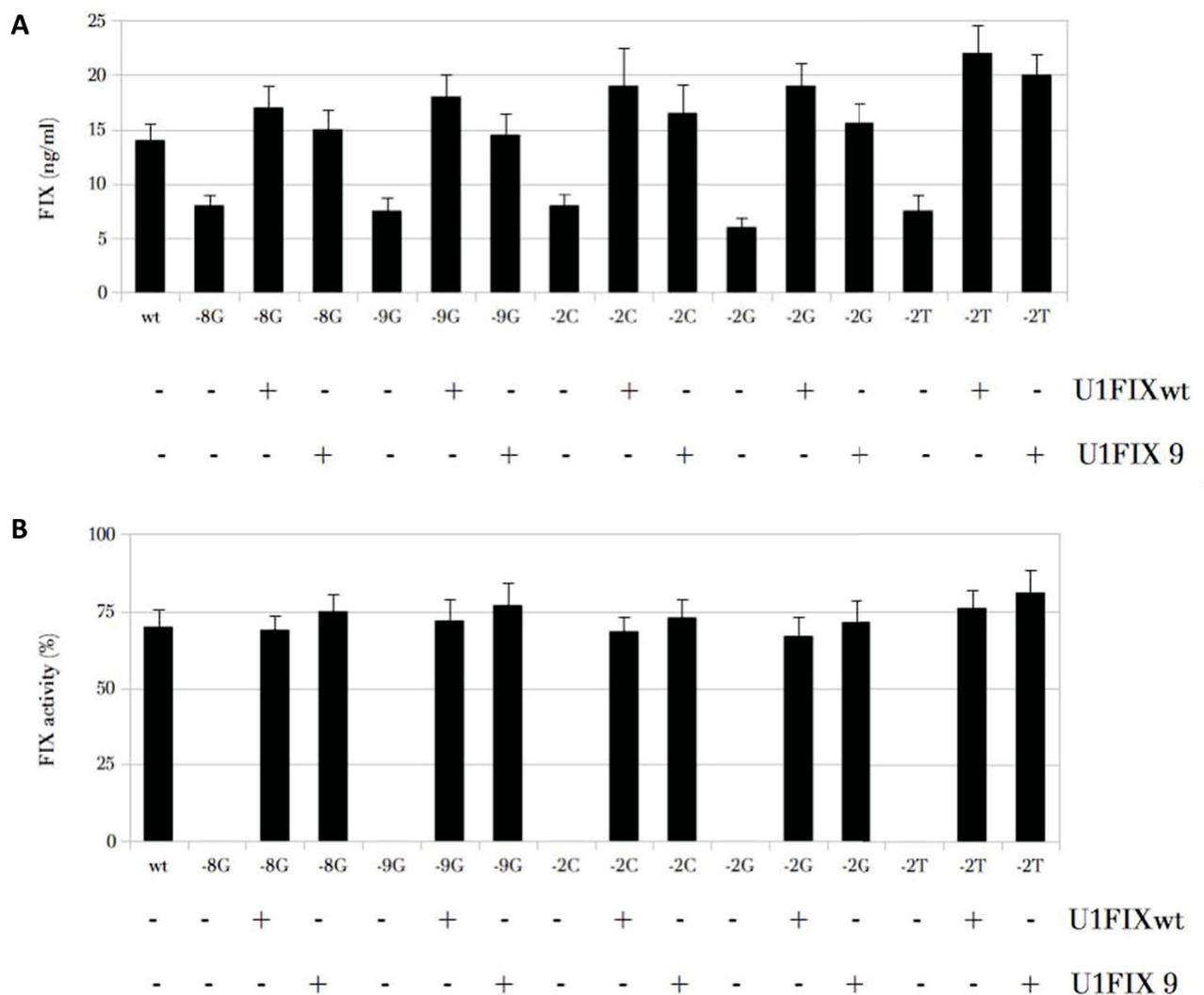


**Figure 77: Rescue by U1FIX wt and U1FIX9 snRNAs of FIX antigen and activity levels in BHL cells supernatant**
*A,ELISA assay showing concentration of secreted factor IX protein in supernatant of BHK cells transfected with pSC-FIX exon 5 wt, pSC-FIX exon 5PPT -8G, pSC-FIX exon 5PPT -9G, pSC-FIX exon 5 -2G, pSC-FIX exon 5 -2C or pSC-FIX exon 5 -2T either alone or with U1FIXwt or U1FIX 9.*
*B, Factor IX specific activity based on aPTT assay of secreted factor IX protein in media of BHK cells transfected with pSC-FIX exon 5wt, pSC-FIX exon 5PPT -8G, pSC-FIX exon 5PPT -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.*

# V.   Discussion and Conclusions

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

Current treatment for coagulation factor deficiencies is based on the intravenous administration of the missing proteins (replacement therapy), either plasma derived or produced by recombinant DNA technology (Pipe, High et al. 2008), in response to bleeding episodes or prophylactically (Carcao and Aledort 2004).

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

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

Enormous efforts have been pushed on the substitutive gene therapy that consists in the viral or non-viral mediated delivery of a copy of the defective gene (or better of the coding DNA sequence) into the patient's cells, thus triggering stable endogenous expression of the missing protein (Murphy and High 2008; Petrus, Chuah et al. 2010).

Among all mutations found in patients, and considering that the frequency of mutations type varies considerably between individual genes, it is estimated that more than 30% of pathogenic mutations cause disease through aberrant splicing mechanisms. It is now clear that substitutions that had for long time been regarded as silent synonymous changes in protein coding regions may have some very severe consequences on splicing process, and thus be the cause of disease (Pagani and Baralle 2004; Buratti, Baralle et al. 2006). With this as background, the quest for new tools able to modulate the splicing mechanism is growing up.

An emerging area of research is represented by the correction of the gene expression of the mutated gene by modulation of the messenger RNA (mRNA) processing through molecules acting at RNA level. They are represented by siRNA, shRNA, miRNA, antisense oligonucleotide and recently by antisense U7/U1snRNA. They act at RNA level modulating prevalently the degradation of the target RNA through various mechanisms or inducing the formation of an aberrant splicing event on the target molecule that results in a functional protein isoform (Hammond and Wood 2011).

Most recently, the usage of a modified small nuclear ribonucleoprotein U1 (U1snRNA) as modulator of spliceosome assembly has emerged as a promising approach for the treatment of splicing-related diseases. It has been successfully explored for the treatment of human genetic

disorders other than coagulation factor diseases (Wood, Yin et al. 2007; Bonetta 2009; Wilton and Fletcher 2011). Notably, this approach would permit restoration of gene expression while maintaining the gene promoter regulation in the cells belonging to the physiological tissue. Moreover, it has the potential to circumvent some limitation due to the large size of certain human disease genes, and could be also effective to address dominant-negative disease forms.

To evaluate the efficacy of a modified U1snRNA in the treatment of coagulation factor disorders caused by aberrant splicing, we chose the *F7* IVS7+5g/a mutation (F7 9726+5g/a) that was found to be associated to undetectable plasma FVII levels and a severe bleeding symptomatology in several Italian FVII deficient patients. The mutation occurs at position +5 of the donor splice site (5'ss) of intron 7 of *F7* gene, which is located in the first of six highly homologous 37bp repeats. This region therefore possesses a corresponding number of strong cryptic 5'ss, thus complicating the selection of the correct 5'ss by the spliceosome machinery. In normal conditions, only the most upstream 5' donor splice site is used, whereas the pseudo-sites remain silent, suggesting a oriented scanning mechanism (Borensztajn, Sobrier et al. 2006).

To explore the effects of the IVS7+5g/a mutation, the analysis of liver-extracted RNA would be the ideal approach. Since this way is not feasible, we moved toward the utilization of a minigene approach, which help to overcome the cloning of the entire gene into a vector. We prepared an extended minigene spanning genomic region exon 6 through exon 8 and analyzed its expression in a human liver cell line (Hep3B) to mimic proper cell environment. We exploited also COS-1 cells to replicate major findings. In both cell line, the mutation *F7* IVS7+5g/a induced exon 7 skipping and usage of the first downstream cryptic 5' donor splice site, both predicting for FVII mRNA frameshift and premature translation termination. Exon 7 skipping accounted for the majority of transcripts (80%), suggesting the loss of exon recognition caused by the mutation. Since no traces of normal transcripts have been revealed in agarose gel analysis of RNA products, and since total absence of FVII is considered incompatible with life, we exploited a more sensible approach, the denaturing capillary electrophoresis of fluorescently labeled RT-PCR products. This methodologic approach revealed traces of normal transcripts (~0.2%) in the mutant, that explain residual FVII levels in homozygous patients and the association of the mutation with a severe but not lethal FVII deficiency.

The mechanism through which the F7 IVS7+5g/a change might account for the aberrant splcing pattern is the reduced complementarity between the 5'-tail of the U1-snRNA and the mutated IVS7 5' donor splice site. During the earliest splicing steps, this would lead to inefficient recognition of the mutated 5'ss and skipping of the exon 7 or alternatively usage of the cryptic 5'ss in the second intronic repeat.

