Role and Mechanism of Action of Exonic Splicing Regulatory Sequences

Thesis

UNRESTRICTED

Role and Mechanism of Action of Exonic Splicing Regulatory Sequences

Elisa Goina

A Thesis Submitted in Fulfilment of the Requirements of the Open University, (UK) for the Degree of Doctor of Philosophy

International Centre for Genetic Engineering and Biotechnology, ICGEB

Trieste, Italy

Director of Studies: Dr Franco Pagani, M.D.

External Supervisor: Prof. Angus I. Lamond PhD, FRSE

Second Supervisor: Prof. Francisco E. Baralle, M.D., Ph.D.

Submission date: 25 September 2008
Date of award: 31 December 2008

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGMENTS

This work was carried out in the Molecular Pathology / Human Molecular Genetics group at the International Centre for Genetic Engineering and Biotechnology (ICGEB) in Trieste, under the supervision of Dr. Franco Pagani.

Firstly I would like to thank my boss Franco Pagani for allowing me to work in his group, for his support in research work during these years and for his professionalism and patience.

A very special "thanks" goes to Prof. Baralle for his help and constant presence in solving all the problems encountered during my PhD period.

I am also very grateful to Prof. Lamond for his valuable discussions and advice during our meetings and the writing of my PhD thesis.

Thanks to Prof. Chabot who carefully read and corrected my thesis.

I wish to thank the marvelous Human Molecular Genetics people, Signora B., Gably, Timba, Eu and Ify, for their great friendship and for the splendid atmosphere that they create in the lab!!!

I want also to thank some other people that spent the first period of my PhD with me in the Molecular Pathology lab. Thanks to Cristiana for always being present, to Maka and Michi for still being friends and to Natasha for her strong support.

Last, but not least, to my family and all my friends...really thank you!!!

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ABBREVIATIONS

The standard abbreviations used in this dissertation follow IUPAC rules. All the abbreviations are defined also in the text when they are introduced for the first time. The abbreviations mentioned only once are not included in this list.

bp	Base pairs
ATM	Ataxia-telangiectasia
BRCA1	BReast CAncer 1 gene
CERES	Composite Exonic Regulatory Elements of Splicing gene
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
cDNA	copy DNA
DAZAP1	Deleted in AZoospermia Associated Protein 1
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphate (A, C, G and T)
DTT	Dithiothreitol
EDTA	Ethylenediamine tetra-acetic acid
ESE	Exonic Splicing Enhancer
ESS	Exonic Splicing Silencer
hnRNP	Heterogenous nuclear ribonucleoprotein
Hep3B	Human hepatocarcinoma cells
IPTG	Isopropyl-d-thiogalactopyranoside
ISE	Intronic Splicing Enhancer
ISS	Intronic Splicing Silencer
kb	Kilobase

kDa	Kilodalton
nt	Nucleotides
N	Nucleotide (A or C or G or T)
NF1	Neurofibromatosis type I gene
PBS	Phosphate buffer saline
PCR	Polymerase Chain Reaction
Pu	Purine (G or A)
Py	Pyrimidine (T or C)
RNA	Ribonucleic acid
RNA Pol II	RNA polymerase II
RRM	RNA Recognition Motif
SDS	N-lauroylsarcosine sodium salt
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
SMN	Survival of Motor Neuron
snRNA	Small nuclear RNA
snRNP	small nuclear RiboNucleoProtein
SR	Serine-arginine-rich protein
ss	splice site
TBE	Tris-borate-EDTA
U2AF	U2 snRNP Auxiliary Factor

ABSTRACT

Exonic mutations can result in altered protein function by affecting exon recognition during splicing. To understand this mechanism I have extensively evaluated the exonic regulatory elements affected by the disease G to T mutation at position +6 of BRCA1 exon 18. This substitution induces the exclusion of the exon and it has been suggested that it disrupts an ASF/SF2-dependent enhancer. Using a pulldown assay with an internal standard, I show that WT and T6 sequences bind ASF/SF2 with similar efficiency, which is significantly lower compared to the binding to a typical enhancer derived from the fibronectin EDA exon. Consistent with the absence of an ASF/SF2 enhancing effect, siRNA depletion of ASF/SF2 did not induce exon WT exclusion indicating that ASF/SF2 is not essential for BRCA1 exon 18 splicing. However, depletion or overexpression of ASF/SF2 resulted in a decrease or increase of T6 exon 18 inclusion, respectively. By a series of exonic deletions in the WT and T6 BRCA1 minigenes, an enhancer sequence was identified in position 23-32 which mediates the observed T6-specific, ASF/SF2-dependent splicing activation. Furthermore, extensive mutations analysis indicated that the T6 mutant creates a sequence with a predominantly inhibitory function. Double site-directed mutations showed that the point mutations affecting the "TAG" sequence, between +6 and +8 positions, completely restore normal splicing. Indeed, RNA protein interaction and siRNA experiments showed that the skipping of T6 BRCA1 exon 18 is due to the creation of a silencer element. This sequence specifically binds to the proteins hnRNP A1/A2 and to DAZAP1. By siRNA experiments, DAZAP1 and hnRNP A1/A2 are shown to be involved redundantly in the regulation of the defective BRCA1 exon 18. The present results support a gain-of-function model for the

BRCA1 T6 exon 18 and indicate that the binding of the hnRNP A1/A2 and DAZAP1 is the primary determinant of exon skipping.

1 INTRODUCTION

The accuracy of intron removal and exon junction during pre-mRNA splicing is determined principally by the recognition of the splicing consensus sequences at the 5' and 3' exon splice sites. More discrete elements such as exonic or intronic splicing enhancers and silencers are also required for regulating pre-mRNA splicing (Cartegni, Chew et al. 2002). Several lines of evidence indicated that nucleotide modifications can be deleterious for normal pre-mRNA splicing by disrupting these regulatory sequences or by creating new cryptic ones. As a result, each nucleotide variant has to be considered as a potential candidate for promoting splicing alterations even if located in intronic sequences or classified as a translationally-silent mutation (Pagani and Baralle 2004).

Over recent years numerous disease associated point mutations have been associated with a defective splicing process making the study of the connection between aberrant splicing and occurrence of disease a central issue in the medical research field (Faustino and Cooper 2003; Garcia-Blanco, Baraniak et al. 2004; Pagani and Baralle 2004). The study of splicing mutations mechanisms increased our general knowledge of different splicing components and improved the understanding of the various network of interactions between cis-acting elements and trans-acting factors. However, the determinants of splicing regulation are still poorly understood as emphasized by many examples of splicing systems that are not easy to model due to their enormous complexity (Singh 2007).

Therefore, one of the key issues raised in molecular diagnosis is the accurate interpretation of the biological consequences of nucleotide variants and their related effect on splicing process. A correct classification of splicing mutations is essential

for understanding structure–function relationships at the protein level, for considering the individual phenotypic risk in familiar diseases and for developing new therapeutic strategies (Cartegni, Chew et al. 2002).

1.1 Pre-mRNA splicing reaction.

1.1.1 Overview of pre-mRNA splicing processing.

Splicing reaction is an essential step in the post-transcriptional regulation of gene expression in eukaryotes. This process is structurally and functionally associated with the nucleus and occurs in a complex termed the spliceosome (Maniatis and Reed 2002). Most eukaryotic genes are transcribed as long precursor molecules containing the coding portions of the transcript (exons) separated by multiple non-coding intervening sequences (introns). In order to generate mature transcripts introns must be precisely removed from the pre-mRNA and exons have to be joined together (Staley and Guthrie 1998).

The first system used to study mRNA splicing and the structure of the pre-mRNA molecule was the late stage of adenovirus infection. The presence of introns was in fact discovered in the non-coding regions of the adenovirus (Berget, Moore et al. 1977; Chow, Gelinas et al. 1977). Subsequent to the characterization of RNA splicing process in viruses, the presence of introns was also reported as a general feature of eukaryotic genes. Jeffreys and Flavell in fact described the presence of a "large insert" in the coding sequence of rabbit beta-globin gene in 1977. A year after, Chambon and colleagues reported that the chicken ovalbumin gene contains seven coding sequences (exons) separated by six intervening sequences (introns) showing the exact positions where the coding sequences for ovalbumin mRNA are interrupted inside the genome (Breathnach, Benoist et al. 1978). Moreover they found that the

sequences at exon-intron boundaries carry common features, probably with the function of unique excision-ligation common points to all boundaries (Breathnach, Benoist et al. 1978). Interestingly, these consensus sequences were also present in vertebrate, plant and yeast suggesting that the splicing process is evolutionarily conserved (Padgett, Grabowski et al. 1986).

In mammalian cells, the spliceosome machinery is present within nuclear boundaries (Maniatis and Reed 2002) showing a dynamic but precise nuclear-localization pattern that is usually termed "speckled pattern". Nuclear speckles function as storage and modification compartments that can provide splicing factors to active transcription sites. Several studies showed that splicing factors are recruited from speckles to sites of transcription and when transcription is inhibited splicing factors accumulate in enlarged, rounded speckles (Lamond and Spector 2003).

On the other hand, a clear functional role for the spliceosome in extra-nuclear domains has not been clearly described. However recent studies have reported splicing activity in isolated rat neuronal dendrites (Glanzer, Miyashiro et al. 2005) and in human anucleate platelets (Denis, Tolley et al. 2005), suggesting the presence of a "nonnuclear" splicing in these specialized cells.

The complete formation of mature RNA (mRNA) requires not only the splicing reaction but also other processes and modifications. The mRNA biogenesis in fact comprises a complex network of dynamic and functional interactions (Fig. 1.1) between the elongating RNA Pol II and different processing machines responsible for capping, polyadenylation, nuclear surveillance and mRNAs export (Bentley 2002; Bentley 2005). Distinct and highly complex cellular machines can carry out each of these steps in the gene-expression process even if growing evidence suggests the existence of "gene-expression factories" in which these individual processes are functionally connected. Indeed, several studies have delineated functional

relationships among these processes leading to the idea that they occur co-transcriptionally (Bentley 2002; Bentley 2005).

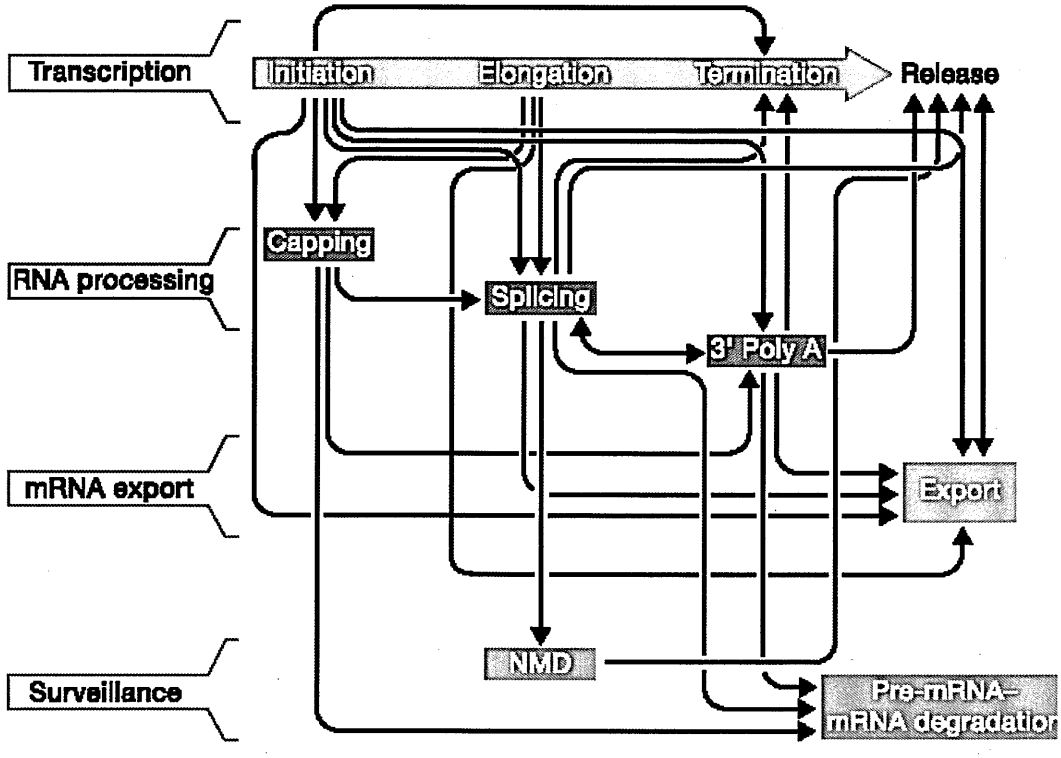

Figure 1.1: A complex network of coupled interactions in gene expression.
The major processes of gene expression are indicated on the left and each step of the transcription is shown at the top of the picture by the red arrow. "Release" indicates the release of the mature mRNA from the site of transcription. The splicing step is only part of the entire RNA processing that includes also the capping and the polyadenilation, as shown in blue. It is important to notice that splicing is also connected with other gene expression steps like transcription, mRNA export, nonsense mediated decay (NMD) and RNA degradation as indicated by the black arrows. The black arrows show in fact physical and/or functional coupling between different steps in gene expression. Each arrow is documented by published studies. Figure adapted from Maniatis and Reed (Maniatis and Reed 2002).

1.1.2 The chemistry of splicing reaction.

Splicing of pre-mRNA requires the interaction and cooperation of several small RNAs and a large number of proteins in order to remove introns and join exons. However, splicing from a chemical point of view is quite simple: it occurs in two transesterification reactions (Lamond 1993; Moore and Sharp 1993).

Conserved motifs in the nucleotide sequences, at the intron-exon boundaries, act as essential splicing signals in the chemical part of the splicing reaction. A "GU" at the exon-intron junction defines the 5'splice site (5'ss) and an "AG" at the other intron-exon junction together with the polyprimidine tract and the A nucleotide at the branch point identify the 3' splice site (3'ss).

During the first step of the splicing reaction, the 2'-hydroxyl group of an A residue at the branch site attacks the phosphate at the 5'ss (Fig. 1.2). This leads to cleavage of the 5' exon 1 from the intron and the concerted ligation of the intron 5' end to the branch-point 2'-hydroxyl. This step produces two reaction intermediates, a detached 5' exon 1 and an intron/3'-exon 2 fragment in a lariat configuration containing a branched A nucleotide at the branch point (Fig. 1.2).

The second transesterification step consists of the attack of the 3'-hydroxyl of the detached exon1 on the phosphate at the 3' end of the intron. This step results in the ligation via phosphodiester bond of the two exons and the release of the intron, still in the form of a lariat (Fig. 1.2). The lariat is then "debranched" to give a linear excised intron, which is usually rapidly degraded.

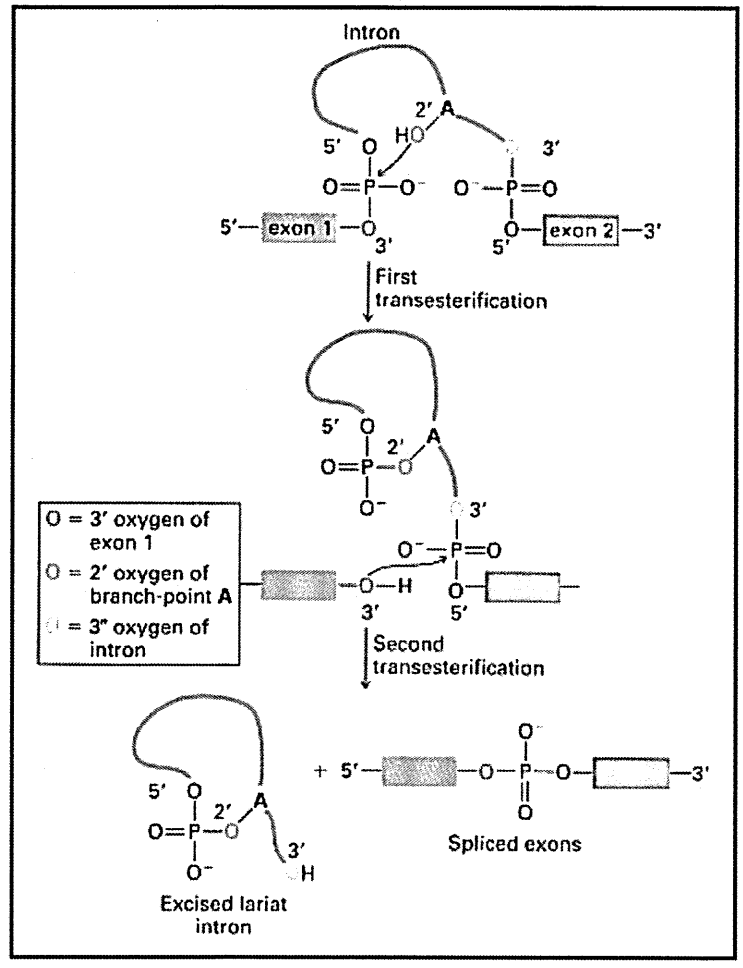

Figure 1.2: Schematic representation of the RNA splicing reaction.

Exons are represented by the pink and red rectangles. The specific adenine nucleotide of the branch point is indicated with an "A". Splicing takes place in two catalytic steps that involves two consecutive trans-esterification reactions.

In the first step of splicing, an adenine residue of the branch point in the intron sequence carries out a nucleophilic attacks the 5' splice site of the exon1 and cuts the sugar-phosphate backbone of the RNA at this point. This reaction generates two splicing intermediates: the free exon 1 and the lariat-exon 2.

In the second reaction, the ester bond between the 5' phosphorous of exon 2 and the 3' oxygen (light blue) of the intron is exchanged for an ester bond with the 3' oxygen of exon 1, releasing the intron as a lariat structure and joining the two exons. Figure adapted from Lodish (Lodish 2000).

1.1.3 Pre-mRNA Splicing is catalyzed by the Spliceosome.

The pre-mRNA splicing reaction catalysis requires the formation of a highly dynamic macromolecular complex called spliceosome composed of several small nuclear ribonucleoproteins (snRNPs) and a large number non-snRNP splicing factors (Rappsilber, Ryder et al. 2002; Jurica and Moore 2003). Several studies, in a variety of systems, have revealed that the essential components of the snRNPs include five uridine-rich RNAs that are classified as snRNAs-U1, U2, U4, U5 and U6. Each snRNP consists of a small stable RNA bound by a specific set of proteins, plus numerous other less stably associated splicing factors (Will and Lührmann 2001). In addition to the five snRNPs, the spliceosome involves many non-snRNP factors but their exact nature and number is not known (Rappsilber, Ryder et al. 2002; Jurica and Moore 2003).

Spliceosome formation consists of a precise and ordered assembly/disassembly of the different snRNP particles and non-snRNP splicing factors onto the pre-mRNA substrate prior splicing catalysis. Several kinds of interactions like RNA-RNA base pairing, RNA-protein and protein-protein binding, together with different conformational rearrangements are carried out for the production of a mature RNA. In mammals, distinct intermediates in spliceosome assembly pathway have been detected (Fig. 1.3). The E, A and B complexes contain unspliced pre-mRNA whereas the C complex has the products of catalytic step I of the splicing reaction (exon 1 and lariat-exon 2). Two complexes, one containing the spliced exons (D complex) and the other containing the excised lariat intron (i complex), are the products of catalytic step II of the splicing reaction (Reed 2000). Although the earliest assembly step (E complex) is energy independent, subsequent spliceosomal steps require NTP hydrolysis (Staley and Guthrie 1998).

The E (early) or commitment complex assembly starts with the recognition of the 5'ss driven by the U1 snRNP in an ATP-independent manner (Fig.1.3a). This interaction is mediated by base pairing of the U1 snRNA with the 5'ss, as well as by protein-protein and protein-pre-mRNA interaction involving U1-70K and U1-C proteins (Will, Rumpler et al. 1996). The other important event that occurs during the E complex formation is the recognition of the 3'ss. In higher eukaryotes, the AG at intron/exon junction together with the adjacent polypyrimydine tract is identified through interactions with the dimeric U2 Auxiliary Factor (U2AF). The subunit U2AF65 recognizes the polypyrimidine tract (Valcarcel, Gaur et al. 1996) and the AG dinucleotide at the 3'ss interacts with the U2AF35 subunit (Wu, Romfo et al. 1999). The branch point, usually located 20-40 nucleotides (nt) upstream the 3'ss within the intron, is recognized by the branch point binding protein (BBP/SF1) (Berglund, Chua et al. 1997).

The formation of the A complex is characterized by the recruitment of the U2 snRNP that replaces the BBP/SF1 factor to the conserved branch site sequence (Fig.1.3b). This U2–branch site ATP-dependent binding is mediated by U2 snRNA base pairing and by U2 snRNP protein–pre-mRNA interactions through SF3b and SF3a subunits (Gozani, Feld et al. 1996). However, U2 snRNP was also identified as a component of a purified, functional E complex (Hong, Bennett et al. 1997). The U2 snRNP seems to bind loosely to the pre-mRNA in the E complex via the integral U2-snRNP-associated protein SF3b, and then through an ATP-dependent process leads to stable binding to the branch point replacing SF1 (Das, Zhou et al. 2000).

Then the ATP-dependent addition of U4/U6•U5 snRNPs, in which the U4 and U6 snRNAs are base paired, to the spliceosomal complex characterizes the transition from the A to B complex (Fig.1.3c). However recent studies reported that the tri-snRNPs are able to interact with the 5'ss and the upstream 5' exon at earlier step of

spliceosome assembly (Maroney, Romfo et al. 2000). Although B complex contains all of the snRNPs components required for splicing, it lacks a catalytic center. In order to activate the spliceosome, the complex B undergoes marked RNA-RNA rearrangements that involve the displacement of U1 snRNP at the 5'ss by U6 snRNP, the disruption of the U4/U6 base pairing interaction and the formation of an intricate network of interactions among the U6, U2 snRNAs and the pre-mRNA (Fig.1.3d). The U2 and U6 snRNAs are extensively base-paired with one another and with the branch site sequence and 5'ss respectively providing a structural basis for juxtaposing the branch site and 5'ss for the first catalytic step (Reed 2000; Boehringer, Makarov et al. 2004). All these rearrangements contribute to the fidelity of 5'ss recognition and support the contact with the branch point generating the activated B complex (B* complex) (Turner, Norman et al. 2004). The formation of B* complex promotes the first catalytic step of splicing in order to generate the free 5' exon and the lariat-3' exon intermediates. This step is followed by the formation of the C complex, in which the second catalytic step of splicing reaction takes place (Fig.1.3e). The U5 snRNA, together with the U2 and U6 snRNAs, is involved in aligning the exons for the second catalytic step through a highly conserved stem loop (O'Keefe, Norman et al. 1996). In addition the presence of a large highly conserved component of the U5 snRNP associated protein, the Prp8 factor, seems to stabilise these interactions. Prp8 had previously been shown to crosslink to both the 5'ss and 3'ss, as well as to the exons flanking these two splice sites (Umen and Guthrie 1995). In addition, Collins and Guthrie have reported that a specific region of this factor affects a tertiary interaction between both the 5'ss and 3'ss and U6 snRNA (Collins and Guthrie 1999).

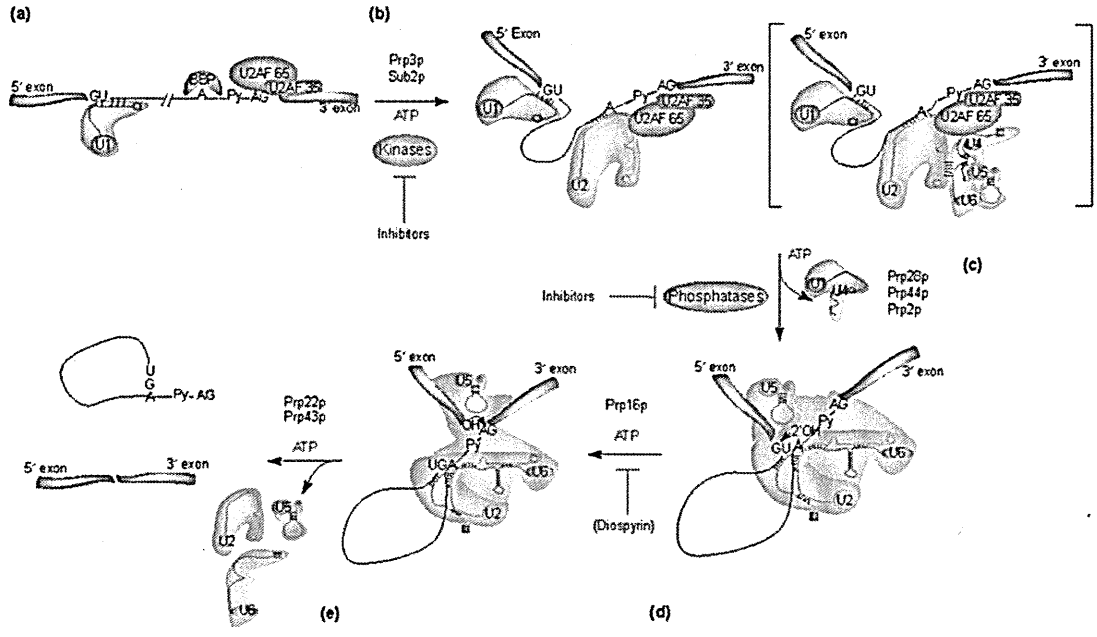

Figure 1.3: Spliceosome assembly.

The spliceosome assembles onto the pre-mRNA in a stepwise manner.

(a) The E (early) or commitment complex assembly starts with the recognition of the 5'ss the branch point and 3'ss mediated by the U1 snRNP, BBP/SF1 and U2AF factors, respectively. (b) The A complex is then generated by the recruitment of the U2 snRNP to the branch point adenosine in an ATP-dependent manner. Inhibitors of kinases and helicases can block this step. (c) The U4/U6•U5 tri-snRNP a full spliceosome is required for the transition to the complex B. This complex is shown in square brackets because the interactions of U4/U6•U5 with specific regions of the pre-mRNA are not well defined. (d) Dissociation of U4 and U1 from the later complex leads to spliceosome activation, B* complex. This step requires dephosphorylation of proteins that can be inhibited by phosphatase inhibitors. (e) Rearrangement involving U2, U5 and U6 snRNPs allows the formation of the C complex. The final step is characterized by the release of splicing products and recycling of snRNPs. Figure adapted from Tazi, Durand et al. (Tazi, Durand et al. 2005).

It is still debatable whether RNA, proteins or both, ultimately catalyzes splicing. Although a large number of proteins are known to be part of spliceosome, there is now growing evidence that the catalytic component of the spliceosome may be RNA based (Valadkhan and Manley 2001; Valadkhan, Mohammadi et al. 2007). Firstly, *in vitro*-transcribed U2 and U6 RNA fragments in the absence of any protein components are able to form intermolecular base pairing, similar to those obtained in the spliceosome. Secondly, the addition of a small branch-site like RNA oligonucleotides to this U2/U6 RNA complex was able to trigger, in the complete absence of proteins, a nucleophilic attack of the branchpoint adenosine on a distinct nucleotide within U6. The final product is a lariat-like structure between U6 and the branchpoint oligonucleotides (Valadkhan and Manley 2001; Valadkhan, Mohammadi et al. 2007). Nevertheless, *in vivo*, protein components of the spliceosome play an essential role in the assembly and stabilization of a catalytically active spliceosome.

1.2 Recognition of the exon: the cis-acting elements.

Exon recognition is principally achieved through the splice site, small consensus sequences located at the intron/exon junctions (Fig. 1.4). The canonical consensus sequence called 5'splice site marks the exon/intron junction at the 5' end of the intron. In mammals only the first two nucleotides GU of the intron are conserved although they are not the only important bases (Aebi, Hornig et al. 1987). At the other intron/exon boundary there is the 3'splice site region that comprises three conserved elements: the branch point, the polypyrimidine tract and the conserved AG dinucleotide at the extreme 3' end of the intron (Shapiro and Senapathy 1987).

The definition of splice site sequence as strong or weak depends on the higher or lower level of similarity to the standard consensus sequences that characterize the exon/intron boundaries. Over the last years, several computational methodologies have been developed to predict the position of 5'ss and 3'ss and the strength of the identified motif. These methods are based on nucleotide frequencies matrices, neural networks and interdependencies between nucleotides at different positions of the consensus sequences (Shapiro and Senapathy 1987).

It is important to notice that the sequences that match the 5'ss and 3'ss are very common within intron sequences and sometimes even stronger than the real splice sites. These intronic splice sites normally define pseudo-exons that are not included in mature mRNAs (Sun and Chasin 2000). The fact that the splicing machinery is able to recognize the correct splice sites, even if they are weak, distant from each other and surrounded by several pseudo 5'ss and 3'ss, means that these consensus elements are by no means sufficient to define exon-intron junctions. In fact, the accurate recognition of splice sites *in vivo* is the result of a combinatorial regulatory mechanism that comprises additional cis-acting regulatory elements (Fig. 1.5).

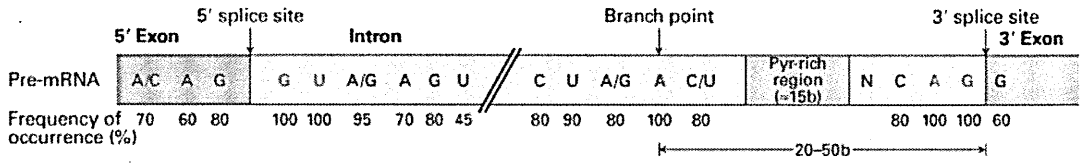

Figure 1.4: Schematic representation of exon-intron boundaries.

The two exons are represented by the red and pink boxes respectively. Between them are reported the consensus sequences present within an intron. The arrows indicate the position of the 5' (GU) and 3' (AG) splice site and the branch point (A). The polypyrimidine tract, rich in pyrimidines, is highlighted by a blue rectangle. There are several conserved nucleotides near the sequences surrounding the intron-exon junctions that act as essential splicing signals. The frequency of each nucleotide in an alignment of conserved sequences from 1,683 human introns is indicated below. N means any base. To note only the universally conserved nucleotides are the dinucleotide cores of the 5' and 3' splice together with the branch point (A) showed 100% of frecquency of occurrence. Figure adapted from Lodish (Lodish 2000).

1.2.1 The 5' splice site (5'ss).

The 5'ss motif in higher eukaryotes consists of nine partially conserved nucleotides, MAG/GURAGU (M indicates A or C, R indicates purines and the slash the exon-intron boundary) at the exon-intron junction, spanning from position -3 to +6 (Fig. 1.4). The underlined GU dinucleotide is universally conserved and mutating one of these two nucleotides completely abolishes splicing (Langford, Klinz et al. 1984). Recognition of the 5'ss involves a nearly perfect Watson-Crick base pairing with the U1 snRNA (Horowitz and Krainer 1994) and guides the early assembly of the spliceosome machinery upon the intron. There are strong evidences for the critical contribution of the base pairing between 5' consensus motif and U1 snRNA to selection of the 5'ss (Zhuang and Weiner 1986; Seraphin, Kretzner et al. 1988;

Siliciano and Guthrie 1988). Studies of alternative splicing have shown that mutations that improve the match of weak splice sites to the consensus can lead to the constitutive recognition of alternatively skipped exon (Huh and Hynes 1993; Muro, Iaconcig et al. 1998).

1.2.2 The 3' splice site (3'ss).

The intronic element that identifies the 3'ss usually appears several thousand bases downstream of the 5'ss. The 3' border of the intron is defined by three different conserved elements (Fig. 1.4): the branch point, the polypyrimidine tract and the terminal AG dinucleotide (Reed 1989). Together these elements make up a loosely defined 3'ss region, which may be extend up to 100 nt along the intron.

The branch site: The branch point in human introns is highly degenerated, and the general consensus sequence, apart from the universally conserved "A", is defined by the YNYURAC motif (R-purine,Y-pyrimidine). Most branch points have been mapped within 18-40 nt of the 3'ss (Reed and Maniatis 1988). However, in the case of rat α-tropomyosin gene intron 2, the branch point is located 172 nt upstream from the 3'ss and nevertheless is essential for the regulation of alternative splicing (Smith and Nadal-Ginard 1989). During the early spliceosome assembly, the SF1 factor recognises the branch site. Although the sequence specificity of SF1 is low, it can discriminate between sequences with general consensus branch site and sequences missing the highly conserved adenosine (Berglund, Chua et al. 1997). Subsequently the recognition of the branch site involves a base pairing with the U2 snRNA, in order to form the spliceosome A complex.

The polypyrimidine tract: The polypyrimidine track 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. This sequence is more pronounced in mammals than in yeast, where the nucleotides composition and the length can vary significantly. The 65-kDa subunit of U2AF binds the polypyrimidine tract during the formation of the early spliceosome complex (Zamore, Patton et al. 1992; Kielkopf, Lucke et al. 2004). Due to its position, the polypyrimidine tract is able to affect the recognition of the branch point sequences (Reed 1989). It has been shown that progressive deletion of the polypyrimidine tract abolishes spliceosome assembly and lariat formation while elongating its length can improve the splicing efficiency (Roscigno, Weiner et al. 1993).

The canonical sequence that characterizes the polypyrimidine tract is composed by a poly-uridine tract, whereas, a poly-cytidine tract is not functional (Roscigno, Weiner et al. 1993). The proximity of the polypyrimidine tract to the 3'ss is important when the pyrimidine length is limiting. Shortening the number of continuous uridines implies the localization of these uridines immediately adjacent to the 3'ss AG. Vice versa, a polypyrimidine tract containing high number of uridines is a competitive pyrimidine stretch regardless of the distance between the branch point and polypyrimidine tract itself (Coolidge, Seely et al. 1997). A functional polypyrimidine tract does not absolutely require continuous uridines. In constitutively included exon 3 of the human Apolipoprotein-AII (Apo AII) the "polypyrimidine tract" is composed of a $(GT)_{16}$ dinucleotide repeats (Shelley, Sharpe et al. 1985).

The terminal AG dinucleotide: The terminal AG dinucleotide defines the 3' border of the intron, just downstream the polypyrimidine. This site is characterised by the short YAG/G sequence (Y denotes pyrimidines; the slash indicates the intron-exon boundary and the underlined nucleotides are conserved). Even if it is essential for the second step of splicing catalyses, no base-pairing interactions with snRNAs are

involved in recognizing this sequence. During the early step of spliceosome assembly is recognised by the U2AF35 subunit (Wu, Romfo et al. 1999).

1.2.3 Minor introns class.

Metazoan genes have been found to contain a rare class of introns that carry highly conserved splice sites that differ from the canonical consensus sequences and are excised by distinct spliceosome (Hall and Padgett 1994). This minor class of introns seems to be most prevalent in human and vertebrates, representing 0.15%-0.34% of all introns (Burge, Padgett et al. 1998). Because these new introns have an AT and AC ends, which deviates from the nearly invariant GT-AG, they are termed AT-AC introns. However, more extensive genomic database searches revealed that AT-AC termini are not the only defining feature of minor-class introns. In fact, most minor-class introns carry the canonical GT-AG ends, and, very rarely, major-class introns have AT-AC termini (Sharp and Burge 1997). Several studies showed that a minor intron is characterized by a longer and more tightly constrained consensus sequence at the 5'ss and at the branch site, as well as by the lack of a polypyrimidine tract upstream of the 3'ss (Sharp and Burge 1997; Burge, Padgett et al. 1998).