To verify the correction efficacy of engineered U1snRNAs with improved complementarity to the IVS7 5'ss we carried out complementation assays. In particular we designed for U1 snRNAs variants designed to bind the mutated donor splice site (U1+5a) or multiple sites in the IVS7 repeats (U1Mut1 and U1Mut2). Evaluation of transcripts in co-transfected cells revealed that all mutated U1-snRNAs reduced exon skipping (> 50% reduction), suggesting that their binding at or near the 5' ss favored exon 7 definition. Among them, only the U1+5a was able to redirect the spliceosome assembly at the correct 5' donor splice site (mutated one) and inducing the generation of normal transcripts. We observed a dose-dependency between the observed rescue and the concentration of U1+5a used. Notably, the normal transcripts were 50±3% of the aberrant form with an 1,5X molar excess of U1+5a. The inability of the overexpressed U1wt to rescue splicing suggests that the correction is not mediated by the increased levels of cellular U1-snRNA, but by the ability of the U1+5a to redirect the spliceosome assembly to the mutated site. The specificity of U1+5a was further highlighted by the observation that it did not alter the splicing of the Wt minigene.

The key issue for the evaluation of the therapeutic potential of the U1+5a is the assessment of rescue at protein and function levels. To address this issue, we have devised a new full-length splicing-competent minigene. The evaluation at RNA of this novel cellular model of severe coagulation factor VII deficiency caused by the F7 IVS7+5g/a mutation, corroborated our previous findings, with the production of exon 7 skipping and partial intron retention in the mutant minigene. The levels of correct transcript in complementation assays with a molar excess of U1+5a snRNA (1,5X) accounted for 48±4% of the 37-bp aberrant form, in accordance with findings with the short minigene. Since exon 7 skipping form accounts for approximately 60% of all transcripts, we estimate that the correctly processed FVII form would represent approximately 20% of all FVII mRNAs. The assessment of rescue at protein and function levels reveals that the treatment with U1+5a is able to partially restore the FVII antigen and activity levels, that are virtually undetectable before the U1+5a loading (from protein and activity levels undetectable to 5±2,8ng/ml for antigen, and 9.5±3.2% of FVII-wt for PT activity). Because FVII deficiency would significantly benefit even from a tiny increase in coagulation factor VII levels, the correction efficacy observed has potential therapeutic implications and prompted us to further investigate the U1+5a mediated rescue.

Until now, no one has tested the effect of modified U1 snRNAs to rescue correct splicing *in vivo*. For FVII deficiency, no mouse models for splicing defects is available, so we chose to circumvent this issue by developing a model of human FVII deficiency into the mouse context. In particular, we used a normal inbred mouse strain and based our approach on the possibility to specifically detect human FVII mRNA and protein in mouse. The model developed is based on the injection in mice of a full-length splicing-competent minigene (wt or bearing the mutation F7 IVS7+5g/a) able

to produce the human FVII protein only upon correct splicing. The rescue efficacy of the U1+5a was evaluated either by hydrodynamic injection or by AAV-mediated liver transduction of the engineered U1. The FVII expression was driven by a liver specific promoter to mimic properly cell environment. Instead, the U1+5a gene was under control of its own promoter, assuring an high expression level. To increase liver transduction, serotypes (AAV2 and AAV8) with high liver tropism were used to drive the expression of genes into the mouse liver.

In a short term evaluation of the U1+5a-mediated rescue, evaluated through hydrodynamic injection experiments, the human FVII antigen level increased from undetectable in mice injected with the mutated plasmid alone to ~180ng/ml in mice treated even with the U1+5a. As expected, the human FVII level observed was dose-dependent. Compared to the human FVII level reached in mice injected with the wt minigene, the rescue effect was evaluated to be around 8.4%. The assessment of human FVII activity in mouse plasma was not achievable due to the its reduced levels compared to those of mouse FVII. At RNA level the evaluation of U1+5a-mediated rescue reveals that normal transcripts accounted for 26±9% of the partial intron retention. Notably, no traces of exon 7 skipping were observed in mice injected with the mutated minigene alone. The different splicing pattern observed in cellular and mouse model could reside into different cellular splicing environments. It is in fact known that a single pre-mRNA can generate multiple different mRNAs if expressed in different cell types (Gooding and Smith 2008; Guo, Bharmal et al. 2010).

To evaluate the prolonged correction effects we exploited the AAV-mediated delivery of both U1+5a and the mutated FVII splicing competent minigene. At protein level, the U1+5a-mediated correction was appreciable and dose-dependent, and was maintained up to 6 weeks after injection. The rescue reaches the maximum after two weeks post injection (~30ng/ml), due to the features of the AAV serotype used, and is maintained up to 6 weeks at levels (5-6 ng/ml) that, if translated into patients, would be close to the therapeutic threshold.

It must be noticed that the co-infection approach led to underestimate of rescue efficiency. In fact, the rescue occurs only in cells transduced at the same time by the vectors carrying the mutated plasmid and the modified U1+5a. Indeed, the rescue depends on the ratio between the two injected vectors, one carrying the substrate (mutated minigene) and one bearing the effector (the mutated U1snRNA). This conclusion is supported by the observation that by increasing the amount of the FVII minigene, the substrate, we observed an increase of the rescued hFVII levels. This is consistent with the hypothesis that increasing the amount of the injected mutated minigene, and thus the number of transduced liver cells, makes the rescue event more appreciable. This observation highlights the need for appropriate genetically modified animal models to properly estimate the therapeutic potential of engineered U1snRNAs. On the other hand, the approach we developed, able to provide evidence for the rescue *in vivo*, represents a valuable and relatively fast strategy to assess

therapeutic interventions in the cases in which the mutation-specific animal model is not available or very difficult to obtain.

The prolonged expression of the U1+5a through the AAV turned out to be lethal at doses above $6*10^{11}$ vg/mouse, very likely because of liver failure, as suggested by animal autopsy. It must be noticed that, in our cellular models, the U1+5a did not display a significant detrimental impact on cell viability or proliferation, very likely because of the short time window of the experiments.

The observation *in vivo* might be explained by i) alteration of the splicing-related machinery induced by overexpression of the U1snRNA it-self and/or ii) off-target effects of the mutated U1snRNA on splicing of other mRNAs.

It has been shown in cellular models that the administration of U1 adaptors designed to achieve gene silencing by targeting the 3'UTR results in the reduction of the expression of many pre-mRNA species through a mechanism of U1snRNP sequestration (Vickers, Sabripour et al. 2011). Reduction of U1snRNP could affect splicing outcome by favoring interactions of exons with the more favorable splice site consensus sequences when the U1snRNP is present in limiting quantities. Since the cellular levels of U1 in excess of what is required for splicing may be required to suppress premature cleavage and polyadenylation in introns, changes in U1 levels could have deleterious effects on the overall pre-mRNA processing (Kaida, Berg et al. 2010; Vickers, Sabripour et al. 2011). Moreover, a study on microRNA/shRNA expression in mice showed that the lethal effects was ascribable to the oversaturation of microRNA/shRNA pathways, in particular to the Exportin 5 saturation (Grimm, Streetz et al. 2006). The toxic effects observed in this study were similar to those observed by us with identical doses of the AAV8-U1+5a in mice, and were dose-related and due to liver failure (Grimm, Streetz et al. 2006). Since even the U1snRNP assembly pathway involves a cytosol phase, which translocation by e for nucleus is mediated by a multiple proteins complex, a overexpression of the U1snRNA could saturate the export/import proteins complex (with focus on proteins CBC, PHAX, Exportin 1, RanGTP, SPN1 and Importin β). This is consistent with the data obtained by overexpression of the exportin, which counteracted the toxic effects of the overexpression of the microRNA/shRNA (Yi, Doehle et al. 2005; Grimm 2011). Altogether these data point toward a toxic effects of the U1+5a related to oversaturation of the U1snRNP processing pathway rather than its binding to other mRNAs. This is consistent with the observation that low doses of AAV8-U1+5a were safe but able to rescue hFVII levels in mouse plasma.

The toxicity related to modified U1snRNAs could be circumvented by reducing their expression in the target cells for instance by optimizing the AAV doses or the regulatory elements of the U1snRNA. In our system, the U1+5a expression was driven by its own promoter, known to be a very strong polymerase II promoter. In fact, inside the cell it has been evaluated that there are about

$1*10^6$ copies of U1snRNAs (Lund and Dahlberg 1984; Egloff, O'Reilly et al. 2008). As shown by Giering et all. (Giering, Grimm et al. 2008), who expressed snRNAs under tissue-specific pol II promoter to circumvent the toxic effect observed to their specific overexpression, the expression of the modified U1snRNA by a weaker and tissue-specific pol II promoter could reduce the amount of engineered U1snRNA. Considered the coagulation factors, the expression of U1snRNA could be driven by a liver specific hAAT promoter (Hafenrichter, Ponder et al. 1994; Guo, Wang et al. 1996), thus narrowing the tissue of expression.