1.3 Additional cis-acting elements.

In order to generate a correct and mature mRNAs, the exons have to be precisely identified among the intronic sequences by the spliceosome. Paradoxically the mammalian junctions that define an exon are weakly conserved and more degenerate with respect to yeast canonical cis-elements. These elements are necessary but are by no means sufficient to define exon/intron junctions. Early experiments in 1987 have shown that internal exonic sequences far from the 5' and 3'ss were essential for exon recognition (Mardon, Sebastio et al. 1987). Thus, in addition to the conventional splicing signals located at exon-intron boundaries, other cis-elements in pre-mRNA that regulate splicing have been extensively characterized (Cartegni, Chew et al. 2002). These elements localized both in exon and intron sequences, act by stimulating or inhibiting the use of specific splice sites (Fig. 1.5). Depending on their location and their function, these elements are referred as exonic splicing enhancers (ESE) and silencers (ESS) or intronic splicing enhancers (ISE) or silencers (ISS) (Cartegni, Chew et al. 2002). Several studies have revealed a non random distribution of these splicing regulatory sequences. There is a higher density of enhancers in real exons than in pseudoexons and introns, and vice versa for the silencers sequences (Sironi, Menozzi et al. 2004; Zhang and Chasin 2004). Although these additional elements are often conserved between species, highly degenerate sequence motifs characterized these elements, making difficult their identification. It has to be noticed that a cis-acting element can have overlapping functions as reported for the cystic fibrosis transmembrane conductance regulator gene (CFTR) exons 9 and 12 (Pagani, Buratti et al. 2003; Pagani, Stuani et al. 2003). In this case, it may be more appropriate to talk about composite exonic regulatory elements of splicing (CERES).

Several features such as nearby sequence composition, the intronic/exonic position (Wang and Burge 2008) and the secondary structure can deeply influence the nature of these elements (Buratti and Baralle 2004). For instance, a cis-acting element could be strictly connected with its context location and when moved in a heterologous context the element fails to regulate splicing. In other cases the location of a regulatory element relative to the pre-mRNA deeply influences the nature of the element itself (Wang and Burge 2008). For example, G triplets, through the hnRNP H (heterogenous nuclear ribonucleoprotein H) binding, commonly enhance splicing when located downstream of the 5'ss within intron sequence (McCullough and Berget 1997; Chou, Rooke et al. 1999; Caputi and Zahler 2001) but they behave as splicing silencers when located in exons (Chen, Kobayashi et al. 1999; Caputi and Zahler 2001). Splicing elements can operate in an additive way, since the presence of several copies increase the effect on splicing regulation, either improving the affinity of the associated factor (Dominguez and Allain 2006) or increasing the copy numbers of the recruited factor (Wang, Rolish et al. 2004). Different regulatory elements can also act in a cooperative fashion. The exonic UAGG motif and intronic GGGG motif overlapping the 5'ss can function cooperatively in the silencing of the brain-specific CI cassette exon of the glutamate NMDA R1 receptor gene (Han, Yeo et al. 2005). All these different types of auxiliary cis-acting elements contribute significantly to recognition of the correct exons and in addition to their known role in alternative splicing, contribute significantly to the regulation of constitutive splicing (Cartegni, Chew et al. 2002; Black 2003).

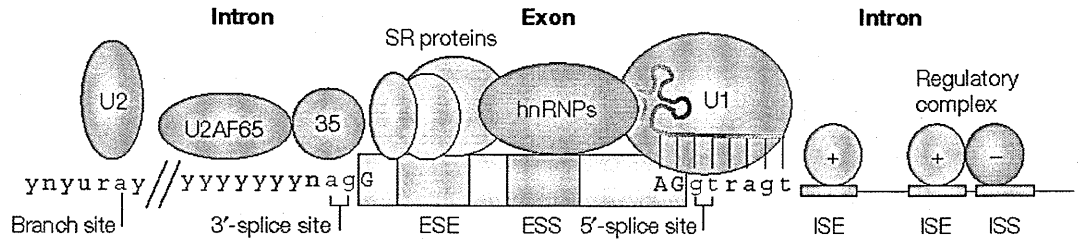

Figure 1.5: Regulatory elements in pre-mRNA splicing

Schematic representation of the possible distribution of canonical and additional splicing cis-acting elements. The canonical splicing signals that define the exon boundaries are relatively short and poorly conserved sequences. Only the GU at 5' end and the AG at 3' end together with the the branch-point adenosine are conserved (the consensus nucleotides are highlighted in red).

The polypyrimidine tract of variable length (represented by "y", cytosine or thymine) is reported upstream of the 3'-splice site. The basal components of the splicing machinery bind to the consensus sequences and promote assembly of the splicing complex. The U1 snRNP binds to the 5'-splice site, the U2 snRNP recognizes the branch site and the U2AF localizes to polypirimide tract and 3'splice site sequences.

Additional enhancer and silencer elements in the exons (ESE; ESS) and/or introns (ISE; ISS) allow the correct splice sites to be distinguished from the many cryptic splice sites that have identical signal sequences.

Trans-acting splicing factors can interact with enhancers and silencers and can accordingly be divided into two main groups: members of the serine arginine (SR) family of proteins and of the heterogeneous nuclear ribonucleoprotein particles (hnRNPs). In general, SR protein binding at ESE facilitates exon recognition whereas hnRNPs are inhibitory. Figure adapted from Pagani and Baralle (Pagani and Baralle 2004).

1.3.1 Splicing enhancers.

Exonic splicing enhancers (ESEs) were identified and extensively studied as regulators of alternative splicing (Black 2003) but they have also been implicated in some constitutive splicing event (Lavigueur, La Branche et al. 1993; Schaal and Maniatis 1999). Enhancer sequences have been the subject of many studies and most, but not all, are know to be recognized by members of the serine-arginine-rich protein family (SR) (Fig.1.5) (Graveley 2000). ESEs, throughout SR proteins binding (Fig.1.6), drive the exon definition by recruiting splicing factors and/or by antagonizing the action of nearby splicing silencer elements (Cartegni, Chew et al. 2002).

Diverse array of sequences can function as enhancers: most of them contain extended purine-rich sequences (more than 65% purine rich), but a second class of sequences lacking stretches of purines was also described (Tian and Kole 1995). In fact, studies conducted by Schaal and Maniatis identified pyrimidine-rich enhancers that are more than 67% pyrimidine-rich and function such as strong enhancers (Schaal and Maniatis 1999). Several groups, through the functional systematic evolution of ligands by exponential enrichment (SELEX) identified other functional ESE sequences. The SELEX carried out both *in vivo* (Coulter, Landree et al. 1997) and *in vitro* (Tian and Kole 1995; Schaal and Maniatis 1999; Boukis and Bruzik 2001), allowed the discovery of several both purine- and non-purine-rich enhancer. Furthermore this method was utilized to identify sequences that can act as ESEs in response to specific SR proteins characterizing SR-protein-specific sequence motifs (Liu, Zhang et al. 1998; Liu, Chew et al. 2000). These ESE motifs identified contain short (6–8nt), degenerate and sometimes partially overlapping sequences. The frequencies of individual nucleotides at each position were used to derive score

matrices in order to predict the location of SR protein-specific putative ESEs (Liu, Zhang et al. 1998; Liu, Chew et al. 2000) and design an ESEfinder web-based program (http://rulai.cshl.edu/tools/ESE/) (Cartegni, Wang et al. 2003).

In order to identify enhancer sequences some *ab initio* computational approaches have also been developed. For instance, the RESCUE-ESE (Relative Enhancer and Silencer Classification by Unanimous Enrichment) identifies putative ESE motifs by selecting hexamers that are enriched in exons against introns and weak against strong splice site scores (Fairbrother, Yeh et al. 2002; Fairbrother, Yeo et al. 2004). Similarly, the PESX approach (Putative Exonic Splicing Enhancers/Silencers) allows the identification of putative exonic splicing regulatory elements by comparing over-represented octamers in internal non-coding exons versus unspliced pseudo exons and the 5' untranslated regions of intronless genes. Oligonucleotides that are sufficiently overrepresented in noncoding exons are selected as PESE elements (Zhang and Chasin 2004; Zhang, Kangsamaksin et al. 2005).

A typical example of an exon bearing an exonic splicing enhancer (ESE) is the fibronectin EDA exon. Using a minigene transfection approach it was determined that the central part of the EDA exon is required for its inclusion into mRNA (Mardon, Sebastio et al. 1987). Then, Lavigueur *et al.* mapped the element required for EDA exon inclusion to a 9 nt purine-rich motif (5'-GAAGAAGAC-3') (Lavigueur, La Branche et al. 1993). Another group identified the same purine-rich sequence as a positive cis-acting element through transfection experiments (Caputi, Casari et al. 1994). In addition the ESE within the EDA exon has been found to interact with splicing factors belonging to a family of SR proteins (Lavigueur, La Branche et al. 1993).

A great number of intronic enhancers have been also characterized (Zheng 2004), but fewer large-scale screenings have been performed for these elements, compared with

the exonic ones. Most of the intronic enhancer regions have been identified by the study of disease-causing point mutations. Several of these mutations have been described to occur in the +20 to +40 regions downstream of the 3'ss causing the exclusion of the neighbouring exon from the mRNA. These sequences could be necessary in constitutive exon recognition or be involved in specific tissues or developmental stages (Venables 2007). A well studied ISE is the G run motif, which often occurs in clusters and acts by enhancing adjacent 5'ss or 3'ss (McCullough and Berget 1997) and the intronic CA repeats which in several cases can enhance splicing of upstream exons (Hui, Hung et al. 2005).

1.3.2 Splicing silencer.

In addition to cis-acting elements able to promote exon inclusion, other sequences able to inhibit splicing, exonic or intronic silencers, have been described. About one third of randomly selected human genomic DNA fragments (~100 bases long) were found to have inhibitory activity *in vivo*, when inserted within the middle exon sequence of the three-exon minigene (Fairbrother and Chasin 2000). These data suggested that silencer elements are very frequent, even though only limited number has been characterized by mutational analysis (Zheng 2004). Most described silencers are intronic elements, but several exonic silencer have been also reported (Pozzoli and Sironi 2005).

A typical silencer motif is still less characterized: the sequence can be purine or pyrimidine-rich and binds a diverse array of splicing factors (Fairbrother and Chasin 2000). Silencer sequences are generally interacting with the heterogeneous nuclear ribonucleoprotein (hnRNP) family (Fig.1.5) (Cartegni, Chew et al. 2002).

A silencer sequence can regulate splicing in different ways: by antagonizing the function of a nearby ESE or by recruiting factors that interfere with the splicing machinery by direct binding, throughout exon looping or by nucleation and cooperative binding (Cartegni, Chew et al. 2002; Matlin, Clark et al. 2005). Previous studies based on exon repression mechanism have shown that inhibition typically occurs during the initial ATP-independent recognition of splice sites. However, the dynamic nature of the splicing machinery suggests that any of the intermediates along the spliceosome assembly pathway are potential targets for biologically relevant regulation (House and Lynch 2006).

Most of inhibitory sequences have been identified through the study of single or a few regulated splicing events and the use of biochemical and cell transfection experiments (Pozzoli and Sironi 2005). Conversely, recent works have provided lists of putative silencers and methods that allowed a systematic identification of these splicing repressors. An *in vivo* fluorescence-based method was developed to systematically identify ESSs (Wang, Rolish et al. 2004). To this end, a three exon minigene system was used in which the first and the last exons encoded the complete green fluorescent protein (GFP) and a library of random decanucleotides was inserted in the central exon. The cells positive for GFP signals were selected and the putative silencer sequence were characterised (Wang, Rolish et al. 2004; Wang, Xiao et al. 2006). Zhang and colleagues proposed a computational approach for screening ESSs. The PESS (putative exonic splicing silencers) were identified in oligonucleotides underrepresented in noncoding exons (Zhang and Chasin 2004; Zhang, Kangsamaksin et al. 2005). The silencer elements identified by wide approaches do not relate to any specific exons/mRNA and have been only in part validated experimentally. These sequences might provide a useful resource to

compare putative silencers that will be identified in specific studies on splicing regulation (Pozzoli and Sironi 2005).

Several studies, by investigating the complex HIV-1 RNA alternative splicing, discovered the presence of exonic inhibitory elements within the tat coding exons. The first splicing silencer element identified in exon 2 specifically inhibits splicing at the upstream 3'ss flanking this exon (Amendt, Hesslein et al. 1994). By sequence analogy with this ESS element in tat exon 2 another negative cis-acting region within tat-rev exon 3 has been identified (Amendt, Si et al. 1995).

Similar inhibitory elements within introns, intronic splicing silencers (ISSs), have been less systematically studied. Many ISSs have been identified through the study of individual alternatively spliced genes (Zheng 2004). A well-known ISS is the binding sites for the polypyrimidine tract binding protein (PTB) splicing factors (Zheng 2004; Matlin, Clark et al. 2005). PTB normally recognizes pyrimidine-rich elements within introns and can function either by antagonizing U2AF65 action or by creating a region of silencing across the down regulated exon (Wagner and Garcia-Blanco 2001; Sauliere, Sureau et al. 2006). A particular example of intronic splicing silencer was described for CFTR exon 9 alternative splicing in which this extended region in intron 9 acts like a silencer by recruiting SR proteins (Pagani, Buratti et al. 2000).

1.3.3 RNA secondary structure.

The natural tendency of RNAs to fold in highly stable secondary and tertiary structures was shown to play an important role in exon definition for particular transcripts (Buratti and Baralle 2004). Numerous studies regarding the ability of RNA secondary structures to affect the splicing process focused on the conserved

regions at the exon boundaries (5'ss, 3'ss, and branch site). Although in each case different peculiarities are involved, structural elements can hind or permit the accessibility of a particular sequence by the splicing factors only to the acceptor site, to the donor site, or both. In the particular case of human tau exon 10 a stem–loop structure at the 5'ss directly affects the activity of the 5'ss (Grover, Houlden et al. 1999; Varani, Hasegawa et al. 1999). The stabilization of the secondary structure decreases exon inclusion whereas the destabilization of this stem–loop increases exon 10 inclusion (Donahue, Muratore et al. 2006). The RNA secondary structure can also affect the relative distance between splicing elements determining a considerable variation in the splice site usage (Buratti and Baralle 2004).

In addition to splicing consensus sequences, the secondary structure can affect cis-acting sequences such as exonic/intronic splicing enhancers (ESE/ ISE) or silencer (ESS/ISS). For example, in the fibronectin EDA splicing the secondary structure has been shown to play an important role in determining the splicing efficiency (Muro, Caputi et al. 1999). In this case, the peculiar secondary structure of the EDA exon exposes an ESE element, binding site for ASF/SF2, in a part of a loop region. The nearby ESS sequence seems to be determinant for stabilising the secondary structure of the RNA in such a way as to allow binding of SR proteins (Muro, Caputi et al. 1999).

Similarly, for the Survival of Motor Neuron 1 and 2 (SMN1 and 2) genes, Miyaso and colleagues have identified an ISE element consisting of a conserved stem loop structure in intron 7 (Miyaso, Okumura et al. 2003). Disruption of this structure leads to loss of binding of a yet unidentified trans-acting factor affecting the splicing process. An *in silico* analysis has shown that this ISE seems to be present in several intron sequences, raising the possibility that this structurally defined regulatory element may play a more general role in the splicing process (Miyaso, Okumura et

al. 2003). Other secondary structures have been described to participate to SMN2 exon 7 regulation such as a stem-loop element near its 3'ss region (Singh, Androphy et al. 2004) and a structure named terminal stem–loop 2 (TSL2) associated with the weak 5' ss (Singh, Singh et al. 2007).

These emerging studies on RNA structure propose the idea that many pre-mRNA sequences are capable of folding in well-defined secondary structures *in vivo*, which in turn affect the splicing process (Buratti, Baralle et al. 2006). A recent analysis in fact suggested that enhancer and silencer motifs are significantly enriched in single-stranded regions of pre-mRNA (Hiller, Zhang et al. 2007), implying that pre-mRNA secondary structure may play a general role in determining cis-acting element function.

1.4 Protein components: trans-acting factors.

Protein components involved in splicing can belong to the small nuclear ribonucleoprotein particles (snRNPs) or to the non-snRNP splicing factors (Fig. 1.5). Most of these proteins share very similar structural characteristics such as different number of RNA binding motifs and/or protein binding domains. These proteins usually target short sequence elements adjacent to sites of regulation.

The snRNP are the main components of the catalytic macromolecular complex called the spliceosome. The non-snRNP RNA binding proteins can be widely expressed or have a restricted expression patterns being involved in general or tissue-specific splicing events respectively (David and Manley 2008). Two families of RNA binding proteins, the SR proteins and the hnRNP factors, have been described as main components of distinct regulatory complexes with functional specificity in the splicing process (David and Manley 2008). However, additional tissue-specific factors have been described as playing a role in regulating particular splicing events

like Nova-1/2 (Jensen, Dredge et al. 2000), nPTB (Markovtsov, Nikolic et al. 2000) in brain and CELF proteins in striated muscle and brain (Ladd, Charlet et al. 2001).

1.4.1 Small Nuclear Ribonucleoprotein Particles (snRNPs).

The spliceosomal small nuclear ribonucleoprotein particles (snRNP) form part of the catalytic macromolecular complex called the spliceosome. The functional components of the spliceosome are the snRNPs U1, U2, U4, U5 and U6, responsible for splicing the vast majority of pre-mRNA introns (U2-type introns) (Patel and Steitz 2003). A group of less abundant snRNPs, U11, U12, U4atac and U6atac, but functionally analogous to the U1, U2, U4 and U6 snRNPs, are components of the so-called minor spliceosome and responsible for the excision of a rare class of introns, named U12-type (Hall and Padgett 1996). The U5 snRNP is shared by both spliceosomes (Patel and Steitz 2003).