However, we cannot rule out the possibility that the U1snRNA targets sequences other than the F7 IVS7 donor splice site and dysregulate splicing of other genes with unpredictable effects. One of the research topic was therefore to increase the specificity of the mutated U1snRNA towards the target gene by exploring the effect of modified U1snRNAs designed to bind to intronic regions downstream of the mutated donor splice site. Due to the presence of several highly homologous repeats downstream of the IVS7 5'ss, the 9726+5g/a mutation was not approachable. To address this issue, we exploited a panel of *F9* gene mutations that impair the definition of exon 5.

Through minigene approach we demonstrated that modified U1snRNA with increased complementarity to the canonical or the mutated F9 IVS5 5'ss were able to efficiently rescue the correct FIX splicing in the presence of the donor splice site mutations at position -2 but not at positions -1T, +1A and +2C. This finding indicate that the U1snRNA-mediated rescue is effective in the presence of mutations impairing but not abolishing the 5'ss, namely those producing defective 5'ss. In fact, the three mutations at -2 position were compatible with the presence of residual levels of correct transcript even without treatment whereas the +1A and +2C mutations, affecting the virtually invariant GT dinucleotide (Sheth, Roca et al. 2006), were not. The only example of mutation at +1 position so far rescued by the U1snRNA strategy is the c.165+1G>T mutation in the *FANCC* gene, which was also shown to be associated to be compatible with synthesis of correct transcripts by it-self (Hartmann L, et al 2010). Since the mutation-specific U1 were nor capable to restore normal FVII splicing in the presence of the +1A and +2C mutations, we infer that they likely affect subsequent steps of the splicing process.

The rescue of the 5'ss mutations indicates that the disease-causing mechanism is the impairment of the interaction between U1 snRNA wt and the 5' ss by the reduction of sequence complementarity. However the compensatory U1snRNA targeting the F9 IVS5 5'ss were also able to rescue correct splicing in the presence of two mutation in the polypyrimidin tract (PPT), thus highlighting the key role of this small nuclear RNA in exon definition. We therefore tested the rescue by U1snRNAs designed to bind to intronic sequences downstream of the IVS5 5'ss. It has been shown that, to promote exon definition, the U1snRNP does not necessarily have to perfectly bind at the 5'ss. Some atypical 5'ss are recognized by U1snRNA shifted by one nucleotide (Roca and Krainer 2009) and

Cohen and colleagues (Cohen, Snow et al. 1994) have shown that U1 snRNAs base pairing at distance from the 5' ss can rescue exon inclusion induced by artificial donor site mutations. Intriguingly, albeit with variable efficacy and with the U1FIX9 being among the best, all shifted U1snRNA resulted to correct exon skipping caused by the splicing mutations occurring either at the position -2 of the 5'ss or at position -8 and -9 of the 3'ss splice sites, thus leading to the definition of Exon Specific U1 (ExSpeU1). The variable efficacy of the different modified U1snRNAs could reside either in the different maturation of the nascent U1snRNA or in the presence of unknown regulatory sequences into region covered by shifted U1snRNAs. In fact modification of U1 snRNAs 5'-tail, either with point mutations (Zhuang and Weiner 1986; Susani, Pangrazio et al. 2004) or introduction of several tens of nucleotides (De Angelis, Sthandier et al. 2002; Denti, Rosa 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). Due to the modifications introduced in their 5'-tail, some engineered U1 snRNPs might not be effectively processed or exported into the nucleus.

To analyze the rescue effect at protein level, we focused our attention to U1FIX9 snRNA and a new designed full-length splicing-competent minigene. The new plasmid was less efficient in comparison to the *in vivo* situation and this is probably due to its sequence context (Romano, Marcucci et al. 2002). Through this approach we demonstrated that the ExSpeU1 U1FIX 9 induced a nearly complete exon inclusion in the presence of the three synonymous mutations at position -2 and of the two PPT transversions (-8G and -9G). Restoration of correct splicing corresponded to synthesis of the full-length factor IX protein, with secreted protein levels and coagulant activity comparable to that of the FIXwt construct.

Taken together these results have provided for the first time experimental evidence for the ability of modified U1snRNA to rescue correct splicing and protein biosynthesis in the presence of mutations located either in the donor splice site or acceptor splice site of coagulation F7 and F9 genes. Noticeably, we demonstrated that a single ExSpeU1, targeting gene specific intronic sequences, efficiently restore splicing in the presence of different mutations, thus extending the potential therapeutic application to a group of patients. Finally, the data in the mouse model represent a proof-of-principle that the U1snRNA-mediated rescue is effective *in vivo*.

The rescue of secreted coagulation FVII and FIX levels obtained both *in vitro* and *in vivo* could have, if translated in patients, a therapeutic impact, thus encouraging further studies aimed at proposing the modified U1snRNA, and particularly the ExSpeu1, as innovative treatment for coagulation factor disease caused by aberrant splicing. This approach could be extended to other human genetic disorders associated to aberrant splicing, a frequent cause of clinically severe disease forms.

# VI. Publications

The results from this PhD thesis have been published in three papers:

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

> Small nuclear U1-RNAs (snRNAs), the spliceosome components selectively recognizing donor splice sites (5'ss), were engineered to restore correct mRNA processing in a cellular model of severe coagulation factor VII (FVII) deficiency, caused by the IVS7 9726 + 5g/a change. Three U1-snRNAs, complementary to the mutated 5'ss (U1 + 5a) or to neighboring sequences were expressed with FVII minigenes in a hepatoma cell line. The U1-snRNAs reduced from 80% to 40% the exon 7 skipping, thus increasing exon definition. The U1 + 5a construct also dramatically increased recognition of the correct 5'ss over the 37-bp downstream cryptic site preferentially activated by the mutation, thus inducing appreciable synthesis of normal transcripts (from barely detectable to 50%). This effect, which was dose-dependent, clearly demonstrated that impaired recognition by the U1-snRNA was the mechanism responsible for FVII deficiency. These findings suggest compensatory U1-snRNAs as therapeutic tools in coagulation factor deficiencies caused by mutations at 5'ss, a frequent cause of severe defects.

Pinotti M, **Balestra D**, Rizzotto L, Maestri I, Pagani F, Bernardi F. (2009). "Rescue of coagulation factor VII function by the U1+5A snRNA." Blood **113**(25): 6461-6464.

> Our previous studies with genomic minigenes have demonstrated that an engineered small nuclear RNA-U1 (U1+5a) partially rescued coagulation factor VII (FVII) mRNA processing impaired by the 9726+5G>A mutation. Here, to evaluate the U1+5a effects on FVII function, we devised a full-length FVII splicing-competent construct (pSCFVII-wt). This construct drove in COS-1 cells the synthesis of properly processed FVII transcripts and of secreted functional FVII (23 +/- 4 ng/mL), which were virtually undetectable upon introduction of the 9726+5G>A mutation (pSCFVII-9726+5a). Cotransfection of pSCFVII-9726+5a with pU1+5a resulted in a partial rescue of FVII splicing and protein biosynthesis. The level increase in medium was dose dependent and, with a molar excess (1.5x) of pU1+5a, reached 9.5% plus or minus 3.2% (5.0 +/- 2.8 ng/mL) of FVII-wt coagulant activity. These data provide the first insights into the U1-snRNA-mediated rescue of donor splice sites at protein level, thus further highlighting its therapeutic implications in bleeding disorders, which would benefit even from tiny increase of functional levels.

Fernandez Alanis E †, Pinotti M†, Dal Mas A†, **Balestra D**, Cavallari N, Rogalska M E, Bernardi F, Pagani F. "An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects." <u>Human Molecular Genetics</u> accepted (09-02-2012).
†The authors wish to be known that, in their opinion, the first three authors should be regarded as join First Authors

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

# VII. References

Achsel, T., H. Brahms, et al. (1999). "A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro." *EMBO J* **18**(20): 5789-5802.

Adams, D. J., L. van der Weyden, et al. (2001). "ZNF265--a novel spliceosomal protein able to induce alternative splicing." *J Cell Biol* **154**(1): 25-32.

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

Alexander, B., R. Goldstein, et al. (1951). "Congenital SPCA deficiency: a hitherto unrecognized coagulation defect with hemorrhage rectified by serum and serum fractions." *J Clin Invest* **30**(6): 596-608.

Anson, D. S., K. H. Choo, et al. (1984). "The gene structure of human anti-haemophilic factor IX." *EMBO J* **3**(5): 1053-1060.

Asokan, A., D. V. Schaffer, et al. (2012). "The AAV Vector Toolkit: Poised at the Clinical Crossroads." *Mol Ther*.

Asparuhova, M. B., G. Marti, et al. (2007). "Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing Tat and Rev exon skipping." *J Gene Med* **9**(5): 323-334.

Ast, G. (2004). "How did alternative splicing evolve?" *Nat Rev Genet* **5**(10): 773-782.

Aznarez, I., E. M. Chan, et al. (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." *Hum Mol Genet* **12**(16): 2031-2040.

Babich, A., J. R. Nevins, et al. (1980). "Early capping of transcripts from the adenovirus major late transcription unit." *Nature* **287**(5779): 246-248.

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

Balleisen, L., G. Assmann, et al. (1985). "Epidemiological study on factor VII, factor VIII and fibrinogen in an industrial population--II. Baseline data on the relation to blood pressure, blood glucose, uric acid, and lipid fractions." *Thromb Haemost* **54**(3): 721-723.

Baralle, D. and M. Baralle (2005). "Splicing in action: assessing disease causing sequence changes." *J Med Genet* **42**(10): 737-748.

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

Baralle, M., D. Baralle, et al. (2003). "Identification of a mutation that perturbs NF1 agene splicing using genomic DNA samples and a minigene assay." *J Med Genet* **40**(3): 220-222.

Battle, D. J., M. Kasim, et al. (2006). "The SMN complex: an assembly machine for RNPs." *Cold Spring Harb Symp Quant Biol* **71**: 313-320.

Battle, D. J., C. K. Lau, et al. (2006). "The Gemin5 protein of the SMN complex identifies snRNAs." *Mol Cell* **23**(2): 273-279.

Bauer, K. A. (1997). "Activation of the factor VII-tissue factor pathway." *Thromb Haemost* **78**(1): 108-111.

Behrens, S. E. and R. Luhrmann (1991). "Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells." *Genes Dev* **5**(8): 1439-1452.

Bennett, J. S., G. Vilaire, et al. (1982). "Identification of the fibrinogen receptor on human platelets by photoaffinity labeling." *J Biol Chem* **257**(14): 8049-8054.

Berget, S. M. (1995). "Exon recognition in vertebrate splicing." *J Biol Chem* **270**(6): 2411-2414.

Berglund, J. A., K. Chua, et al. (1997). "The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC." *Cell* **89**(5): 781-787.

Berkner, K., S. Busby, et al. (1986). "Isolation and expression of cDNAs encoding human factor VII." *Cold Spring Harb Symp Quant Biol* **51 Pt 1**: 531-541.

Bernardi, F., P. Arcieri, et al. (1997). "Contribution of factor VII genotype to activated FVII levels. Differences in genotype frequencies between northern and southern European populations." *Arterioscler Thromb Vasc Biol* **17**(11): 2548-2553.

Bernardi, F., G. Marchetti, et al. (1996). "Factor VII gene polymorphisms contribute about one third of the factor VII level variation in plasma." *Arterioscler Thromb Vasc Biol* **16**(1): 72-76.

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

Bindereif, A. and M. R. Green (1987). "An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly." *EMBO J* **6**(8): 2415-2424.

Bjoern, S., D. C. Foster, et al. (1991). "Human plasma and recombinant factor VII. Characterization of O-glycosylations at serine residues 52 and 60 and effects of site-directed mutagenesis of serine 52 to alanine." *J Biol Chem* **266**(17): 11051-11057.

Black, D. L., B. Chabot, et al. (1985). "U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing." *Cell* **42**(3): 737-750.

Bom, V. J. and R. M. Bertina (1990). "The contributions of Ca2+, phospholipids and tissue-factor apoprotein to the activation of human blood-coagulation factor X by activated factor VII." *Biochem J* **265**(2): 327-336.

Bonetta, L. (2009). "RNA-based therapeutics: ready for delivery?" *Cell* **136**(4): 581-584.

Borensztajn, K., M. L. Sobrier, et al. (2006). "Oriented scanning is the leading mechanism underlying 5' splice site selection in mammals." *PLoS Genet* **2**(9): e138.

Borensztajn, K., M. L. Sobrier, et al. (2003). "Factor VII gene intronic mutation in a lethal factor VII deficiency: effects on splice-site selection." *Blood* **102**(2): 561-563.

Bowen, D. J. (2002). "Haemophilia A and haemophilia B: molecular insights." *Mol Pathol* **55**(2): 127-144.

Branlant, C., A. Krol, et al. (1982). "U2 RNA shares a structural domain with U1, U4, and U5 RNAs." *EMBO J* **1**(10): 1259-1265.

Bringmann, P. and R. Luhrmann (1986). "Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents." *EMBO J* **5**(13): 3509-3516.

Bromberg, M. E., W. H. Konigsberg, et al. (1995). "Tissue factor promotes melanoma metastasis by a pathway independent of blood coagulation." *Proc Natl Acad Sci U S A* **92**(18): 8205-8209.

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

Butenas, S. and K. G. Mann (1996). "Kinetics of human factor VII activation." *Biochemistry* **35**(6): 1904-1910.

Cao, W. and M. A. Garcia-Blanco (1998). "A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding." *J Biol Chem* **273**(32): 20629-20635.

Carbon, P., S. Murgo, et al. (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**(1): 71-79.

Carcao, M. D. and L. Aledort (2004). "Prophylactic factor replacement in hemophilia." *Blood Rev* **18**(2): 101-113.

Carmel, I., S. Tal, et al. (2004). "Comparative analysis detects dependencies among the 5' splice-site positions." *RNA* **10**(5): 828-840.

Chaing, S. H., A. Wallmark, et al. (1994). "A NlaIII polymorphism within the human factor VII gene." *Hum Genet* **93**(6): 722-723.

Chanfreau, G. and A. Jacquier (1993). "Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron." *EMBO J* **12**(13): 5173-5180.

Cho, S., A. Hoang, et al. (2011). "Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly." *Proc Natl Acad Sci U S A* **108**(20): 8233-8238.

Cohen, J. B., J. E. Snow, et al. (1994). "Suppression of mammalian 5' splice-site defects by U1 small nuclear RNAs from a distance." *Proc Natl Acad Sci U S A* **91**(22): 10470-10474.

Cooper, D. N. (1991). "The molecular genetics of familial venous thrombosis." *Blood Rev* **5**(1): 55-70.

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

Coppola, J. A., A. S. Field, et al. (1983). "Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped." *Proc Natl Acad Sci U S A* **80**(5): 1251-1255.

Dahlback, B. (2000). "Blood coagulation." *Lancet* **355**(9215): 1627-1632.

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

Davie, E. W. (2003). "A brief historical review of the waterfall/cascade of blood coagulation." *J Biol Chem* **278**(51): 50819-50832.

Davie, E. W., K. Fujikawa, et al. (1991). "The coagulation cascade: initiation, maintenance, and regulation." *Biochemistry* **30**(43): 10363-10370.

Davie, E. W. and O. D. Ratnoff (1964). "Waterfall Sequence for Intrinsic Blood Clotting." *Science* **145**: 1310-1312.

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

Denti, M. A., A. Rosa, et al. (2006). "Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model." *Proc Natl Acad Sci U S A* **103**(10): 3758-3763.

Derian, C. K., W. VanDusen, et al. (1989). "Inhibitors of 2-ketoglutarate-dependent dioxygenases block aspartyl beta-hydroxylation of recombinant human factor IX in several mammalian expression systems." *J Biol Chem* **264**(12): 6615-6618.

Dietz, H. C., D. Valle, et al. (1993). "The skipping of constitutive exons in vivo induced by nonsense mutations." *Science* **259**(5095): 680-683.

Domitrovich, A. M. and G. R. Kunkel (2003). "Multiple, dispersed human U6 small nuclear RNA genes with varied transcriptional efficiencies." *Nucleic Acids Res* **31**(9): 2344-2352.

Dorner, A. J., D. G. Bole, et al. (1987). "The relationship of N-linked glycosylation and heavy chain-binding protein association with the secretion of glycoproteins." *J Cell Biol* **105**(6 Pt 1): 2665-2674.

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

Egloff, S., D. O'Reilly, et al. (2007). "Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression." *Science* **318**(5857): 1777-1779.

Egloff, S., D. O'Reilly, et al. (2008). "Expression of human snRNA genes from beginning to end." *Biochem Soc Trans* **36**(Pt 4): 590-594.

Eichinger, S., P. M. Mannucci, et al. (1995). "Determinants of plasma factor VIIa levels in humans." *Blood* **86**(8): 3021-3025.

Eigenbrot, C., D. Kirchhofer, et al. (2001). "The factor VII zymogen structure reveals reregistration of beta strands during activation." *Structure* **9**(7): 627-636.

Eliceiri, G. L. and T. Gurney, Jr. (1978). "Subcellular location of precursors to small nuclear RNA species C and D and of newly synthesized 5 S RNA in HeLa cells." *Biochem Biophys Res Commun* **81**(3): 915-919.

Eliceiri, G. L. and M. S. Sayavedra (1976). "Small RNAs in the nucleus and cytoplasm of HeLa cells." *Biochem Biophys Res Commun* **72**(2): 507-512.

Erdmann, D. and J. Heim (1995). "Orphan nuclear receptor HNF-4 binds to the human coagulation factor VII promoter." *J Biol Chem* **270**(39): 22988-22996.

Fabrizio, P., S. Esser, et al. (1994). "Isolation of S. cerevisiae snRNPs: comparison of U1 and U4/U6.U5 to their human counterparts." *Science* **264**(5156): 261-265.

Ford, E., M. Strubin, et al. (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 Dev* **12**(22): 3528-3540.

Fragall, C. T., A. M. Adams, et al. (2011). "Mismatched single stranded antisense oligonucleotides can induce efficient dystrophin splice switching." *BMC Med Genet* **12**: 141.

Furie, B. and B. C. Furie (1988). "The molecular basis of blood coagulation." *Cell* **53**(4): 505-518.

Furie, B. and B. C. Furie (1992). "Molecular and cellular biology of blood coagulation." *N Engl J Med* **326**(12): 800-806.