Each snRNP particle consists of a snRNA molecule complexed with a set of eight Sm or Sm-like proteins and several specific factors (Will and Lührmann 2001). The snRNAs U1, U2, U4, U5 and U6 are characterized by their small size, metabolic stability and a high degree of sequence conservation (Kambach, Walke et al. 1999). The snRNAs U1, U2 and U6 to a minor extent have shown great complementarities to the splicing consensus sequences. The U1 snRNP initiates spliceosome assembly by binding to the 5'ss through base pairs between the single-stranded 5'terminal sequence of the U1 RNA and a conserved stretch of six nucleotides at the 5'ss. However, this interaction is switched subsequently for a mutually exclusive base pairing interaction with U6 snRNA. Before the 5'ss recognition, the U6 snRNA is extensively base paired (amounting to >20 base pairs) with U4 snRNA forming a unique U4/U6 particle. The U5 snRNA is then assembled in an ATP dependent

reaction with the U4/U6 snRNPs giving the U4/U6•U5 three-snRNPs particle (Konarska and Sharp 1987). This triple snRNP is believed to escort U6 to the spliceosome through its binding to U4 (Staley and Guthrie 1998). The U2 snRNA is involved in the branch point recognition, binding to the consensus CUPuAPy through a highly conserved stretch presents in its snRNA during the complex A formation (Staley and Guthrie 1998). The U5 snRNA has been reported to contain a conserved secondary structure and an essential loop involved in directly contacting exons sequences at the 5'ss and 3'ss (Turner, Norman et al. 2004).

All the snRNAs are transcribed by RNA polymerase II while U6, and presumably U6atac snRNA, are transcribed by the RNA polymerase III (Will and Lührmann 2001). The assembly of the U6 snRNP probably takes place entirely in the nucleus while the other pre-snRNAs are transported to the cytoplasm where snRNP assembly is initiated (Fischer, Sumpter et al. 1993). The pre-snRNAs export is dependent on a number of factors, including the m7G cap, the cap-binding complex (CBC), the export receptor CRM1/Xpo1 and RanGTP (Will and Lührmann 2001).

Following their export to the cytoplasm, the snRNA precursors bind with seven Sm proteins B', B, D1, D2, D3, E, F and G in order to form the structural core of snRNPs. The Sm proteins form three distinct heteromeric complexes prior to their interaction with the highly conserved Sm site of the U1, U2, U4 and U5 snRNAs (Raker, Plessel et al. 1996). After the formation of the Sm core, the 7-methyl guanosine (m7G) cap of these snRNAs is hypermethylated. Properly assembled Sm core, cap methylation and 3'end processing of the snRNAs are required for the subsequent nuclear import (Fischer, Sumpter et al. 1993).

The U6 snRNP composition differs from the other particles and contains a set of Sm-like (Lsm) proteins. The Sm-like proteins that belong to the Sm protein family are specifically required for the assembly of U6 snRNA. This subclass of Sm-like

proteins shares the conserved structural motif characteristic of all Sm-proteins family members (Mayes, Verdone et al. 1999).

Besides Sm and Sm-like proteins, other particle-specific proteins associate with snRNAs. For instance the mammalian U1 snRNP consists of ten different proteins, seven common Sm proteins, and other three (U1-70K, U1-A and U1-C) specific for U1 (Stark, Dube et al. 2001). U1-70K and U1-A proteins bind directly to the RNA and are involved in the splice site recognition and selection, while U1-C associates via protein-protein interactions with U1-70K and other Sm proteins (Nelissen, Will et al. 1994). A subset of U2 snRNP proteins also plays a critical role in tethering the U2 snRNP to the pre-mRNA. These proteins include the heteromeric splicing factors SF3a and SF3b (Brosi, Hauri et al. 1993) and bind the 20 nt region just upstream of the branch site in a sequence-independent manner (Gozani, Feld et al. 1996). At least five different proteins have been described as associating with U4/U6 snRNP, including a 15.5 KDa protein, polypeptides of 20, 60 and 90 KDa that form complex with one another (Teigelkamp, Achsel et al. 1998). U5 snRNP particle presents also a complex protein composition. Of the six U5 snRNP-specific proteins three are NTPases, U5 220 KDa and Brr2p are members of the DEXD/H box family, whereas Snu114p is a GTPase connected to translation elongation factor EF-2. Among these proteins it has been suggested that the U5 220 KDa factor acts as a protein cofactor for RNA-based catalysis in the spliceosome (Turner, Norman et al. 2004).

1.4.2 U2 snRNP auxiliary factor (U2AF).

U2 snRNP auxiliary factor (U2AF) is an essential pre-mRNA splicing factor complex that comprises two subunits of 65-kDa (U2AF65) and of 35-kDa (U2AF35). The U2AF commits the pre-mRNA to the first ATP-dependent step

(complex A) of the splicing process by binding the 3'ss and stabilizing the U2 snRNP association to the branch point (Kielkopf, Lucke et al. 2004).

The U2AF65 subunit specifically recognizes the polypyrimidine tract via its C-terminal RNA-binding domain, which consists of two canonical RNA recognition motifs (RRMs) and one U2AF homology motif (UHM) (Zamore, Patton et al. 1992; Kielkopf, Lucke et al. 2004). U2AF65 also stabilizes the U2 snRNP binding to the branch site through its N-terminal arginine- and serine-rich (RS) domain and promoting the base pairing (Valcarcel, Gaur et al. 1996). The U2AF35 binding domain of U2AF65 is located in the proline-rich region between the RS motif and RNA-binding motifs (Kielkopf, Rodionova et al. 2001).

The small subunit U2AF35 is involved the recognition of the essential AG dinucleotide at 3'ss during the earliest stage of spliceosome assembly (Merendino, Guth et al. 1999; Wu, Romfo et al. 1999). U2AF35 is made of a central UHM motif flanked by two zinc fingers in the N-terminal region (Birney, Kumar et al. 1993; Kielkopf, Lucke et al. 2004) while the C-terminus contains an RS domain and a glycine tract. U2AF35 binds both U2AF65 and the pre-mRNA through its RRM domain (Zhang, Zamore et al. 1992; Kellenberger, Stier et al. 2002).

Both U2AF subunits bind to intronic sequences only during the early steps of spliceosome assembly and subsequently are replaced by the U5 snRNP during the formation of the B complex (Wu, Romfo et al. 1999; Turner, Norman et al. 2004). In addition to constitutive splicing, several studies support the involvement of U2AF in the regulation of alternative spliced exons. In some cases, the SR proteins bound to a splicing enhancer together with U2AF35, are involved in recruiting U2AF65 to weak pyrimidine tracts (Graveley, Hertel et al. 2001). Although substantial evidence describes an important role for U2AF35 in enhancer dependent splicing, other experiments are inconsistent with this model (Graveley, Hertel et al. 2001). Recent

experiments have demonstrated that the small U2AF subunit is necessary for efficient enhancer-dependent splicing but its RS domain is not required for this function (Shepard, Reick et al. 2002).

1.4.3 Serine Arginine rich proteins (SRs).

SR proteins are a family of structurally related RNA binding proteins, highly conserved in metazoan cells that play multiple roles in splicing and in general mRNA metabolism (Graveley 2000; Huang and Steitz 2005). They contain one or two N-terminal RNA-recognition motifs (RRM) and a specific C-terminal domain rich in repeating arginines and serines, the "RS" domain (Birney, Kumar et al. 1993). The RNA binding and the RS domains are modular structures and they can be exchanged between different SR proteins. Differences among SR proteins structure depend on the length of the RS domain and the second RRM domain; when it is present, the sequence is often divergent from the canonical consensus sequence (Graveley 2000). In addition to SR proteins other factors involved in pre-mRNA metabolism can contain a RS domain. These proteins are usually referred as SR-related proteins and they include both subunits of U2AF, the U1-70K protein and several ATPases (Boucher, Ouzounis et al. 2001).

The structural organization of SR proteins suggests a model for their function. The RRM mediates sequence-specific binding to the mRNA, whereas the RS domain seems to be involved mainly in protein-protein interactions (Graveley 2000; Cartegni, Chew et al. 2002). However, a recent work has reported that the RS domains can also be involved in direct RNA contacts during splicing. A specific interaction between the RS domain and the branch point was described as promoting spliceosome formation (Shen and Green 2004).

The sequence-specific binding to pre-mRNAs is crucial for the function of SR proteins in the earliest step of spliceosome assembly (Graveley 2000; Sanford, Ellis et al. 2005). The binding specificity of individual SR protein has been studied using the SELEX technique showing that the consensus sequences recognized are highly degenerated (Liu, Zhang et al. 1998; Liu, Chew et al. 2000). Other SR proteins, indicating a redundancy in their functions, can also bind a target sequence recognized by one SR protein. In agreement with this observation, RNAi depletion of most SR proteins in *C. elegans* resulted in no detectable phenotype, suggesting a functional overlap among these factors (Longman, Johnstone et al. 2000). However, the lethality caused by loss of a SR protein in *Drosophila* (Ring and Lis 1994) or in the chicken DT40 cell line (Wang, Takagaki et al. 1996) argue against a large redundancy of SR proteins. Similarly deletion of SRp20, ASF/SF2 and SC35in the germ line of mice led to embryonic lethality indicating an essential role of this factors in the early embryonic development (Jumaa, Wei et al. 1999; Wang, Xu et al. 2001; Xu, Yang et al. 2005).

The SR proteins are required both for constitutive and alternative splicing events (Sanford, Ellis et al. 2005). Two non-exclusive models have been proposed to explain the role of SR protein in pre-mRNA splicing (Fig.1.6). One model is based on the ability of these splicing factors to bind exonic splicing enhancers (ESEs) and through their RS domain to recruit and stabilize U1 snRNP and U2AF binding to the 5' and 3'ss respectively. SR proteins have also been suggested to allow protein-protein interactions across introns binding the U1-70K factor and the U2AF35 and stimulating the usage of the 5' and 3'ss (Fig.1.6a) (Cartegni, Chew et al. 2002; Sanford, Ellis et al. 2005). In addition SR proteins have also been described in the recruitment of the U4/U6 U5 tri-snRNP to the spliceosome during the formation of the B complex (Roscigno and Garcia-Blanco 1995). The second model proposes that

a SR protein, bound to an ESE, can antagonize the negative effect of a juxtaposed silencer element (Fig.1.6b) (Cartegni, Chew et al. 2002; Sanford, Ellis et al. 2005). Due to their involvement in splice site selection, specifically in promoting the selection of the proximal 5' (Mayeda and Krainer 1992; Caceres, Stamm et al. 1994) and 3'ss (Caceres, Stamm et al. 1994; Bai, Lee et al. 1999), SR proteins have been reported to be important players in regulating alternative splicing. They exert a role in promoting U2AF binding to weak 3'ss (Wu and Maniatis 1993; Zuo and Maniatis 1996) or antagonizing the activity of negative splicing factor such as hnRNP A1 (Caceres, Stamm et al. 1994) or other SR proteins (Gallego, Gattoni et al. 1997). SR proteins have shown to promote exon inclusion when bound to a target site within exons (ESEs); however, in some cases they can act in a negative fashion. The negative effect on splicing can be mediated by the binding to an intronic sequence (ISS) (Ibrahim el, Schaal et al. 2005; Buratti, Stuani et al. 2007) or by the inhibitory property of the protein itself, as reported for SRp 38 (Barnard, Li et al. 2002).

SR protein activity is regulated through phosphorylation/dephosphorylation at multiple positions within the RS domain (Stamm 2008). This post-translational modification is a crucial step for the splicing organization inside the cell nucleus by affecting the RNA-binding activity and sub nuclear localization of the SR proteins (Misteli and Spector 1997). While localized predominantly in the nucleus, some (not all) SR proteins shuttle continuously between the nucleus and the cytoplasm (Caceres, Screaton et al. 1998). The RS domain phosphorylation is required for the translocation of the SR proteins from the cytoplasm to the nucleus and also for the recruitment of these factors from nuclear speckles ("splicing factor compartments") to the sites of pre-mRNA synthesis (Bourgeois, Lejeune et al. 2004).

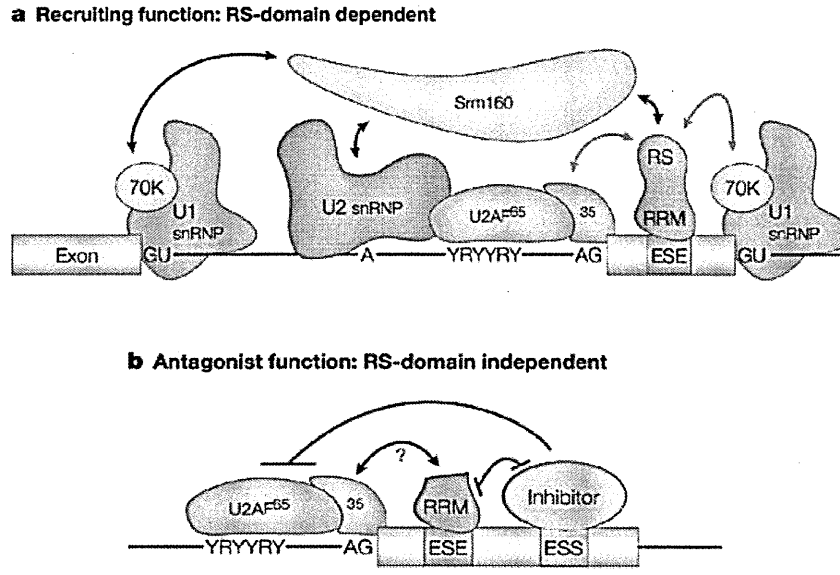

Figure 1.6: Models of SR protein action in ESE dependent splicing.

a) RS-domain-dependent mechanism. A SR protein bound to an exonic splicing enhancer (ESE) through its RNA-recognition motifs (RRM) contacts the splicing and through its RS domain recruits and stabilizes U1 snRNP and U2AF binding to the 5' and 3'ss, respectively. SR proteins can also allow protein-protein interactions across introns binding the U1-70K factor and the U2AF35 stimulating the usage of the 5' and 3'ss. For some ESE-dependent pre-mRNAs, indirect interactions (black arrows) are bridged by the splicing co-activator SRm160, which stimulates splicing of some ESE-dependent pre-mRNAs and also interacts with the U2 snRNP.

b) RS-domain-independent mechanism. In this case, the main function of the SR protein bound to an ESE is to antagonize the negative effect on splicing of an inhibitory protein that is bound to a juxtaposed exonic splicing silencer (ESS). The SR protein is shown without its RS domain, although this domain is normally present and might still promote U2AF binding. Inhibitory interactions are shown in red and the putative stimulatory interactions are double-headed arrow. These models are not mutually exclusive, and some splicing might involve a combination of these mechanisms. Figure adapted from Cartegni, Chew et al. (Cartegni, Chew et al. 2002).

Phosphorylation is also important for specific RNA recognition, since the high positive charge of unphosphorylated RS domains masks the specificity of the RNP domains and enhances non-specific binding (Tacke, Chen et al. 1997; Stamm 2008). While the functions of SR proteins in pre-mRNA splicing have been extensively studied, recent work has demonstrated their roles in numerous other steps of mRNA metabolism including mRNA nuclear export, mRNA stability and translation (Huang and Steitz 2005; Sanford, Ellis et al. 2005).

ASF/SF2: One of the best known SR proteins is the alternative splicing factor/splicing factor 2 (ASF/SF2). Different groups have highlighted the two basic properties of this SR protein. ASF/SF2 was described as an essential splicing factor necessary for the early step of splicing (Krainer, Conway et al. 1990) and was also characterized as an alternative splicing factor able to drive splice site selection (Ge and Manley 1990). ASF/SF2 can promote the recruitment of U1 snRNP to 5'ss (Kohtz, Jamison et al. 1994) to help 5'ss and 3'ss bridging (Wu and Maniatis 1993), and plays a role in splicing regulation, through binding to exonic splicing enhancers (Sun, Mayeda et al. 1993).

ASF/SF2, together with other SR proteins, is involved in additional roles in gene expression. For example, ASF/SF2 remains associated with the spliced mRNA and is able to shuttle between the nucleus and the cytoplasm (Caceres, Screaton et al. 1998), suggesting a role in mRNA export (Huang and Steitz 2005). In addition ASF/SF2 seems to regulate the mRNA stability by binding to the 3'UTR and enhancing RNA degradation in the cytoplasm (Lemaire, Prasad et al. 2002). ASF/SF2 can also stimulate translation of reporter mRNAs by associating with translating ribosomes (Sanford, Gray et al. 2004).

Despite these advances in understanding the functions of ASF/SF2 less is known about the physiological roles of this protein. Depletion of ASF/SF2 by RNAi resulted

in lethality in *C. elegans* (Longman, Johnstone et al. 2000) and tissue-specific deletion in mice resulted in defects in the developing heart (Xu, Yang et al. 2005). ASF/SF2 showed also an unexpected role in maintaining genomic stability by protecting cells from the deleterious effects of R-loop formation (Li and Manley 2005). In addition a recent work found that ASF/SF2 is an oncoprotein with roles in both the establishment and the maintenance of cell transformation (Karni, de Stanchina et al. 2007). In particular, ASF/SF2 has been found to control alternative splicing of the oncogene Ron which modulates cell motility (Ghigna, Giordano et al. 2005).

1.4.4 Heterogeneous ribonuclear proteins (hnRNPs).

The heterogeneous ribonuclear proteins (hnRNPs) family is a class of diverse RNA-binding proteins that associate with nascent pre-mRNA. These factors remain associated with pre-mRNA until its processing is completed and with mRNAs during export from nucleus to cytoplasm (Izaurralde and Mattaj 1995).

Approximately >20 proteins have been identified by two-dimensional gel electrophoresis of human hnRNP complexes with molecular weight ranging from 34 (hnRNP A1) to 120 kDa (hnRNP U) (Dreyfuss, Matunis et al. 1993; Dreyfuss, Kim et al. 2002). Some hnRNPs are extremely abundant (~100 million copies per nucleus), while others are present in lower amount (Kamma, Portman et al. 1995; Markovtsov, Nikolic et al. 2000). For many of these proteins multiple isoforms are produced by alternative splicing processes. This diversity is further increased by post-translational modifications of potential physiological significance, including phosphorylation, arginine methylation and SUMOylation (Dreyfuss, Kim et al. 2002; Martinez-Contreras, Cloutier et al. 2007).