Geib, T. and K. J. Hertel (2009). "Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs." *PLoS One* **4**(12): e8204.

Giannelli, F. and P. M. Green (1996). "The molecular basis of haemophilia A and B." *Baillieres Clin Haematol* **9**(2): 211-228.

Giering, J. C., D. Grimm, et al. (2008). "Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic." *Mol Ther* **16**(9): 1630-1636.

Gilbert, W. (1978). "Why genes in pieces?" *Nature* **271**(5645): 501.

Glaus, E., F. Schmid, et al. (2011). "Gene therapeutic approach using mutation-adapted U1 snRNA to correct a RPGR splice defect in patient-derived cells." *Mol Ther* **19**(5): 936-941.

Gooding, C., F. Clark, et al. (2006). "A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones." *Genome Biol* **7**(1): R1.

Gooding, C. and C. W. Smith (2008). "Tropomyosin exons as models for alternative splicing." *Adv Exp Med Biol* **644**: 27-42.

Gorman, L., D. Suter, et al. (1998). "Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs." *Proc Natl Acad Sci U S A* **95**(9): 4929-4934.

Goyenvalle, A. and K. E. Davies (2011). "Engineering exon-skipping vectors expressing U7 snRNA constructs for Duchenne muscular dystrophy gene therapy." *Methods Mol Biol* **709**: 179-196.

Goyenvalle, A., A. Vulin, et al. (2004). "Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping." *Science* **306**(5702): 1796-1799.

Gozani, O., R. Feld, et al. (1996). "Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A." *Genes Dev* **10**(2): 233-243.

Green, F., C. Kelleher, et al. (1991). "A common genetic polymorphism associated with lower coagulation factor VII levels in healthy individuals." *Arterioscler Thromb* **11**(3): 540-546.

Greenberg, D., C. H. Miao, et al. (1995). "Liver-specific expression of the human factor VII gene." *Proc Natl Acad Sci U S A* **92**(26): 12347-12351.

Gridasova, A. A. and R. W. Henry (2005). "The p53 tumor suppressor protein represses human snRNA gene transcription by RNA polymerases II and III independently of sequence-specific DNA binding." *Mol Cell Biol* **25**(8): 3247-3260.

Grimm, C., B. Stefanovic, et al. (1993). "The low abundance of U7 snRNA is partly determined by its Sm binding site." *EMBO J* **12**(3): 1229-1238.

Grimm, D. (2011). "The dose can make the poison: lessons learned from adverse in vivo toxicities caused by RNAi overexpression." *Silence* **2**: 8.

Grimm, D., K. L. Streetz, et al. (2006). "Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways." *Nature* **441**(7092): 537-541.

Guo, W., S. J. Bharmal, et al. (2010). "Titin diversity--alternative splicing gone wild." *J Biomed Biotechnol* **2010**: 753675.

Guo, Z. S., L. H. Wang, et al. (1996). "Evaluation of promoter strength for hepatic gene expression in vivo following adenovirus-mediated gene transfer." *Gene Ther* **3**(9): 802-810.

Hafenrichter, D. G., K. P. Ponder, et al. (1994). "Liver-directed gene therapy: evaluation of liver specific promoter elements." *J Surg Res* **56**(6): 510-517.

Hagen, F. S., C. L. Gray, et al. (1986). "Characterization of a cDNA coding for human factor VII." *Proc Natl Acad Sci U S A* **83**(8): 2412-2416.

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

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

Hansen, C. B., C. Pyke, et al. (2001). "Tissue factor-mediated endocytosis, recycling, and degradation of factor VIIa by a clathrin-independent mechanism not requiring the cytoplasmic domain of tissue factor." *Blood* **97**(6): 1712-1720.

Hartmann, L., K. Neveling, 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." *Am J Hum Genet* **87**(4): 480-493.

Hartmann, L., S. Theiss, et al. (2008). "Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases?" *Front Biosci* **13**: 3252-3272.

Hasselback, R. and P. F. Hjort (1960). "Effect of heparin on in vivo turnover of clotting factor." *J Appl Physiol* **15**: 945-948.

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

Helenius, A. (1994). "How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum." *Mol Biol Cell* **5**(3): 253-265.

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

Heyd, F. and K. W. Lynch (2010). "Phosphorylation-dependent regulation of PSF by GSK3 controls CD45 alternative splicing." *Mol Cell* **40**(1): 126-137.

High, K. A. (2011). "Gene therapy for haemophilia: a long and winding road." *J Thromb Haemost* **9 Suppl 1**: 2-11.

Hirsch, H. A., L. Gu, et al. (2000). "The retinoblastoma tumor suppressor protein targets distinct general transcription factors to regulate RNA polymerase III gene expression." *Mol Cell Biol* **20**(24): 9182-9191.

Hitomi, Y., K. Sugiyama, et al. (1998). "Suppression of the 5' splice site mutation in the Nagase analbuminemic rat with mutated U1snRNA." *Biochem Biophys Res Commun* **251**(1): 11-16.

Hong, W., M. Bennett, et al. (1997). "Association of U2 snRNP with the spliceosomal complex E." *Nucleic Acids Res* **25**(2): 354-361.

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

Houdayer, C., C. Dehainault, et al. (2008). "Evaluation of in silico splice tools for decision-making in molecular diagnosis." *Hum Mutat* **29**(7): 975-982.

Hu, C. and G. S. Lipshutz (2012). "AAV-based neonatal gene therapy for hemophilia A: long-term correction and avoidance of immune responses in mice." *Gene Ther*.

Huang, Q., M. R. Jacobson, et al. (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." *Mol Cell Biol* **17**(12): 7178-7185.

Humphries, S. E., A. Lane, et al. (1992). "The study of gene-environment interactions that influence thrombosis and fibrinolysis. Genetic variation at the loci for factor VII and plasminogen activator inhibitor-1." _Arch Pathol Lab Med_ **116**(12): 1322-1329.

Hunault, M., A. A. Arbini, et al. (1999). "Characterization of two naturally occurring mutations in the second epidermal growth factor-like domain of factor VII." _Blood_ **93**(4): 1237-1244.

Hunault, M., A. A. Arbini, et al. (1997). "The Arg353Gln polymorphism reduces the level of coagulation factor VII. In vivo and in vitro studies." _Arterioscler Thromb Vasc Biol_ **17**(11): 2825-2829.

Iakhiaev, A., U. R. Pendurthi, et al. (1999). "Catabolism of factor VIIa bound to tissue factor in fibroblasts in the presence and absence of tissue factor pathway inhibitor." _J Biol Chem_ **274**(52): 36995-37003.

Incitti, T., F. G. De Angelis, et al. (2010). "Exon skipping and duchenne muscular dystrophy therapy: selection of the most active U1 snRNA antisense able to induce dystrophin exon 51 skipping." _Mol Ther_ **18**(9): 1675-1682.

Izaurralde, E., J. Lewis, et al. (1995). "A cap-binding protein complex mediating U snRNA export." _Nature_ **376**(6542): 709-712.

Izaurralde, E., J. Stepinski, et al. (1992). "A cap binding protein that may mediate nuclear export of RNA polymerase II-transcribed RNAs." _J Cell Biol_ **118**(6): 1287-1295.

Jacobs, E. Y., I. Ogiwara, et al. (2004). "Role of the C-terminal domain of RNA polymerase II in U2 snRNA transcription and 3' processing." _Mol Cell Biol_ **24**(2): 846-855.

Jamison, S. F. and M. A. Garcia-Blanco (1992). "An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract." _Proc Natl Acad Sci U S A_ **89**(12): 5482-5486.

Jamison, S. F., Z. Pasman, et al. (1995). "U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements." _Nucleic Acids Res_ **23**(16): 3260-3267.

Jesty, J. and Y. Nemerson (1974). "Purification of Factor VII from bovine plasma. Reaction with tissue factor and activation of Factor X." _J Biol Chem_ **249**(2): 509-515.

Jin, J., L. Perera, et al. (2001). "Four loops of the catalytic domain of factor viia mediate the effect of the first EGF-like domain substitution on factor viia catalytic activity." _J Mol Biol_ **307**(5): 1503-1517.

Jurica, M. S. and M. J. Moore (2003). "Pre-mRNA splicing: awash in a sea of proteins." _Mol Cell_ **12**(1): 5-14.

Kaida, D., M. G. Berg, et al. (2010). "U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation." _Nature_ **468**(7324): 664-668.

Kandels-Lewis, S. and B. Seraphin (1993). "Involvement of U6 snRNA in 5' splice site selection." _Science_ **262**(5142): 2035-2039.

Kane, W. H. and E. W. Davie (1988). "Blood coagulation factors V and VIII: structural and functional similarities and their relationship to hemorrhagic and thrombotic disorders." _Blood_ **71**(3): 539-555.

Kaufman, R. J. (1998). "Post-translational modifications required for coagulation factor secretion and function." _Thromb Haemost_ **79**(6): 1068-1079.

Kelly, C. R., C. D. Dickinson, et al. (1997). "Ca2+ binding to the first epidermal growth factor module of coagulation factor VIIa is important for cofactor interaction and proteolytic function." _J Biol Chem_ **272**(28): 17467-17472.

Kemball-Cook, G., D. J. Johnson, et al. (1999). "Crystal structure of active site-inhibited human coagulation factor VIIa (des-Gla)." _J Struct Biol_ **127**(3): 213-223.

Kent, O. A., D. B. Ritchie, et al. (2005). "Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway." _Mol Cell Biol_ **25**(1): 233-240.

Ketterling, R. P., J. B. Drost, et al. (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." _Hum Mutat_ **13**(3): 221-231.

Khachigian, L. M., V. Lindner, et al. (1996). "Egr-1-induced endothelial gene expression: a common theme in vascular injury." *Science* **271**(5254): 1427-1431.

Kisiel, W., K. Fujikawa, et al. (1977). "Activation of bovine factor VII (proconvertin) by factor XIIa (activated Hageman factor)." *Biochemistry* **16**(19): 4189-4194.

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

Kitao, S., A. Segref, et al. (2008). "A compartmentalized phosphorylation/dephosphorylation system that regulates U snRNA export from the nucleus." *Mol Cell Biol* **28**(1): 487-497.

Koeberl, D. D., C. D. Bottema, et al. (1990). "Mutations causing hemophilia B: direct estimate of the underlying rates of spontaneous germ-line transitions, transversions, and deletions in a human gene." *Am J Hum Genet* **47**(2): 202-217.

Kohler, A. and E. Hurt (2007). "Exporting RNA from the nucleus to the cytoplasm." *Nat Rev Mol Cell Biol* **8**(10): 761-773.

Kolb, S. J., D. J. Battle, et al. (2007). "Molecular functions of the SMN complex." *J Child Neurol* **22**(8): 990-994.

Komiyama, Y., A. H. Pedersen, et al. (1990). "Proteolytic activation of human factors IX and X by recombinant human factor VIIa: effects of calcium, phospholipids, and tissue factor." *Biochemistry* **29**(40): 9418-9425.

Konarska, M. M., J. Vilardell, et al. (2006). "Repositioning of the reaction intermediate within the catalytic center of the spliceosome." *Mol Cell* **21**(4): 543-553.

Krawczak, M., J. Reiss, et al. (1992). "The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences." *Hum Genet* **90**(1-2): 41-54.

Krawczak, M., N. S. Thomas, et al. (2007). "Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing." *Hum Mutat* **28**(2): 150-158.

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

Kumar, A. and D. S. Fair (1993). "Specific molecular interaction sites on factor VII involved in factor X activation." *Eur J Biochem* **217**(2): 509-518.

Lamond, A. I., M. M. Konarska, et al. (1987). "A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation." *Genes Dev* **1**(6): 532-543.

Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." *Nature* **409**(6822): 860-921.

Lawson, J. H., S. Butenas, et al. (1992). "The evaluation of complex-dependent alterations in human factor VIIa." *J Biol Chem* **267**(7): 4834-4843.

Lesser, C. F. and C. Guthrie (1993). "Mutations in U6 snRNA that alter splice site specificity: implications for the active site." *Science* **262**(5142): 1982-1988.

Liao, X. L., L. Kretzner, et al. (1990). "Universally conserved and yeast-specific U1 snRNA sequences are important but not essential for U1 snRNP function." *Genes Dev* **4**(10): 1766-1774.

Liu, J. L. and J. G. Gall (2007). "U bodies are cytoplasmic structures that contain uridine-rich small nuclear ribonucleoproteins and associate with P bodies." *Proc Natl Acad Sci U S A* **104**(28): 11655-11659.

Liu, S., R. Rauhut, et al. (2006). "The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP." *RNA* **12**(7): 1418-1430.

Lopez-Bigas, N., B. Audit, et al. (2005). "Are splicing mutations the most frequent cause of hereditary disease?" *FEBS Lett* **579**(9): 1900-1903.

Lopez, J. A., D. W. Chung, et al. (1988). "The alpha and beta chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence." *Proc Natl Acad Sci U S A* **85**(7): 2135-2139.

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

Macfarlane, R. G. (1964). "An Enzyme Cascade in the Blood Clotting Mechanism, and Its Function as a Biochemical Amplifier." _Nature_ **202**: 498-499.

Madhani, H. D. and C. Guthrie (1992). "A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome." _Cell_ **71**(5): 803-817.

Madocsai, C., S. R. Lim, et al. (2005). "Correction of SMN2 Pre-mRNA splicing by antisense U7 small nuclear RNAs." _Mol Ther_ **12**(6): 1013-1022.

Malca, H., N. Shomron, et al. (2003). "The U1 snRNP base pairs with the 5' splice site within a penta-snRNP complex." _Mol Cell Biol_ **23**(10): 3442-3455.

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

Mann, K. G. (1999). "Biochemistry and physiology of blood coagulation." _Thromb Haemost_ **82**(2): 165-174.

Mann, K. G., M. E. Nesheim, et al. (1990). "Surface-dependent reactions of the vitamin K-dependent enzyme complexes." _Blood_ **76**(1): 1-16.

Mannucci, P. M., S. Duga, et al. (2004). "Recessively inherited coagulation disorders." _Blood_ **104**(5): 1243-1252.

Margaritis, P., V. R. Arruda, et al. (2004). "Novel therapeutic approach for hemophilia using gene delivery of an engineered secreted activated Factor VII." _J Clin Invest_ **113**(7): 1025-1031.

Mariani, G., F. H. Herrmann, et al. (2000). "Clinical manifestations, management, and molecular genetics in congenital factor VII deficiency: the International Registry on Congenital Factor VII Deficiency (IRF7)." _Blood_ **96**(1): 374.

Mariani, G., F. H. Herrmann, et al. (2005). "Clinical phenotypes and factor VII genotype in congenital factor VII deficiency." _Thromb Haemost_ **93**(3): 481-487.

Mariani, G., L. Lo Coco, et al. (1998). "Molecular and clinical aspects of factor VII deficiency." _Blood Coagul Fibrinolysis_ **9 Suppl 1**: S83-88.

Massenet, S., L. Pellizzoni, et al. (2002). "The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway." _Mol Cell Biol_ **22**(18): 6533-6541.

Mayes, A. E., L. Verdone, et al. (1999). "Characterization of Sm-like proteins in yeast and their association with U6 snRNA." _EMBO J_ **18**(15): 4321-4331.

McVey, J. H., E. Boswell, et al. (2001). "Factor VII deficiency and the FVII mutation database." _Hum Mutat_ **17**(1): 3-17.

Meade, T. W., A. P. Haines, et al. (1983). "Menopausal status and haemostatic variables." _Lancet_ **1**(8314-5): 22-24.

Medlin, J. E., P. Uguen, et al. (2003). "The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA." _EMBO J_ **22**(4): 925-934.

Meilahn, E. N., L. H. Kuller, et al. (1992). "Hemostatic factors according to menopausal status and use of hormone replacement therapy." _Ann Epidemiol_ **2**(4): 445-455.

Mermoud, J. E., P. T. Cohen, et al. (1994). "Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism." _EMBO J_ **13**(23): 5679-5688.

Mittal, V., M. A. Cleary, et al. (1996). "The Oct-1 POU-specific domain can stimulate small nuclear RNA gene transcription by stabilizing the basal transcription complex SNAPc." _Mol Cell Biol_ **16**(5): 1955-1965.

Montejo, J. M., M. Magallon, et al. (1999). "Identification of twenty-one new mutations in the factor IX gene by SSCP analysis." _Hum Mutat_ **13**(2): 160-165.

Mouaikel, J., U. Narayanan, et al. (2003). "Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron." _EMBO Rep_ **4**(6): 616-622.

Mount, S. M., I. Pettersson, et al. (1983). "The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro." _Cell_ **33**(2): 509-518.

Muller, B. and D. Schumperli (1997). "The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism." *Semin Cell Dev Biol* **8**(6): 567-576.

Mullis, K. B. and F. A. Faloona (1987). "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." *Methods Enzymol* **155**: 335-350.

Murphy, S. L. and K. A. High (2008). "Gene therapy for haemophilia." *Br J Haematol* **140**(5): 479-487.

Nagai, K., Y. Muto, et al. (2001). "Structure and assembly of the spliceosomal snRNPs. Novartis Medal Lecture." *Biochem Soc Trans* **29**(Pt 2): 15-26.

Nakagaki, T., D. C. Foster, et al. (1991). "Initiation of the extrinsic pathway of blood coagulation: evidence for the tissue factor dependent autoactivation of human coagulation factor VII." *Biochemistry* **30**(45): 10819-10824.

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

Nelsestuen, G. L. and J. W. Suttie (1972). "The purification and properties of an abnormal prothrombin protein produced by dicumarol-treated cows. A comparison to normal prothrombin." *J Biol Chem* **247**(24): 8176-8182.

Nemerson, Y. (1966). "The reaction between bovine brain tissue factor and factors VII and X." *Biochemistry* **5**(2): 601-608.

Nicolaisen, E. M., L. Thim, et al. (1993). "FVIIa derivatives obtained by autolytic and controlled cathepsin G mediated cleavage." *FEBS Lett* **317**(3): 245-249.