The structure of hnRNP proteins is modular and consists of one or more RNA binding domains associated with an auxiliary domain often involved in protein-protein interactions (Dreyfuss, Matunis et al. 1993). For instance, the hnRNP A/B proteins contain two RNP domains at the N-terminus and a Gly-rich auxiliary domain at the carboxy end. HnRNP E1-E2 proteins contain three KH domains (Ostareck-Lederer et al., 1998). The hnRNP H family members contain two (2H9) or three (H, H' and F) quasi RNA recognition motifs (qRRMs) and one or two glycine rich auxiliary domains (Honore, Rasmussen et al. 1995).

The hnRNP proteins show general RNA-binding specificity and individual proteins display preference for specific sequences that tend to coincide with sites of functional importance in pre-mRNA processing. However, hnRNP proteins generally do not bind specific sites exclusively but recognize different RNA sequences with a wide spectrum of affinities. This RNA binding ability is further modulated by cooperative protein-protein interactions (Dreyfuss, Matunis et al. 1993; Dreyfuss, Kim et al. 2002).

The hnRNP proteins frequently mediate splicing repression, particularly through binding to exonic splicing silencer elements or by sterical interference with other protein interaction (Cartegni, Chew et al. 2002). Nevertheless, depending on the position of the splicing regulatory elements hnRNPs can also associate with enhancer elements to help exon inclusion (Caputi and Zahler 2002) and a generic role for hnRNP A1 and F/H proteins in intron definition has been recently proposed (Martinez-Contreras, Fisette et al. 2006).

Although many of the hnRNPs localized in the nucleus, a subset of these proteins shuttles continuously between nucleus and cytoplasm. This indicates a role of these proteins in nuclear export and in other cytoplasm processes (Pinol-Roma and Dreyfuss 1992; Martinez-Contreras, Cloutier et al. 2007).

hnRNP A1: Among the hnRNP proteins, the most abundant and extensively studied is hnRNP A1 (Mayeda and Krainer 1992). hnRNP A1 binds RNA through two RNA recognition motifs at its N terminus while the C-terminal domain comprises several RGG repeats, which also contribute to RNA binding. *In vivo*, four arginine residues within the RGG repeats of hnRNP A1 are methylated and are thought to influence the RNA-binding properties (Kim, Merrill et al. 1997). The C-terminus also includes the M9 motif, involved in hnRNP A1 nuclear import and export (Izaurralde, Jarmolowski et al. 1997).

Although at steady state hnRNP A1 is predominantly nuclear, it shuttles rapidly between nuclear and cytoplasmic compartments (Pinol-Roma and Dreyfuss 1992). The shuttling of hnRNP A1 is subject to regulation and it has been proposed to have a role in cell proliferation, survival, and differentiation of normal and transformed cells (Iervolino, Santilli et al. 2002). The endogenous hnRNP A1 is weakly phosphorylated in cells grown under normal conditions while upon osmotic shock a cytoplasmic accumulation of hnRNP A1 is induced, concomitant with an increase in its phosphorylation state (van der Houven van Oordt, Diaz-Meco et al. 2000).

Since the early 1990s hnRNP A1 has been associated with the regulation of alternative splicing process (Caceres and Kornblihtt 2002). The first hnRNP A1-dependent ESS was identified in studies of HIV tat exon 2 regulation (Amendt, Hesslein et al. 1994; Amendt, Si et al. 1995). HnRNP A1 splicing repression and the existence of A1-dependent ESSs have been documented in other numerous examples in humans (Del Gatto-Konczak, Olive et al. 1999; Kashima and Manley 2003; Disset, Bourgeois et al. 2006).

Despite its extensive characterization only a small number of high affinity sites have been obtained by SELEX (Burd and Dreyfuss 1994). Interestingly a recent study, based on searching for endogenous hnRNP A1 target RNAs, reported a specific

binding for the pri-miR-18a. This result highlighted a new role for hnRNP A1 as auxiliary factor involved in miRNA processing (Guil and Caceres 2007).

Several mechanisms have been suggested to explain hnRNP A1-mediated splicing repression. The hnRNP A1 antagonizes in a concentration-dependent way the activity of SR proteins on the selection of 5'splice site and can bind to certain exon splicing silencers to prevent the use of adjacent 3'splice site (Mayeda and Krainer 1992; Smith and Valcarcel 2000). Another mechanism was elucidated studying the alternative splicing of the hnRNP A1 transcript itself. The hnRNP A1 protein can function as splicing repressor, influencing the alternative splicing of its own pre-mRNA, by binding to a conserved intronic element present in both introns surrounding the alternative exon 7b (Chabot, Blanchette et al. 1997; Blanchette and Chabot 1999; Hutchison, LeBel et al. 2002). Cooperation between the two A1 complexes on these sites was suggested as promoting ''looping-out'' of the intervening RNA, including exon 7b, thereby inhibiting splicing.

hnRNP A1 is also implicated in various post-splicing activities, such as mRNA export (Izaurralde, Jarmolowski et al. 1997), mRNA stability (Hamilton, Burns et al. 1997) and cap-dependent and internal ribosome entry site-mediated translation (Svitkin, Ovchinnikov et al. 1996; Bonnal, Pileur et al. 2005). In addition it was reported that hnRNP A1 has the ability to disrupt the higher order structure of telomeric DNA indicating a role in telomere maintenance (Zhang, Manche et al. 2006). HnRNP A1 in fact binds with high affinity to telomeric single stranded DNA sequences (LaBranche, Dupuis et al. 1998; Dallaire, Dupuis et al. 2000) and interacts with telomerase RNA *in vitro* (Fiset and Chabot 2001). A severe hnRNP A1 reduction in mouse erythroleukemic cells resulted in shortened telomeric repeats, which can be lengthened by restoring hnRNP A1 expression (LaBranche, Dupuis et al. 1998).

1.5 Alternative splicing.

Mechanisms that increase protein diversity include the usage of multiple transcription start sites, alternative pre-mRNA processing, polyadenylation, pre-mRNA editing and post-translational modifications. Among these mechanisms, alternative pre-mRNA splicing is considered the most important source of protein diversity. This process was described in nearly all metazoan organisms as a mode of producing functionally diverse polypeptides from a single gene (Maniatis and Tasic 2002). Alternative splicing in fact generates a large number of mRNAs from the surprisingly low number of human genes encoding proteins with slight or opposing functional differences, with profound biological consequences (Lopez 1998).

The process of alternative splicing is highly regulated in developmental stages and in different tissues (Black 2003). Moreover the selection of the correct splicing variants in a given cell type and/or in a particular cell condition is considered to be regulated by multiple (sometimes overlapping) exonic and/or intronic splicing enhancers and silencers (Cartegni, Chew et al. 2002; Ladd and Cooper 2002). Because a single primary transcript can have several regions that each undergo alternative splicing, the resulting combinatorial effects of selecting different splice sites can be very pronounced, and genes that code for tens to hundreds of different isoforms are common (Graveley 2001). One of the most striking examples of alternative splicing complexity has been described in *Drosophila melanogaster* axon guidance receptor gene, *Down syndrome cell-adhesion molecule* (*Dscam*). The pre-mRNA of this gene can potentially produce 38,016 different mature transcripts by alternative splicing (Schmucker, Clemens et al. 2000). Even if only a subset of these mRNAs is ever produced *in vivo*, this combinatorial use of alternative exons still represents an incredible source of diversity, especially given that the entire *Drosophila* genome consists of only ~14,000 genes (Adams, Celniker et al. 2000).

At first, in humans, splicing was thought to be only a minor processing pathway affecting about 5% of all genes (Sharp 1994), but over time, it became clear that it is very abundant. Bioinformatic studies showed that 59% of the 245 genes present on chromosome 22 are alternatively spliced, and DNA microarray experiments indicate that 74% of all human genes are alternatively spliced (Johnson, Castle et al. 2003), suggesting that alternative splicing of human genes is the rule and not the exception.

Every conceivable pattern of alternative splicing is present in nature (Fig. 1.7). Exons can have different 5'ss or 3'ss that can be alternatively used. Single cassette exons can reside between two constitutive exons such that the alternative exon is either included or skipped. Alternatively, multiple cassette exons can reside between two constitutive exons such that the splicing machinery must choose between them. In these systems, special mechanisms must enforce the exclusive choice (Smith and Nadal-Ginard 1989). Introns can be retained in the mRNA and become translated. The 5-terminal exons of a mRNA can be switched through the use of alternative promoters and alternative splicing. Similarly, the 3-terminal exons can be switched by combining alternative splicing with alternative polyadenylation sites. In addition, these individual patterns can be combined in a single transcription unit to produce a complex array of splice isoforms. Moreover, changes in alternative splicing can modulate transcript expression levels subjecting mRNAs to nonsense-mediated decay (NMD) by creating a stop codon within the coding sequence or by altering the structure of the gene product by inserting, or deleting, novel protein parts (Faustino and Cooper 2003).

The mechanisms that determine which splice site has to be utilized and/or which exon has to be chosen in different cell types or developmental stages have still not been precisely defined.

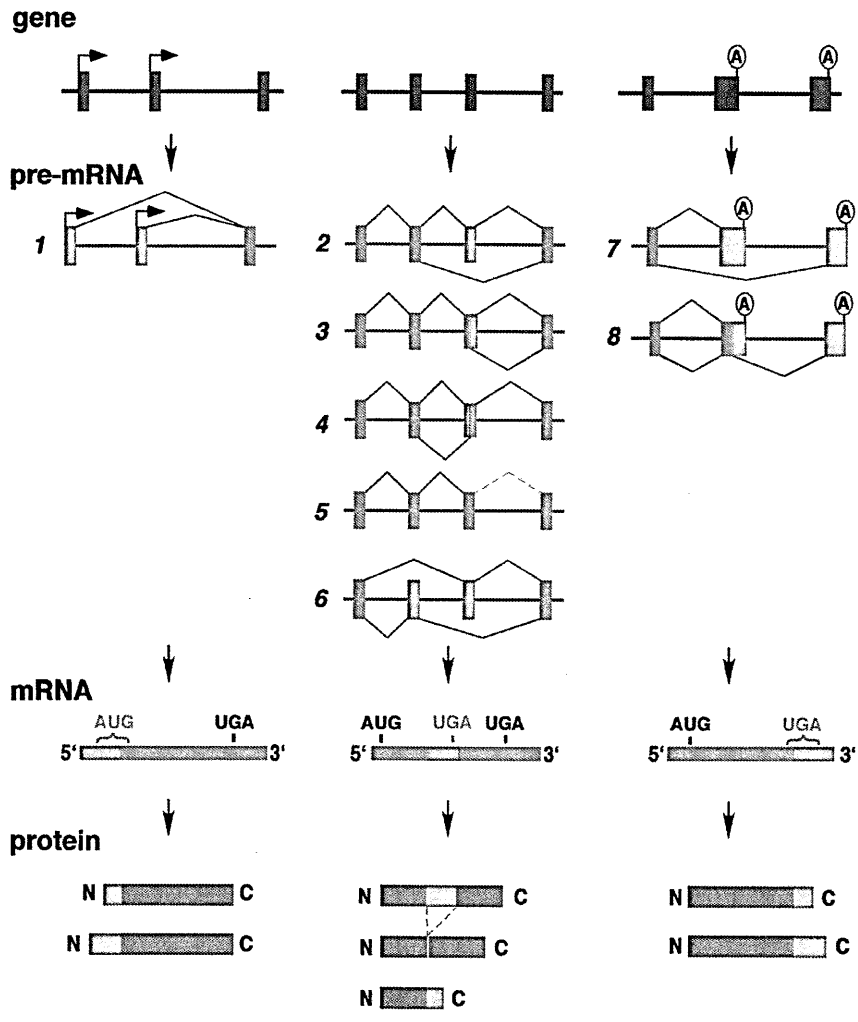

Figure 1.7: Patterns of alternative splicing.

Alternative splicing generates different segments within mRNAs.

Alternative promoters: selection of one of multiple first exons results in variability at the 5' end of the mRNA. Red indicates variable regions within the mRNA and encoded protein (*1*). Alternative splicing of internal exons: the alternative splicing patterns for internal exons include the cassette exon (*2*), alternative 5' splice sites (*3*), alternative 3' splice sites (*4*), intron retention (*5*), and mutually exclusive exons (*6*). Alternative terminal exons: selection of one of multiple terminal exons results from a competition between cleavage at the upstream poly(A) site or splicing to the downstream 3' splice site (*7*). There are also examples of competition between a 5' splice site and a poly(A) site within an upstream terminal exon (*8*). Variability at the 3' end of the mRNA produces either proteins with different C termini or mRNAs with different 3'-UTRs. Figure adapted from Faustino and Cooper (Faustino and Cooper 2003).

Much progress has been made in identifying the "combinatorial code" composed by cis-acting elements and trans-acting factors involved in the regulation of alternative splicing. High-throughput technologies like large-scale sequencing and microarrays analysis are providing opportunities to address these key questions (Ben-Dov, Hartmann et al. 2008).

1.5.1 Combinatorial control of trans-acting factors.

The information that determines an alternative splicing event is very difficult to characterize. Along with the main determinants of exon-intron definition, such as splice sites or enhancer/silencer elements, many more splicing factors have been involved in splicing regulation. Alternative splicing in mammals in fact is largely controlled by combinatorial binding of basal splicing factors to the pre-mRNA. Although several tissue-specific factors associated to a particular alternative splice site event have been identified (Markovtsov, Nikolic et al. 2000; Ladd, Charlet et al. 2001), the model for splice site choice consists of a selective usage of an exon due to the binding of a distinct set of generic splicing factors (Smith and Valcarcel 2000; Mabon and Misteli 2005). This combinatorial model is supported by the observation that *in vivo* and *in vitro* the counteracting activities of multiple antagonistic factors can regulate alternative splicing, suggesting that the physiological concentration of competing splicing factors is important for regulation of splice site selection (Caceres, Stamm et al. 1994; Hanamura, Caceres et al. 1998).

For example, hnRNP A1 can modulate splice site selection by antagonizing the activity of ASF/SF2 protein (Caceres, Stamm et al. 1994). It was observed that high relative concentrations of A1 favour the choice of distal 5'ss, an excess of ASF/SF2 result in the proximal 5'ss choice both *in vitro* and upon overexpression *in vivo*

(Mayeda and Krainer 1992; Caceres, Stamm et al. 1994). The same competitive effect between hnRNP A1 and ASF/SF2 has been reported also in the case of alternative 3'ss selection. In the same manner hnRNP A1 promotes the use of the distal 3'ss, while ASF/SF2 promotes the use of the proximal 3'ss (Caceres, Stamm et al. 1994; Bai, Lee et al. 1999).

In addition to these indirect data, a recent work based on a quantitative single-cell imaging provided new evidence for the combinatorial model for alternative splicing reporting the first *in vivo* evidence for a distinct association of splicing factors with alternative spliced pre-mRNAs (Mabon and Misteli 2005).

1.6 Exon definition and Intron definition model.

In most of the cases, vertebrate genes contain multiple short exons separated by intronic sequences that can be significantly longer. A typical human pre-mRNA molecule can be up to 30 kb long and contains about ten exons separated by much longer and variable intronic sequences (Hawkins 1988). The average length of an internal exon is around 137 nt (Hawkins 1988) and the exons that are more than 300 nt long or are shorter than 50 nt seem not to be favoured by the splicing machinery (Black 1991; Sterner and Berget 1993).

The discrepancy between the lengths of human exons compared to the introns led to the proposal of the "exon definition" model (Fig.1.8A). In this model the exon is the unit recognized by the splicing machinery and the splice sites are first paired across exons, with subsequent spliceosome assembly proceeding through pairing of neighbour exon units (Robberson, Cote et al. 1990; Berget 1995).

Evidence for the exon definition model arises from the observation that the first and the last exon require a special mechanism for their recognition. It has been shown that 5' capping and the 5'ss are necessary to define the first exon (Izaurralde, Lewis

et al. 1994). On the other hand, last exon is defined by 3'ss and polyadenylation signal (Niwa and Berget 1991). Moreover, it was shown that the exon length might affect splicing. A minimal separation of the sites seems to be required to prevent steric hindrance of the factors that recognize the splice sites bordering an exon (Black 1991). This indication is supported by the observation that a constitutively recognized internal exon was skipped by *in vivo* splicing machinery if its size was smaller that 50 nt (Dominski and Kole 1991). Complementary studies showed that the extension of the 18-nt N1 exon of the mouse c-src gene up to 109 nt leads to its constitutive inclusion. This finding suggests that the exon is normally skipped because it is too short to allow spliceosome assembly at both ends simultaneously (Black 1991). Further i*n vitro* analysis on the expansion of internal exons to length above 300 nt causes the activation of cryptic splice site located within the exon or lead to exon skipping indicating that splicing efficiency is affected by length (Robberson, Cote et al. 1990). In fact, less than 1% of the known internal exons in vertebrate are longer than 400 nt (Hawkins 1988). Conversely the experiments obtained by expanding the central exon in the dihydrofolate reductase minigene with random DNA fragments from *E.coli*, demonstrated that its expansion up to 1200 nt does not compromise the correct inclusion into mRNA (Chen and Chasin 1994). To explain these data a compensatory relationship between exon and intron size has been proposed. In fact, Sterner et al. observed that long internal exons are problematic for recognition if they are flanked by long introns, suggesting that these exons might be flanked by short introns (Sterner, Carlo et al. 1996). In this case, the splice sites can be paired across introns rather than exons as proposed by the "intron definition" model (Fig.1.8B). This model in fact suggests that introns are the units recognized by the splicing machinery proposing a sort of scanning mechanism where the 5' and 3' splice signals are initially recognized and paired across the intron

(Guthrie 1991; Berget 1995). Intron definition is thought to be the predominant way of splicing in transcripts containing short introns and long exons (Sterner, Carlo et al. 1996). Evidence supporting this mechanism arises from observations made in yeast, where RNA messengers often have a unique intron and its length is usually below 100 nt (Goguel and Rosbash 1993).

In *Drosophila* there are two different classes of introns: long, vertebrate-like introns that possess a 3' pyrimidine tract and short, yeast-like introns that lack this consensus sequence (Talerico and Berget 1994). 50% of small introns are less than 100 nt and are often flanked by large exons (Hawkins 1988). Mutants of 5'ss and expansion of the size of these small introns showed the induction of intron retention and cryptic splice site activation, respectively supporting the intron definition model as way of recognition (Talerico and Berget 1994).

All these observations suggest mechanistic differences in the process of splice site selection in pre-mRNAs containing small exons or small introns. It has been indicated that splice sites are initially paired across the shortest distance so both exon definition and intron definition might occur in different parts of the pre-mRNA of the same gene (Sterner, Carlo et al. 1996).

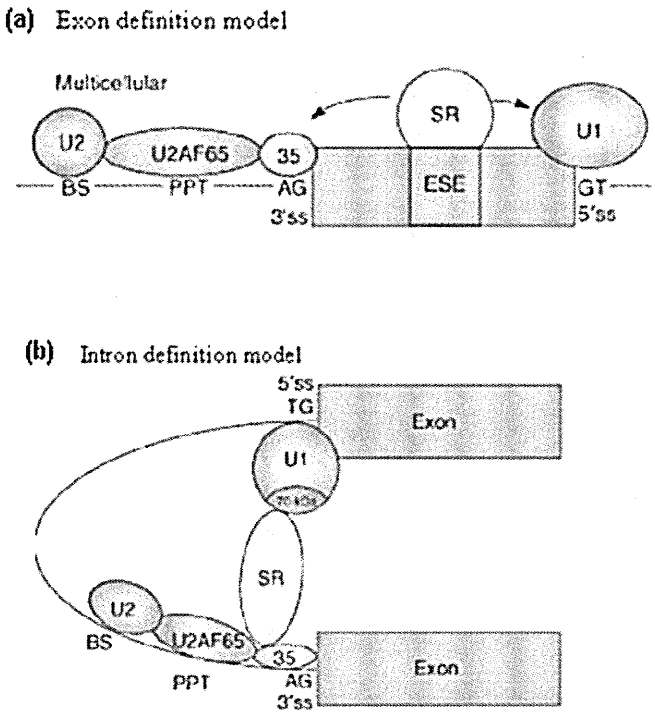

Figure 1.8: Exon definition versus intron definition models.

(A) Exon definition model: a vertebrate gene consists of multiple short exons separated by considerably longer introns. In this model the exon is recognized as a unit during early spliceosome assembly. Multiple factors interact with exonic sequences in order to defining the 5' and the 3' borders of the exon. SR proteins can help this "cross exon" recognition through the binding exonic splicing enhancers (ESEs) and recruiting U1 snRNP to the 5'ss, U2AF65 and U2AF35 subunits to the PPT and to the 3'ss, respectively.

(B) Intron definition model: it has been proposed for the systems in which pre-mRNA has small introns. In this case the intron, rather than exon, is the recognized unit. Multiple factors favour the 5' and the 3' ss pairing at the intron ends. SR proteins function in a "cross intron" recognition complex by bridging together the U1 snRNP bound to the upstream 5'ss and the U2AF65 and U2AF35 subunits bound to the PPT and to the AG of the downstream 3'ss, respectively. Adapted from Ram and Ast (Ram and Ast 2007).

1.7 Splicing is a co-transcriptional process.

The transcriptional process includes the formation of a pre-initiation complex, the transcription initiation, elongation, termination and finally the dissociation of RNA polymerase II (Pol II) from the DNA template. Then the pre-mRNA transcript undergoes several processing steps including capping, splicing, polyadenylation, surveillance and export (Fig. 1.9). *In vitro* each of these modifications can occur independently from the others, as found in reconstituted systems on purified pre-mRNA substrates. However, *in vivo* these reactions influence one another's efficiency revealing a functional relationship among them and indicating that most mRNA processing reactions occur co-transcriptionally (Bentley 2002; Bentley 2005). These observations led to the proposal that the RNA modifications occur in a "gene expression factory" composed of machines linked together for the purposes of efficiency and regulation (Bentley 2002; Bentley 2005).

To date, two models have been proposed to explain the connection between the different steps of RNA transcription and processing. The "recruitment model" derived from the observation that several trans-acting factors can interact directly with the RNA Pol II and thus increasing their own concentration in the proximity of the nascent transcript (Bentley 2005). The carboxy-terminal domain (CTD) of the RNA Pol II in fact has been shown to play a central role in coupling transcription to pre-mRNA processing acting as assembly platform for proteins involved both in transcription and pre-mRNA processes regulation (Bentley 2005). In particular, recent studies reported an important connection between SRp20 and the CTD tail for the EDA alternative exon regulation. Mature mRNAs transcribed by a RNA Pol II, lacking the CTD tail, showed a much higher percentage of EDA inclusion compared to the ones transcribed by the entire RNA Pol II. The CTD-dependent silencing

seems to be mediated by its ability to recruit the SRp20 splicing factor (de la Mata and Kornblihtt 2006).

An alternative to the recruitment model is the "kinetic model", mainly linked to the RNA Pol II transcriptional elongation rate (Kornblihtt, de la Mata et al. 2004). This model was originally proposed by experiments in which RNA Pol II pausing sites were artificially introduced into a gene, delaying the transcription of a splicing inhibitory element and therefore resulting in higher inclusion levels of an alternative exon (Roberts, Gooding et al. 1998). Further evidence supported this model indicating that transcription can affect splicing acting at different cis- and trans-acting levels (Kornblihtt, de la Mata et al. 2004). For instance differences in promoter architecture have been shown to affect the subsequent selection of the fibronectin EDA alternative exon and the CFTR exon 9 (Cramer, Pesce et al. 1997; Cramer, Caceres et al. 1999; Pagani, Stuani et al. 2003). However, *in vivo* most of the genes have a single promoter and the regulation of splicing has been reported to occur through the binding of different transcription factors. In line with this view it has been described that transcriptional activators can affect alternative splicing (Nogues, Kadener et al. 2002). Additionally, the presence of enhancer sequences, like SV40 next to the promoter can stimulate RNA Pol II elongation while the deletion of this enhancer causes a reduction in exon skipping (Kadener, Fededa et al. 2002).

A less explored factor involved in regulation of splicing via RNA Pol II elongation is the chromatin packaging. Changes in its structure due to acetylation can alter splicing. It was reported in fact that cell treatment with a potent inhibitor of histone deacetylation inhibits EDA inclusion (Nogues, Kadener et al. 2002). Recently, a role of SWI/SNF chromatin remodelling complex was also described for alternative splicing. This factor was shown to stimulate the inclusion of alternative exons in the

CD45 gene by decreasing the RNA pol II elongation rate and facilitating the recruitment of the splicing machinery (Batsche, Yaniv et al. 2006).

A more direct proof for the kinetic model derived from studies on RNA Pol II elongation rate (Roberts, Gooding et al. 1998; de la Mata, Alonso et al. 2003). A slow Pol II and/or the presence of internal transcriptional stalling sites, results in an increased inclusion of alternative exon harbouring a weak 3'ss. By contrast, when a highly processive RNA Pol II transcribes the same pre-mRNA, the weak alternative splice site is unable to compete with the stronger downstream 3'ss, which results in skipping of the alternative exon (de la Mata, Alonso et al. 2003). Recently a reciprocal coupling between splicing and transcriptional elongation has also been reported. Splicing proteins have been involved in transcriptional elongation *in vitro* (Fong and Zhou 2001) and specific depletion of SC35 showed RNA Pol II accumulation and attenuated elongation *in vivo* (Lin, Coutinho-Mansfield et al. 2008). Furthermore, it was also described that an efficient RNA Pol II transcription is strictly connected with the presence of intronic sequences (Furger, O'Sullivan et al. 2002). Taking these evidences together a complex view has emerged from the studies focused on the coupling between transcription and pre-mRNA processing suggesting that both recruitment of factors to the CTD and RNA Pol II kinetic are involved in this connection (Kornblihtt, de la Mata et al. 2004).

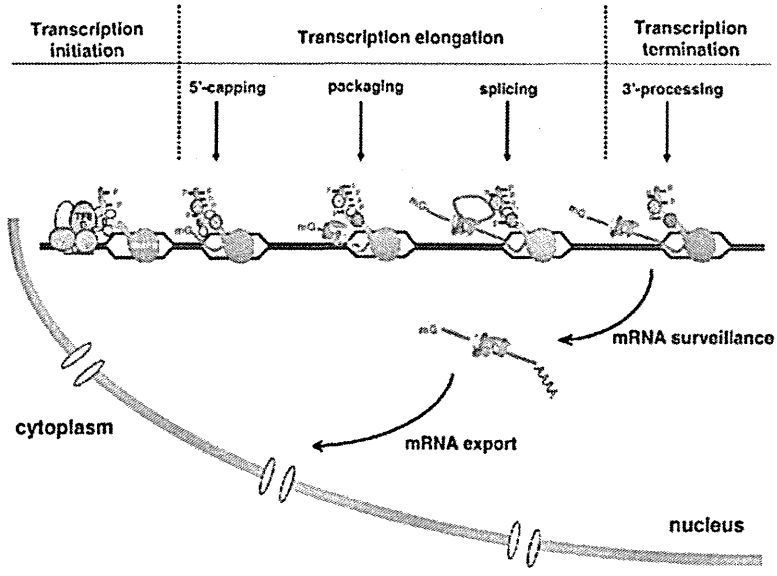

Figure 1.9: Coupling mRNA transcription and processing.

The production of a correct mRNA is a coordinated and multiple-step process that occurs in the nucleus of eukaryotic cells. RNA Pol II, and specifically the C-terminal domain (CTD) of its largest subunit directs these processes in a way that links CTD phosphorylation changes to different factors binding. These proteins in turn affect the subsequent processing steps and/or help to recruit processing/packaging factors to the nascent transcript. The maturation of pre-mRNA includes the capping at its 5' end, the introns removal during splicing, and the cleavage and the polyadenylation of the 3' end. After going through the mRNA surveillance system, the matured mRNA is exported to the cytoplasm for translation. Each stage of RNA Pol II transcription and the steps of co-transcriptional processing are indicated. Figure adapted from Li and Manley (Li and Manley 2006).

1.8 Splicing mutations.

Considering the complexity of the pre-mRNA process, it is not surprising that splicing mutations are directly linked with aberrant splicing processes. As a consequence the study of the network of interactions between defective splicing and occurrence of disease has become a central issue in the medical research field (Cartegni, Chew et al. 2002; Faustino and Cooper 2003; Garcia-Blanco, Baraniak et al. 2004; Pagani and Baralle 2004). More than a decade ago, it was estimated that at least 15% of mutations that cause genetic disease affect pre-mRNA splicing (Krawczak, Reiss et al. 1992). In the specific case of NF1 and ATM genes, the genomic variations which affect splicing process seem to represent up to 50% of all mutations that lead to gene dysfunction (Teraoka, Telatar et al. 1999; Ars, Serra et al. 2000). At the mRNA processing level, splicing alterations can produce different outcomes such as exon skipping, activation of cryptic splice sites or intron retention (Buratti, Baralle et al. 2006). In mammals, exon exclusion from the mature transcript is the most prevalent splicing outcome followed by the inactivation of the donor site (Krawczak, Reiss et al. 1992).

The alteration of splicing processes due to a mutation can produce an erroneous mRNA followed by a subsequent loss of function of its translated protein. In other cases, a nucleotide substitution can modify the normal level of an alternative spliced exon and can result in an imbalance in the different protein isoforms produced (Pagani and Baralle 2004).

It is important to notice that the splicing process can be affected by mutations dispersed throughout any given part of a gene sequence including splice sites, exons and introns (Fig. 1.10) but the functional mechanisms of many of these mutations remains poorly characterised (Pagani and Baralle 2004).

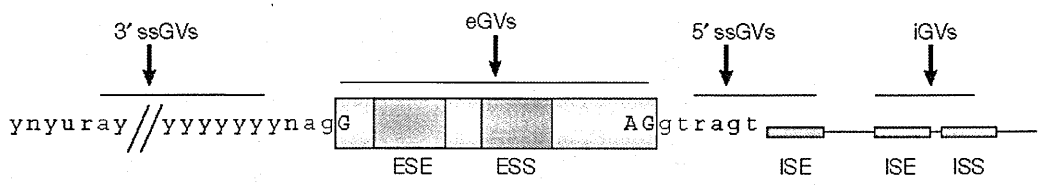

Figure 1.10: Splicing mutations.

Genomic variants (GVs) can affect all the different cis-acting regulatory elements: the canonical 5' and 3' splice site and the accessory elements, enhancers (red rectangles) and silences (blue rectangles). Exonic genomic variants (eGVs), even if they can affect splicing, are usually classified for their effect at the proteins level as missense, nonsense or synonymous mutations. Intronic variants (iGVs) might be located within approximately 50 bp from the splice sites or deep in the introns and ofter are considered polymorphism. Figure adapted from Pagani and Baralle (Pagani and Baralle 2004)

In the specific case of ATM and NF1 genes the 13% and 11% of splicing mutations, respectively, have been erroneously classified as frameshift, missense or nonsense mutations if the analysis had been limited to the effect at the protein level (Pagani and Baralle 2004). Moreover, the presence of mutations can affect the mRNA secondary structure (Buratti and Baralle 2004). In addition to the cis-acting elements variants, mutations associated with trans-acting factors can result in a more general splicing defect whereas proteins involved in a specific splicing event can affect a particular subset of spliced mRNAs. Therefore, components of the basal splicing machinery, if mutated, have the potential to target a huge number of mRNAs (Faustino and Cooper 2003). In general mutations affecting spliceosome components result in cell lethality both in yeast and metazoans (Golling, Amsterdam et al. 2002) but in the particular case of SMN1, involved in snRNPs assembly, the homozygous loss of this gene causes spinal muscular atrophy (SMA) (Wirth 2000).

Detailed knowledge about the connection between splicing mutations and disease can have a significant impact on the diagnosis and for the treatment of genetic diseases. The adequate classification of mutations in term of effect at mRNA and/or protein level is essential for the development of new therapeutic strategies and compounds (Cartegni, Chew et al. 2002; Pagani and Baralle 2004).

1.8.1 Canonical splice site mutations.

Splicing signals are frequent targets of mutations in genetic diseases and cancer (Fig. 1.10). Most of the changes are single-point mutations affecting the conserved consensus signal sequences at the donor or acceptor splice sites (Pagani and Baralle 2004). Consensus sequences are limited to 9 bases at the 5' exon/intron junction and 4 bases at the 3' intron/exon junction (Shapiro and Senapathy 1987). It has been shown that mutations occurring in one of the first two bases (GU) in intron, immediately downstream of 5'ss, as well as AG in intron, immediately upstream of 3'ss completely abolish splicing (Langford, Klinz et al. 1984). Subsequently, a study of more than a hundred splice-site mutations showed that point mutations affecting the 5'ss were more common than those at the 3'ss (62% vs 26%) (Krawczak, Reiss et al. 1992). At the 5'ss, mutations affecting the GT residue at position +1, +2 are the most common (Krawczak, Reiss et al. 1992), followed by mutations at position +5 (Pohlenz, Dumitrescu et al. 2002). Mutations at these positions are thought significantly to reduce the pairing of the donor splice site with the complementary site in the U1 snRNP, which is one of the first steps in the complex process of pre mRNA splicing (Krawczak, Reiss et al. 1992; Kramer 1996). In fact, cotransfection experiments with U1 snRNA mutant complementary to the nucleotide substitution in this region showed rescue the aberrant splicing in the NF1 exon 3 (Buratti, Baralle et

al. 2004). As for the 5'ss, it was shown that mutations affecting the 3'ss at the conserved -2 and -1 sites (A and G dinucleotides respectively) are most common and have the effect of completely abolishing splicing (Langford, Klinz et al. 1984), although mutations affecting the +1 site are also observed (Ikeda, Takagi et al. 2001). Mutation of the adenine residue involved in the lariat formation strongly reduces splicing efficiency of the downstream exon (Reed and Maniatis 1988). Indeed, the use of the normal 3'ss can be restored either improving the match of a cryptic branch site to the branch site consensus or by introducing mutant U2 snRNA with greater complementarity to the cryptic branch site (Zhuang and Weiner 1989).

1.8.2 Exonic and intronic mutations.

Nucleotide substitutions that involve accessory sequences can be classified as loss or gain of function splicing mutations if the splicing element considered is destroyed/weakened or created/enhanced, respectively (Faustino and Cooper 2003; Garcia-Blanco, Baraniak et al. 2004).

A well-studied example in which the effect of a splicing mutation on an exonic regulatory sequence has been analyzed is represented by exon 7 in the SMN1/SMN2 genes (Fig. 1.11). SMN1 and SMN2 represent two nearly identical copies of the survival motor neuron gene in humans. SMN1 predominantly produces full-length transcripts, whereas SMN2 mostly produces transcripts lacking exon 7. Deletion of SMN1 coupled with the inability of SMN2 to compensate for the loss of SMN1 leads to spinal muscular atrophy (SMA), the second most common autosomal recessive genetic disorder in humans (Lefebvre, Burglen et al. 1995; Wirth, Brichta et al. 2006). Comparison between SMN1 and SMN2 revealed a critical C to T substitution at the sixth position (C6U transition in transcript) of exon 7 of SMN2 sufficient to

cause exon skipping (Lorson, Hahnen et al. 1999; Monani, Lorson et al. 1999). It is generally believed that the rescue of SMN2 exon 7 skipping holds the promise for cure of SMA, due to the deletion of the SMN1 gene. However, the mechanism by which C6U promotes SMN2 exon 7 skipping remains a complex issue (Singh 2007). The single nucleotide substitution C6U was reported to create one or a combination of situations including the abrogation of an enhancer, the creation of a silencer or both events together (Singh 2007). Based on *in vitro* experiments and on ESEfinder predictions, an initial analysis suggested the "loss of function model" (Fig.1.11a) as a potential mechanism (Cartegni and Krainer 2002). According to this model, the presence of the C6U variant in SMN2 exon 7 abrogates an essential enhancer element of the exon associated with the splicing factor ASF/SF2 (Cartegni, Hastings et al. 2006). However, several evidences against the loss of function model came from *in vivo* experiments suggesting the "gain of function model" (Fig.1.11b), as the more probable mechanism (Kashima and Manley 2003). In this case, the C6U substitution creates an exonic splicing silencer (ESS) associated with splicing factor hnRNP A1 and with exon 7 skipping (Kashima and Manley 2003). Strongest evidence in support of hnRNP A1 model came from an RNA interference (RNAi) experiment in which depletion of hnRNP A1 promoted exon 7 inclusion in SMN2 (Kashima, Rao et al. 2007). This classical view of mutations influencing an exon recognition by affecting either an ESE or an ESS in some instances does not fully represent the complexities of the splicing regulation.