Nilsen, T. W. (2002). "The spliceosome: no assembly required?" *Mol Cell* **9**(1): 8-9.

Nilsen, T. W. and B. R. Graveley (2010). "Expansion of the eukaryotic proteome by alternative splicing." *Nature* **463**(7280): 457-463.

O'Hara, P. J. and F. J. Grant (1988). "The human factor VII gene is polymorphic due to variation in repeat copy number in a minisatellite." *Gene* **66**(1): 147-158.

O'Hara, P. J., F. J. Grant, et al. (1987). "Nucleotide sequence of the gene coding for human factor VII, a vitamin K-dependent protein participating in blood coagulation." *Proc Natl Acad Sci U S A* **84**(15): 5158-5162.

Ohno, K., A. Tsujino, et al. (2001). "Choline acetyltransferase mutations cause myasthenic syndrome associated with episodic apnea in humans." *Proc Natl Acad Sci U S A* **98**(4): 2017-2022.

Ohno, M., A. Segref, et al. (2000). "PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation." *Cell* **101**(2): 187-198.

Ortolano, S., C. Spuch, et al. (2012). "Present and future of adeno associated virus based gene therapy approaches." *Recent Pat Endocr Metab Immune Drug Discov* **6**(1): 47-66.

Osterud, B. and S. I. Rapaport (1977). "Activation of factor IX by the reaction product of tissue factor and factor VII: additional pathway for initiating blood coagulation." *Proc Natl Acad Sci U S A* **74**(12): 5260-5264.

Ott, I., Y. Miyagi, et al. (2000). "Reversible regulation of tissue factor-induced coagulation by glycosyl phosphatidylinositol-anchored tissue factor pathway inhibitor." *Arterioscler Thromb Vasc Biol* **20**(3): 874-882.

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

Pagani, F., E. Buratti, et al. (2002). "A new type of mutation causes a splicing defect in ATM." *Nat Genet* **30**(4): 426-429.

Pagani, F., E. Buratti, et al. (2000). "Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element." *J Biol Chem* **275**(28): 21041-21047.

Pagani, F., C. Stuani, et al. (2003). "New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12." *Hum Mol Genet* **12**(10): 1111-1120.

Pan, Q., O. Shai, et al. (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." *Nat Genet* **40**(12): 1413-1415.

Patel, A. A. and J. A. Steitz (2003). "Splicing double: insights from the second spliceosome." *Nat Rev Mol Cell Biol* **4**(12): 960-970.

Patel, S. B., N. Novikova, et al. (2007). "Splicing-independent recruitment of spliceosomal small nuclear RNPs to nascent RNA polymerase II transcripts." *J Cell Biol* **178**(6): 937-949.

Patthy, L. (1985). "Evolution of the proteases of blood coagulation and fibrinolysis by assembly from modules." *Cell* **41**(3): 657-663.

Pedersen, A. H., T. Lund-Hansen, et al. (1989). "Autoactivation of human recombinant coagulation factor VII." *Biochemistry* **28**(24): 9331-9336.

Pellizzoni, L., J. Yong, et al. (2002). "Essential role for the SMN complex in the specificity of snRNP assembly." *Science* **298**(5599): 1775-1779.

Petrus, I., M. Chuah, et al. (2010). "Gene therapy strategies for hemophilia: benefits versus risks." *J Gene Med* **12**(10): 797-809.

Peyvandi, F., P. M. Mannucci, et al. (2000). "A novel polymorphism in intron 1a of the human factor VII gene (G73A): study of a healthy Italian population and of 190 young survivors of myocardial infarction." *Br J Haematol* **108**(2): 247-253.

Pfeiffer, R. A., R. Ott, et al. (1982). "Deficiency of coagulation factors VII and X associated with deletion of a chromosome 13 (q34). Evidence from two cases with 46,XY,t(13;Y)(q11;q34)." *Hum Genet* **62**(4): 358-360.

Pike, A. C., A. M. Brzozowski, et al. (1999). "Structure of human factor VIIa and its implications for the triggering of blood coagulation." *Proc Natl Acad Sci U S A* **96**(16): 8925-8930.

Pinotti, M., D. Balestra, et al. (2009). "Rescue of coagulation factor VII function by the U1+5A snRNA." *Blood* **113**(25): 6461-6464.

Pinotti, M., F. Bernardi, et al. (2011). "RNA-based therapeutic approaches for coagulation factor deficiencies." *J Thromb Haemost* **9**(11): 2143-2152.

Pinotti, M., L. Rizzotto, et al. (2008). "U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency." *Blood* **111**(5): 2681-2684.

Pinotti, M., R. Toso, et al. (2000). "Modulation of factor VII levels by intron 7 polymorphisms: population and in vitro studies." *Blood* **95**(11): 3423-3428.

Pinotti, M., R. Toso, et al. (1998). "Molecular mechanisms of FVII deficiency: expression of mutations clustered in the IVS7 donor splice site of factor VII gene." *Blood* **92**(5): 1646-1651.

Pipe, S. W., K. A. High, et al. (2008). "Progress in the molecular biology of inherited bleeding disorders." *Haemophilia* **14 Suppl 3**: 130-137.

Plessel, G., U. Fischer, et al. (1994). "m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain." *Mol Cell Biol* **14**(6): 4160-4172.

Pohlenz, J., A. Dumitrescu, et al. (2002). "Congenital secondary hypothyroidism caused by exon skipping due to a homozygous donor splice site mutation in the TSHbeta-subunit gene." *J Clin Endocrinol Metab* **87**(1): 336-339.

Pollak, E. S., H. L. Hung, et al. (1996). "Functional characterization of the human factor VII 5'-flanking region." *J Biol Chem* **271**(3): 1738-1747.

Pomeranz Krummel, D. A., C. Oubridge, et al. (2009). "Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution." *Nature* **458**(7237): 475-480.

Proudfoot, N. J. (1989). "How RNA polymerase II terminates transcription in higher eukaryotes." *Trends Biochem Sci* **14**(3): 105-110.

Radcliffe, R. and Y. Nemerson (1975). "Activation and control of factor VII by activated factor X and thrombin. Isolation and characterization of a single chain form of factor VII." *J Biol Chem* **250**(2): 388-395.

Radcliffe, R. and Y. Nemerson (1976). "Mechanism of activation of bovine factor VII. Products of cleavage by factor Xa." *J Biol Chem* **251**(16): 4749-4802.

Ragni, M. V., J. H. Lewis, et al. (1981). "Factor VII deficiency." *Am J Hematol* **10**(1): 79-88.

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

Raker, V. A., G. Plessel, et al. (1996). "The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro." *EMBO J* **15**(9): 2256-2269.

Rao, L. V. and S. I. Rapaport (1988). "Activation of factor VII bound to tissue factor: a key early step in the tissue factor pathway of blood coagulation." *Proc Natl Acad Sci U S A* **85**(18): 6687-6691.

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

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

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

Robberson, B. L., G. J. Cote, et al. (1990). "Exon definition may facilitate splice site selection in RNAs with multiple exons." *Mol Cell Biol* **10**(1): 84-94.

Roca, X. and A. R. Krainer (2009). "Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA." *Nat Struct Mol Biol* **16**(2): 176-182.

Roca, X., A. J. Olson, et al. (2008). "Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics." *Genome Res* **18**(1): 77-87.

Romano, M., R. Marcucci, et al. (2002). "Regulation of 3' splice site selection in the 844ins68 polymorphism of the cystathionine Beta -synthase gene." *J Biol Chem* **277**(46): 43821-43829.

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

Rosen, E. D., J. C. Chan, et al. (1997). "Mice lacking factor VII develop normally but suffer fatal perinatal bleeding." *Nature* **390**(6657): 290-294.

Rossi, F., T. Forne, et al. (1996). "Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5' splice site-U1 snRNP interaction." *J Biol Chem* **271**(39): 23985-23991.

Rottingen, J. A., T. Enden, et al. (1995). "Binding of human factor VIIa to tissue factor induces cytosolic Ca2+ signals in J82 cells, transfected COS-1 cells, Madin-Darby canine kidney cells and in human endothelial cells induced to synthesize tissue factor." *J Biol Chem* **270**(9): 4650-4660.

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

Ruskin, B., P. D. Zamore, et al. (1988). "A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly." *Cell* **52**(2): 207-219.

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

Sahashi, K., A. Masuda, et al. (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 Res* **35**(18): 5995-6003.

Sanford, D. G., C. Kanagy, et al. (1991). "Structure of the propeptide of prothrombin containing the gamma-carboxylation recognition site determined by two-dimensional NMR spectroscopy." *Biochemistry* **30**(41): 9835-9841.

Sashital, D. G., G. Cornilescu, et al. (2004). "U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction." Nat Struct Mol Biol 11(12): 1237-1242.

Schaub, M., E. Myslinski, et al. (1997). "Staf, a promiscuous activator for enhanced transcription by RNA polymerases II and III." EMBO J 16(1): 173-181.

Schmid, F., E. Glaus, et al. (2011). "U1 snRNA-mediated gene therapeutic correction of splice defects caused by an exceptionally mild BBS mutation." Hum Mutat 32(7): 815-824.