For example in the case of neurofibromatosis type-1 gene (NF-1) mutations within exon 37 have been reported as causing both the skipping of exons 36-37 together and the skipping of exon 37 alone. These mutations, localized within an ESE region, interfere with the exon 37 definition but also affect the proper recognition of the upstream exon 36 (Messiaen, Callens et al. 1997; Baralle, Skoko et al. 2006).

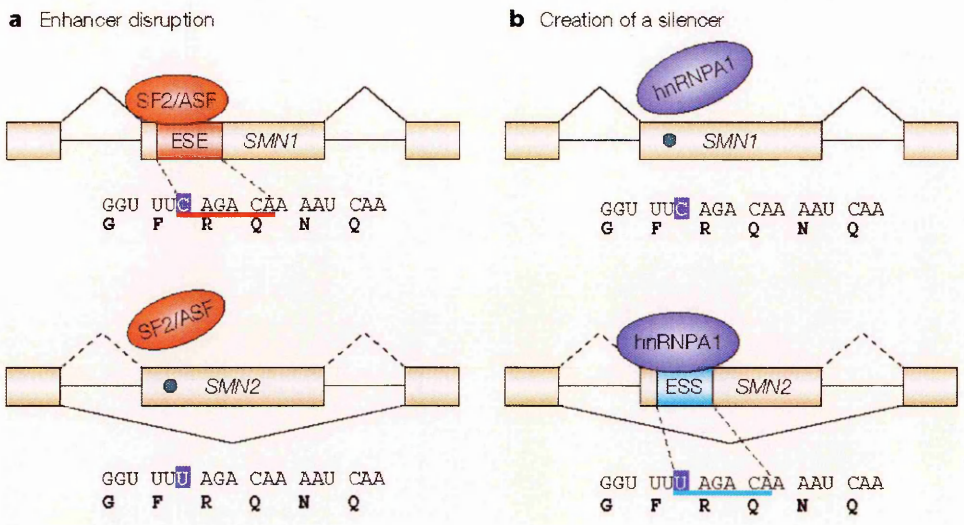

Figure 1.11: Enhancer disruption and creation of silencer models.

SMN1 and SMN2 represent two nearly identical copies of the survival motor neuron gene in humans. Homozygous loss of function of the survival motor neuron 1 (*SMN1*) gene causes spinal muscular atrophy (SMA). SMN2 differs from SMN1 for a translationally silent C→T substitution at +6 position in exon 7 that causes approximately the 80% of exon 7 skipping. The exclusion of this exon from the mRNA produces a truncated unstable protein. Two models have been proposed to explain how this mutation causes exon 7 skipping.

a) In the loss of function model the presence of the single point mutation might cause the inactivation of an exonic splicing enhancer (ESE), located in the 5' region of exon 7, binding site for ASF/SF2 protein.

b) In the gain of function model the C to T substitution creates a new exonic splicing silencer (ESS) able to bind the negative splicing factor hnRNP A1, leading to exon 7 exclusion. Figure taken from Pagani and Baralle (Pagani and Baralle 2004).

In a recent study the ESE region was extensively mapped by point mutation showing that this element is not only required for the definition of exon 37, but also contributes towards the identification of the upstream exon 36. Intriguingly, it has been reported that both these exons need an extremely wide genomic context to obtain the *in vivo* splicing pattern indicating that their correct definition relies on the intrinsic properties of the exons themselves but also on the genomic environment within which they are located (Baralle, Skoko et al. 2006).

As reported for the exon 12 of CFTR gene, the exon inclusion/exclusion regulation can be even more intricate (Pagani, Stuani et al. 2003). Skipping of this exon removes a highly conserved region encoding part of the first nucleotide-binding fold of CFTR producing a non-functional protein (Delaney, Rich et al. 1993).

The systematic study of natural and artificial mutation within the exon 12 sequence allowed the mapping of two particular exonic regulatory elements involved in splicing regulation. Most of the nucleotide substitutions studied in CFTR exon 12 context showed a change in the splicing pattern either increasing exon skipping or promoting exon inclusion (Pagani, Stuani et al. 2003). Whereas exonic regulatory splicing elements found in both constitutive and alternative exons are classically divided into enhancers and silencers (Cartegni, Chew et al. 2002), in the case of CFTR exon 12 the experiments clearly showed an overlapping enhancer and silencer region with neither a pure enhancer nor a silencer behaviour. This result defined a new type of splicing element with a composite regulatory function named CERES (composite exon regulatory element of splicing) (Pagani, Stuani et al. 2003).

The identification of CERES with similar overlapping enhancer and silencer functions in other exons, such as the CFTR exon 9, indicates a common splicing regulatory role of these elements (Pagani, Buratti et al. 2003).

Even if most of the studies have focused on the splice sites and exonic mutations, intronic variants located within approximately 50 bp to the splice sites but also deeply in an intron sequence have been reported to cause splicing defects (Pagani and Baralle 2004). Most commonly, these intronic variants create new splice sites able to define the boundaries of a cryptic exon (Buratti, Baralle et al. 2006). A different mechanism has been proposed in the case of the ATM gene, in which a mutation can affect the splicing efficiency of the intron reported. A four bases deletion in intron 20 in the ATM gene, 2 kb and 0.6 kb from the preceding and following exons, respectively, has led to the discovery of an unknown intronic splicing element involved in the efficient pre-mRNA processing of long introns (Lewandowska, Stuani et al. 2005). The disruption of this intronic processivity element results in the activation of a pseudo exon leading to an erroneous mRNA and consequently to the disease phenotype (Pagani, Buratti et al. 2002).

1.8.3 Secondary structure and mutations.

Changes in RNA secondary structure due to the presence of mutations have been involved in pathogenic processes, as described for the NF-1 gene (Hoffmeyer, Nurnberg et al. 1998; Kaufmann, Leistner et al. 2002), the dystrophin gene (Matsuo, Nishio et al. 1992), and the CFTR gene (Hefferon, Groman et al. 2004). These examples derived from association studies between *in silico* predictions of pre-mRNA secondary structures and splicing activity.

The role of the secondary structure in the generation of human disease has been studied extensively in the case of tau exon 10 mutations (Grover, Houlden et al. 1999; Varani, Hasegawa et al. 1999; Jiang, Cote et al. 2000; Varani, Spillantini et al. 2000). The mutations in the intronic region near the 5'ss of exon 10 strongly

correlate with the disruption of a characteristic stem-loop structure and the stability of this secondary structure determines the ratio of tau exon 10+/- transcripts (Varani, Hasegawa et al. 1999). An increase in the ratio of tau mRNAs containing exon 10 results in the production of an aberrant protein isoform that leads to neuro-degeneration (Grover, Houlden et al. 1999). Extensive mutational analyses (Grover, Houlden et al. 1999) and functional studies to monitor U1 snRNP binding to the splice site have shown that mutations which destabilize the helix result in an increased splice site usage owing to an increase in U1 snRNP binding (Jiang, Cote et al. 2000).

1.8.4 A G to T substitution in BRCA1 exon 18 causes a splicing defect.

In 1998 Mazoyer at al reported the case of a nonsense-mediated exon skipping event in the Breast Cancer 1 gene (BRCA1), in one breast and ovarian cancer family that contains four breast and four ovarian cancer cases (Mazoyer, Puget et al. 1998). By PCR amplification and subsequent agarose gel analysis they identified, in affected family members, a fragment of BRCA1 cDNA covering exons 16–22 able to produce two bands of equal intensity one corresponding to the expected length (596 bp) and one slightly shorter. Direct sequencing of both amplified bands revealed a 78 bp in-frame deletion in the short fragment corresponding to the entire exon 18 exclusion (Mazoyer, Puget et al. 1998). Alternative removal of exon 18 does not occur naturally (Xu, Chambers et al. 1997). Thus, in order to identify the basis of BRCA1 exon 18 skipping, a genomic DNA fragment, encompassing part of intron 17, all of exon 18, and part of intron 19, was amplified by PCR and analyzed for mutations in all the basic splicing consensus sequences: the 3'ss and 5'ss and the predicted branchpoint position. However, the only mutation detected was a G to T substitution

in position +6 of exon 18 (nucleotide 5199, codon 1694), which changes a glutamic acid into a stop codon (Glu1694ter) (Fig. 1.12). Specific cDNA amplification showed that only a wild-type exon 18 sequence was found in all carriers, which implies that the mutant allele produces only mRNA in which exon 18 had been skipped (Mazoyer, Puget et al. 1998). Although the Glu1694ter mutation would be expected to lead to the truncation of 169 amino acids from the BRCA1 protein, it has been showed to cause an in frame deletion of 26 amino acids (Asp1692–Phe1717) due to exon 18 skipping. At the BRCA1 protein level this deletion disrupts the first BRCT domain, specifically removing part of block D, the most highly conserved motif within this domain (Callebaut and Mornon 1997). The block D is organized around a conserved aromatic residue, a Trp that in presence of the G to T mutation is changed to Gly because of the creation of a new junction between exon 17 and 19.

How this nonsense mutation causes exon skipping has been unclear. The ESE disruption was the first proposed mechanism although Glu1694Ter mutation does not occur in a typical purine-rich segment (Liu, Cartegni et al. 2001). BRCA1 exon 18 wild-type and mutant sequences were analyzed by ASF/SF2, SC35, SRp40 and SRp55 motif-scoring matrices (Cartegni, Wang et al. 2003). Multiple high-score motifs of each type are distributed throughout this exon. As reported in figure 1.13 the analysis of the G to T mutated exon at position +6 predicts the specific disruption of the first of three high-score ASF/SF2 motifs.

To study the exon-skipping mechanism, a wild type and mutant minigenes were analyzed for splicing *in vitro*. Whereas the wild-type exon 18 was efficiently included, the G to T mutant exon 18 was totally excluded. The correlation between the first ASF/SF2 high-score motif disruption and the splicing patterns suggested that the nonsense mutation disrupted an ESE (Liu, Cartegni et al. 2001). The study of two artificial BRCA1 exon 18 mutants and four additional high-score heptamers showed

a correlation between ESEfinder prediction and *in vitro* splicing pattern, further indicating the ESE disruption as the putative mechanism (Liu, Cartegni et al. 2001)

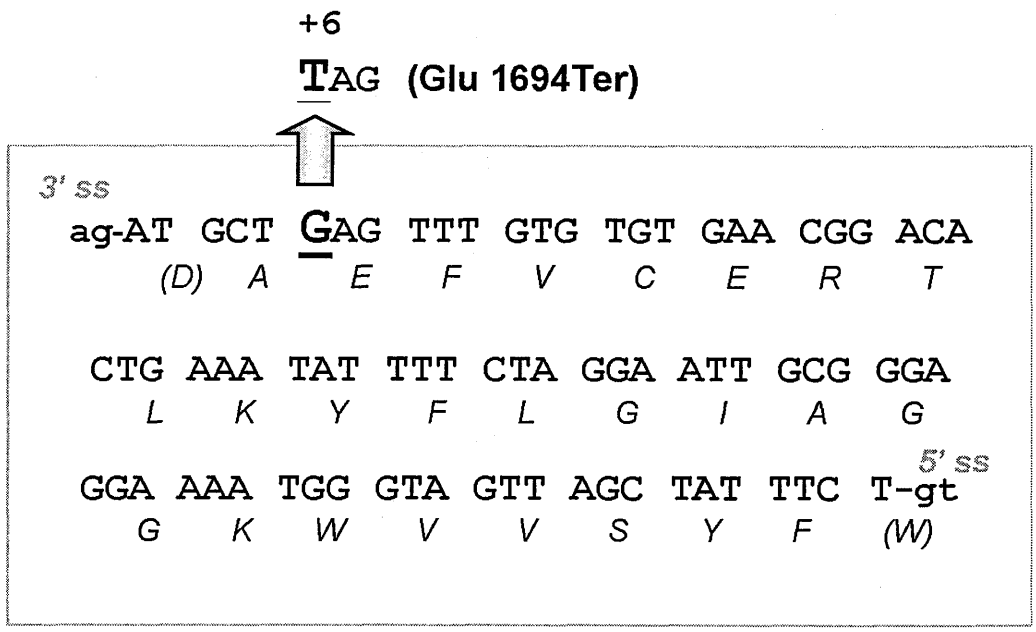

Figure 1.12: BRCA1 exon 18 sequence WT and G6T mutant.

Schematic representation of BRCA1 exon 18 (78pb) sequence (capital letter) flanked by short dinucleotides from the 5' and 3' splice site. The letters below each DNA codon sequence represent the corresponding aminoacid. The arrow indicates the position +6 in which the G to T substitution occurs. The T point mutation creates a stop codon (TAG) and causes the inappropriate skipping of the entire constitutive exon 18 *in vivo*. Exon 18 skipping results in the retention of the same reading frame and the removal of 26 amino acids (Mazoyer, Puget et al. 1998).

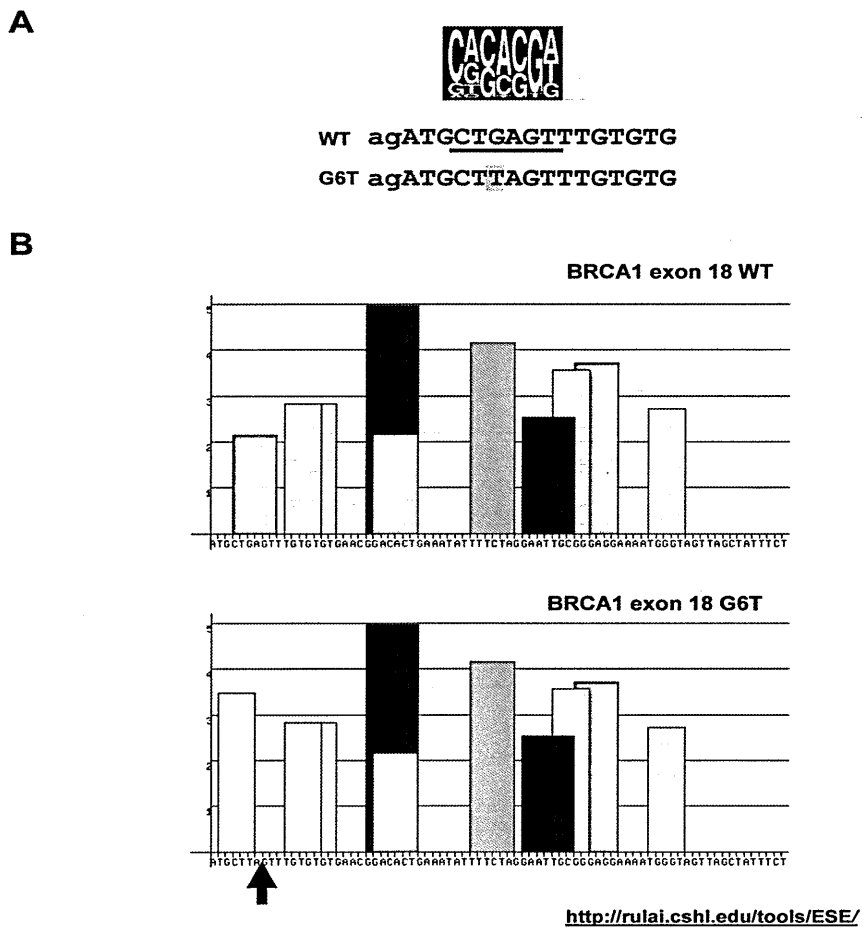

Figure 1.13: High score SR protein motifs in BRCA1 exon 18 WT and G6T mutant.

(A) Effect of point mutation on ASF/SF2 motif score. The graphic representation of ASF/SF2 consensus motif is shown and the height of each letter reflects the frequency of that nucleotide at that particular position. The degenerate ASF/SF2 consensus motif is aligned to the first 14 nt of exon 18 WT and mutated position +6 (grey boxed). The blue horizontal line under the WT sequence indicates the position of the first ASF/SF2 motif recognised within the exon 18 by ESEfinder program. The ASF/SF2 motif scores calculated for the WT and mutated sequence are shown on the right.

(B) SR motif distribution in BRCA1 exon 18 WT and G6T mutant. The 78nt sequence of exon 18 WT (top) and G6T mutated (bottom) were analyzed by ESEfinder program. Multiple high-score motifs of each type are distributed throughout the exon 18, red for ASF/SF2, blue for SC35, green for SRp40 and yellow for SRp55. Exon 18 mutated sequence showed the specific disruption of the first high-score ASF/SF2 motifs (arrow) by the G6T substitution.

1.9 BRCA1 gene.

BRCA1 was the first breast and ovarian cancer susceptibility gene to be identified by linkage analysis of families with multiple cases of early onset breast and ovarian cancer (Miki, Swensen et al. 1994). Indeed, at least half of familial breast cancer cases can be linked to mutations within the BRCA1 gene (Futreal, Liu et al. 1994). Cancer cells isolated from BRCA1-associated tumors show a classic loss-of-heterozygosity pattern at the BRCA1 loci, classifying this gene as a prototypical growth suppressor (Neuhausen and Marshall 1994).