Schneider, C., C. L. Will, et al. (2002). "Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions." Mol Cell Biol 22(10): 3219-3229.

Segref, A., I. W. Mattaj, et al. (2001). "The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export." RNA 7(3): 351-360.

Seligsohn, U., C. K. Kasper, et al. (1979). "Activated factor VII: presence in factor IX concentrates and persistence in the circulation after infusion." Blood 53(5): 828-837.

Seligsohn, U., B. Osterud, et al. (1979). "Activation of human factor VII in plasma and in purified systems: roles of activated factor IX, kallikrein, and activated factor XII." J Clin Invest 64(4): 1056-1065.

Seraphin, B., L. Kretzner, et al. (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." EMBO J 7(8): 2533-2538.

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

Shatkin, A. J. (1976). "Capping of eucaryotic mRNAs." Cell 9(4 PT 2): 645-653.

Sheth, N., X. Roca, et al. (2006). "Comprehensive splice-site analysis using comparative genomics." Nucleic Acids Res 34(14): 3955-3967.

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

Siliciano, P. G. and C. Guthrie (1988). "5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements." Genes Dev 2(10): 1258-1267.

Singh, R. and R. Reddy (1989). "Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA." Proc Natl Acad Sci U S A 86(21): 8280-8283.

Smith, D. J., C. C. Query, et al. (2007). "trans-splicing to spliceosomal U2 snRNA suggests disruption of branch site-U2 pairing during pre-mRNA splicing." Mol Cell 26(6): 883-890.

Smith, D. J., C. C. Query, et al. (2008). ""Nought may endure but mutability": spliceosome dynamics and the regulation of splicing." Mol Cell 30(6): 657-666.

Spurdle, A. B., F. J. Couch, et al. (2008). "Prediction and assessment of splicing alterations: implications for clinical testing." Hum Mutat 29(11): 1304-1313.

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

Stanek, D. and K. M. Neugebauer (2006). "The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze." Chromosoma 115(5): 343-354.

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

Stefanovic, B., W. Hackl, et al. (1995). "Assembly, nuclear import and function of U7 snRNPs studied by microinjection of synthetic U7 RNA into Xenopus oocytes." Nucleic Acids Res 23(16): 3141-3151.

Stevens, S. W., D. E. Ryan, et al. (2002). "Composition and functional characterization of the yeast spliceosomal penta-snRNP." Mol Cell 9(1): 31-44.

Sun, J. S. and J. L. Manley (1995). "A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing." Genes Dev 9(7): 843-854.

Sunnerhagen, M. S., E. Persson, et al. (1993). "The effect of aspartate hydroxylation on calcium binding to epidermal growth factor-like modules in coagulation factors IX and X." *J Biol Chem* **268**(31): 23339-23344.

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

Susani, L., A. Pangrazio, et al. (2004). "TCIRG1-dependent recessive osteopetrosis: mutation analysis, functional identification of the splicing defects, and in vitro rescue by U1 snRNA." *Hum Mutat* **24**(3): 225-235.

Tanner, G., E. Glaus, et al. (2009). "Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA." *Hum Mutat* **30**(2): 255-263.

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

Tarn, W. Y. and J. A. Steitz (1994). "SR proteins can compensate for the loss of U1 snRNP functions in vitro." *Genes Dev* **8**(22): 2704-2717.

Tavassoli, K., A. Eigel, et al. (1998). "Identification of four novel mutations in the factor VIII gene: three missense mutations (E1875G, G2088S, I2185T) and a 2-bp deletion (1780delTC)." *Hum Mutat* **Suppl 1**: S260-262.

Tavassoli, K., A. Eigel, et al. (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." *Hum Mutat* **12**(5): 301-303.

Tazi, J., N. Bakkour, et al. (2009). "Alternative splicing and disease." *Biochim Biophys Acta* **1792**(1): 14-26.

Tazi, J., U. Kornstadt, et al. (1993). "Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing." *Nature* **363**(6426): 283-286.

Tournier, I., M. Vezain, et al. (2008). "A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects." *Hum Mutat* **29**(12): 1412-1424.

Tripathi, V., J. D. Ellis, et al. (2010). "The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation." *Mol Cell* **39**(6): 925-938.

Triplett, D. A., J. T. Brandt, et al. (1985). "Hereditary factor VII deficiency: heterogeneity defined by combined functional and immunochemical analysis." *Blood* **66**(6): 1284-1287.

Tsuji-Wakisaka, K., M. Akao, et al. (2011). "Identification and functional characterization of KCNQ1 mutations around the exon 7-intron 7 junction affecting the splicing process." *Biochim Biophys Acta* **1812**(11): 1452-1459.

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

Uchikawa, H., K. Fujii, et al. (2007). "U7 snRNA-mediated correction of aberrant splicing caused by activation of cryptic splice sites." *J Hum Genet* **52**(11): 891-897.

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

Valadkhan, S. (2005). "snRNAs as the catalysts of pre-mRNA splicing." *Curr Opin Chem Biol* **9**(6): 603-608.

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

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

Valadkhan, S., A. Mohammadi, et al. (2007). "Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing." *RNA* **13**(12): 2300-2311.

van 't Hooft, F. M., A. Silveira, et al. (1999). "Two common functional polymorphisms in the promoter region of the coagulation factor VII gene determining plasma factor VII activity and mass concentration." *Blood* **93**(10): 3432-3441.

Van Arsdell, S. W. and A. M. Weiner (1984). "Human genes for U2 small nuclear RNA are tandemly repeated." *Mol Cell Biol* **4**(3): 492-499.

Vankeerberghen, A., L. Wei, et al. (1998). "Characterization of 19 disease-associated missense mutations in the regulatory domain of the cystic fibrosis transmembrane conductance regulator." *Hum Mol Genet* **7**(11): 1761-1769.

Vibe-Pedersen, K., A. R. Kornblihtt, et al. (1984). "Expression of a human alpha-globin/fibronectin gene hybrid generates two mRNAs by alternative splicing." *EMBO J* **3**(11): 2511-2516.

Vickers, T. A., M. Sabripour, et al. (2011). "U1 adaptors result in reduction of multiple pre-mRNA species principally by sequestering U1snRNP." *Nucleic Acids Res* **39**(10): e71.

Vidal, V. P., L. Verdone, et al. (1999). "Characterization of U6 snRNA-protein interactions." *RNA* **5**(11): 1470-1481.

Wachtel, C. and J. L. Manley (2009). "Splicing of mRNA precursors: the role of RNAs and proteins in catalysis." *Mol Biosyst* **5**(4): 311-316.

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

Wang, E. T., R. Sandberg, et al. (2008). "Alternative isoform regulation in human tissue transcriptomes." *Nature* **456**(7221): 470-476.

Wang, G. S. and T. A. Cooper (2007). "Splicing in disease: disruption of the splicing code and the decoding machinery." *Nat Rev Genet* **8**(10): 749-761.

Wang, H. Y., W. Lin, et al. (1998). "SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells." *J Cell Biol* **140**(4): 737-750.

Warshawsky, I., J. Herz, et al. (1996). "The low density lipoprotein receptor-related protein can function independently from heparan sulfate proteoglycans in tissue factor pathway inhibitor endocytosis." *J Biol Chem* **271**(42): 25873-25879.

Will, C. L., S. Rumpler, et al. (1996). "In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes." *Nucleic Acids Res* **24**(23): 4614-4623.

Wilton, S. D. and S. Fletcher (2011). "RNA splicing manipulation: strategies to modify gene expression for a variety of therapeutic outcomes." *Curr Gene Ther* **11**(4): 259-275.

Wion, K. L., D. Kelly, et al. (1985). "Distribution of factor VIII mRNA and antigen in human liver and other tissues." *Nature* **317**(6039): 726-729.

Wood, M., H. Yin, et al. (2007). "Modulating the expression of disease genes with RNA-based therapy." *PLoS Genet* **3**(6): e109.

Wu, J. Y. and T. Maniatis (1993). "Specific interactions between proteins implicated in splice site selection and regulated alternative splicing." *Cell* **75**(6): 1061-1070.

Wu, S., C. M. Romfo, et al. (1999). "Functional recognition of the 3' splice site AG by the splicing factor U2AF35." *Nature* **402**(6763): 832-835.

Yi, R., B. P. Doehle, et al. (2005). "Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs." *RNA* **11**(2): 220-226.

Yong, J., L. Pellizzoni, et al. (2002). "Sequence-specific interaction of U1 snRNA with the SMN complex." *EMBO J* **21**(5): 1188-1196.

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

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

Zdziarska, J., A. Undas, et al. (2009). "Severe bleeding and miscarriages in a hypofibrinogenemic woman heterozygous for the gamma Ala82Gly mutation." *Blood Coagul Fibrinolysis* **20(5):** 374-376.

Zhang, Y., Y. Deng, et al. (1994). "Tissue factor controls the balance of angiogenic and antiangiogenic properties of tumor cells in mice." *J Clin Invest* **94(3):** 1320-1327.

Zhuang, Y. and A. M. Weiner (1986). "A compensatory base change in U1 snRNA suppresses a 5' splice site mutation." *Cell* **46(6):** 827-835.