The BRCA1 gene, mapped to chromosome 17q12, is 100 Kb in length, contains 24 exons and encodes a large protein of 1863 amino acids which has been shown to localize predominantly in the nucleus (Chen, Li et al. 1996). BRCA1 does not share any significant structural or sequence homology to any known proteins. There are, however, three main features that are important for its overall function. A RING finger motif, containing a conserved pattern of cysteine and histidine residues, located at the extreme N-terminal end of BRCA1 (Miki, Swensen et al. 1994). This motif is thought to be important for its association with a number of proteins and in particular to BRCA1-associated ring domain 1 (BARD1) (Wu, Wang et al. 1996). BARD1 binds to BRCA1 and forms a heterodimer with ubiquitin–ligase activity (Hashizume, Fukuda et al. 2001). The central region of BRCA1, which constitutes about 60% of the protein, is encoded only by exon 11 and contains two nuclear localization signals (Chen, Li et al. 1996). Several proteins, including DNA repair proteins, tumor suppressors and transcription activator/repressor factors associate directly or indirectly with this part of BRCA1 (Deng and Brodie 2000). In the C-terminal part there are two highly conserved BRCT domains that are also present in a group of proteins involved in DNA damage repair and cell cycle control (Bork, Hofmann et al. 1997). BRCT domains behave as phosphopeptide-binding motifs

ideal for protein targeting (Manke, Lowery et al. 2003; Yu, Chini et al. 2003). C-terminal truncations of up to eight aminoacids can be tolerated by the BRCA1 protein but further deletion results in drastic BRCT folding defects (Williams and Glover 2003).

Numerous studies have reported a role of BRCA1 in transcription regulation, even though it is not itself a sequence specific DNA binding transcription factor (Rosen, Fan et al. 2006). BRCA1 was found to associate with the RNA pol II complex, suggesting its involvement in the basal transcription regulation (Scully, Anderson et al. 1997; Haile and Parvin 1999). BRCA1 is linked to RNA pol II, in part, through the RNA helicase A enzyme (Anderson, Schlegel et al. 1998) and binds preferentially to RNA pol II complexes containing a polyphosphorylated p220 subunit (Krum, Miranda et al. 2003). Indeed, BRCA1 regulates the phosphorylation state of RNA pol II p220 subunit through the CDK-activating kinase (CAK) (Moisan, Larochelle et al. 2004). The precise function of BRCA1 in the RNA pol II complex is unclear. However, BRCA1 associates with BARD1 and exhibits an ubiquitin ligase activity (Hashizume, Fukuda et al. 2001), suggesting that it may mediate ubiquitination and degradation of the RNApol II enzyme components, upon transcriptional blockage in response to DNA damage (Parvin 2001).

BRCA1 was also described as recruiting a SWI/SNF-like chromatin-remodeling complex (Bochar, Wang et al. 2000) and mediating a large-scale chromatin-unfolding activity through the cofactor of BRCA1 (COBRA1) (Narita, Yamaguchi et al. 2003).

Moreover, the interaction of BRCA1 with different sequence-specific DNA binding factors also contributes to stimulating or inhibiting transcription. For example BRCA1 stabilizes p53 and promotes its transcriptional activity (Zhang, Somasundaram et al. 1998), interacts with STAT1 factor involved in the growth

arrest following interferon gamma (IFNγ) treatment (Ouchi, Lee et al. 2000), binds directly and represses the estrogen receptor (ER-α) (Fan, Ma et al. 2001) and regulates many other transcription factors (Rosen, Fan et al. 2006).

BRCA1 has also been implicated in maintaining genomic integrity through its involvement in the DNA damage pathway and in the cell cycle control. BRCA1 has been linked with non-transcriptional and transcriptional aspects of the DNA double-strand breaks (DSB) response pathway. The assembly of a DNA repair complex at sites of DNA damage is an extremely rapid BRCA1-dependent function that does not involve transcription. In response to different forms of DNA damage, BRCA1 can be hyperphosphorylated by ATM (ataxia-telangiectasia mutated), ATR (ATM and Rad3-related), and/or the CHK2 kinase (Ting and Lee 2004), which allows its re-localization from transcriptional to DNA repair complexes (Parvin 2001). When BRCA1 interacts with DNA repair factors it forms a BRCA1-associated genome surveillance complex (BASC) that includes proteins involved in DSB repair, mismatch repair, DNA replication and recombination proteins (Wang, Cortez et al. 2000).

If the DNA damage is too severe, the secondary response phase is activated. BRCA1 is also involved in this second step of DSB pathway producing a delay in cell cycle progression and promoting the transcription of certain DNA repair and checkpoint genes (Mullan, Quinn et al. 2006). Indeed, BRCA1 has also been consistently linked to cell cycle control by producing cell cycle arrest at different phases. Several studies have supported a role for BRCA1 in the activation of p21 leading to G1/S arrest (Li, Chen et al. 1999). BRCA1 is also required for intra-S and G2/M phase cell cycle checkpoints (Xu, Kim et al. 2001), in particular by inhibiting the transcription of cyclinB (MacLachlan, Somasundaram et al. 2000) and stimulating the production of a number of G2/M checkpoint regulatory genes (Yarden, Pardo-Reoyo et al. 2002).

It is evident that BRCA1 is a multifunctional protein that can suppress tumor formation through numerous and important cellular mechanisms such as DNA repair pathway, cell cycle control and transcription regulation (Yoshida and Miki 2004). Even if all these diverse roles associated with BRCA1 are so general the loss of BRCA1 function leads to breast and ovarian cancer (Narod and Foulkes 2004). An aberrant transcriptional control by BRCA1 remains the principal explanation for the tissue specificity of these tumors. BRCA1 is transcriptionally associated with the repression of estrogen receptor (ERa) and its loss could lead to an increase in ERa levels, which in turn stimulates cellular estrogen metabolism (Yoshida and Miki 2004). Further studies will provide a better understanding of the tumorigenic pathway involving the loss of BRCA1 and identify targets for treatment of breast and ovarian cancers.

2 AIM OF THE PROJECT

Numerous studies have indicated that exonic mutations independently from their predicted effect on the aminoacid sequence can affect pre-mRNA splicing cis-acting regulatory sequences located within the same exon. As a result, every nucleotide variation must to be considered as a potential candidate to cause splicing alterations even if classified as missense, nonsense or translationally-silent (Cartegni, Chew et al. 2002; Pagani and Baralle 2004).

The aim of this thesis is to increase the present knowledge about the role of exonic sequences in splicing regulation and, moreover provide an explanation of the mechanism of exon skipping caused by exonic mutations.

The BRCA1 exon 18 system was chosen as a model in which a naturally occurring G to T mutation, at position +6, induces exon skipping. This research project was particularly focused on the role of ASF/SF2 in regulating normal and pathological BRCA1 exon 18 splicing and the characterization of cis-acting elements and the trans-acting factors involved.

3 RESULTS

3.1 Regulatory role of ASF/SF2 on WT and T6 mutant BRCA1 exon 18 splicing.

3.1.1 Analysis of ASF/SF2 binding specificity against WT and U6 mutant synthetic RNAs.

The exonic sequence of BRCA1 exon 18 from position +4 to +10 has been previously suggested to be an ESE that interacts with the ASF/SF2 splicing factor (Liu, Cartegni et al. 2001). Moreover, the natural G6 to T mutation was predicted to abrogate this ESE and consequently ASF/SF2 binding. However, this evidence was based mainly on a computer-assisted program (ESEfinder) and the *in vitro* analysis of a limited number of mutants (Liu, Cartegni et al. 2001).

To test directly the interaction between exon 18 and ASF/SF2 and evaluate the role of this splicing factor in the splicing regulation, a classical pull down assay was performed using as a targets two synthetic RNA oligonucleotides (20 nt) corresponding to the first part of exon 18 WT and U6 (Fig. 3.1.1A). These two RNA were covalently linked via their 3' end to agarose beads and incubated in a HeLa nuclear extract. Proteins that remained tightly bound to RNA, after the washing steps, were separated on a SDS-PAGE gel, transferred to PVDF filters and probed for the presence of ASF/SF2 with a specific antibody (Fig. 3.1.1B). The results in figure 3.1.1 show that WT and U6 mutant RNAs bind to ASF/SF2 with a similar efficiency. This experiment was performed three times and confirmed the absence of any significant difference in the binding affinity between the two RNAs. The data obtained by this *in vitro* binding assay showed that despite the presence of the mutation, ASF/SF2 binds to both exon 18 WT and U6 sequences with similar efficiency (Fig. 3.1.1B, lanes 2-3).

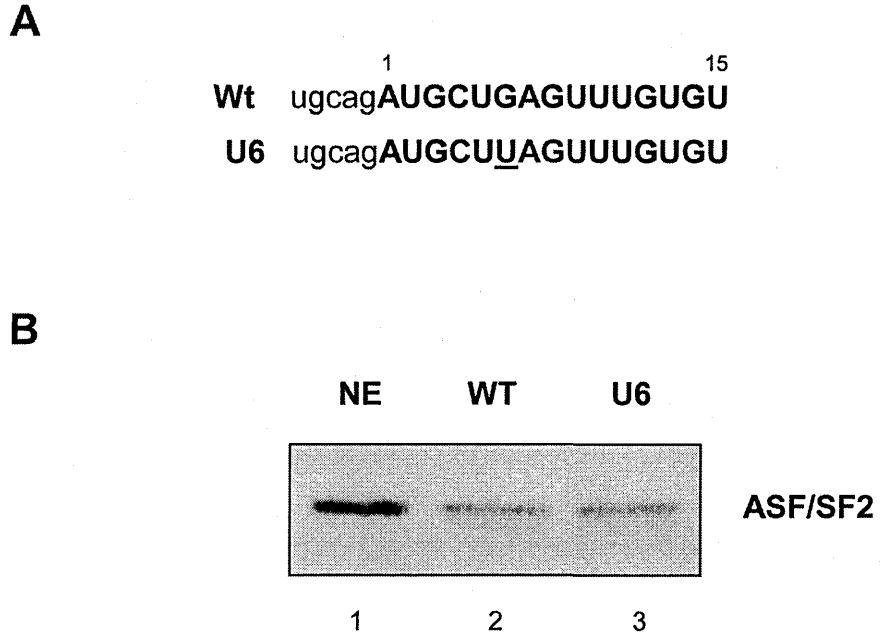

Figure 3.1.1: Western blot of pulldown analysis of BRCA1 exon 18 WT and U6 RNAs to determine the binding of ASF/SF2.

(A) Nucleotide sequences, WT and U6, of BRCA1 exon 18 synthetic RNAs. Exonic and intronic sequences are shown in upper and lower case, respectively, and the mutated nucleotide U in position +6 is underlined.

(B) Western blot analysis of ASF/SF2. A pulldown experiment of BRCA1 exon 18 WT and U6 RNAs was followed by western blot using an antibody against ASF/SF2. Both RNAs were able to pulldown ASF/SF2 equally (lane 2 and 3). The nuclear extract sample (NE) corresponds to 1/20th of the amount used for the pull down assay. The pulldown picture is representative of three independent experiments. This pulldown analysis contains an intrinsic limitation due to the difficult normalization of the amount of RNAs among the different samples. This limitation can affect a precise quantification of the small differences in ASF/SF2 binding between the two samples (see the text).

3.1.2 Analysis of ASF/SF2 binding specificity against WT and U6 mutant and developing a new pulldown strategy.

The classical pulldown method used to test the binding capability of ASF/SF2 factor contains an intrinsic limitation due to the difficult normalization of the amount of RNAs among the different samples. In fact, the binding efficiency might be affected by erroneous quantification of the input RNA, to a different amount of oligonucleotides attached to the beads and/or to a higher level of degradation in one of the samples. All together these limitations can affect a precise quantification of the data, raising the possibilities that small differences in the binding efficiency among the samples cannot be detected.

In order to verify whether ASF/SF2 binds equally to the two RNAs, as observed (Fig. 3.1.1B, lanes 2-3), the original pull down methodology was improved by including an "internal control" for the amount of RNA used during the experiment. In this case, the progress made on the pulldown technique was the possibility to normalise the amount of RNA that remains linked to the agarose beads. The two target sequences, WT and T6 BRCA1 exon 18, were linked to a UG motif already described in literature to be bound efficiently by the specific protein TDP43 (Buratti, Dork et al. 2001). This protein contains two RNA recognition motif (RRM) domains with distinct RNA/DNA binding characteristics and has a high affinity for single-stranded UG stretches, starting from a minimum number of six UG (or TG), and increasing its affinity together with the number of repeats (Buratti and Baralle 2001). As reported in detail in figure 3.1.2A, we transcribed *in vitro* a RNA that comprise: the first part of BRCA1 exon 18 WT or U6, a $(U)_5$ spacer and a sequence composed by six UG dinucleotide repeats. The two transcribed RNAs were linked to agarose beads via 3' end and incubated in a HeLa nuclear extract.

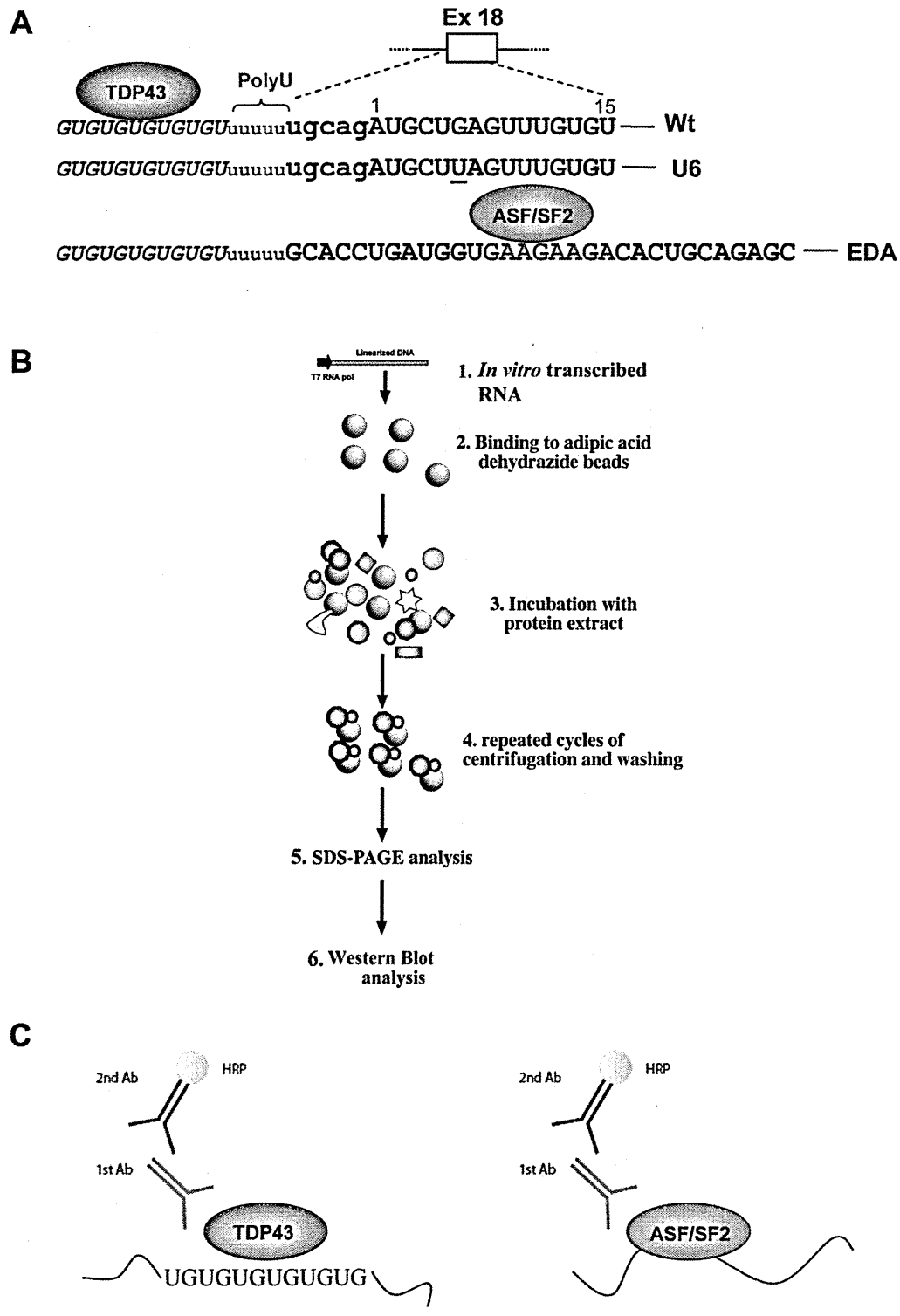

Figure 3.1.2: Modified pulldown analysis (see the chapter 3.1.2)

Proteins that remained tightly bound to RNA after the washing steps were separated on SDS-PAGE gel, transferred to PVDF filters and probed for the presence of the target protein with specific antibodies (Fig. 3.1.2B). During the pulldown assay, the UG repeats, present in each RNA, are specifically recognised by TDP43. By performing western blot against this protein, the pulldown was normalized for the RNA amount. Subsequently, the membrane was probed for ASF/SF2 (Fig. 3.1.2C). This new strategy gave the opportunity of precisely comparing the different samples analysed to the TDP43 standard binding.

As reported in figure 3.1.3, the pull-down assay was followed by Western blot probed with ASF/SF2 and TDP34 antibodies. Relative to the signal of TDP43 the WT and U6 mutant RNAs bound to ASF/SF2 with a similar efficiency (Fig. 3.1.3B, lanes 2-3). Moreover, including an ASF/SF2 purine-rich enhancer as a positive control allowed the binding capability of the BRCA1 sequences to be estimated. In this case, the UG motif was linked to a previously described enhancer sequence located in the alternatively spliced fibronectin EDA exon (Lavigueur, La Branche et al. 1993; Caputi, Casari et al. 1994). This positive control clearly indicated a lower binding capability of ASF/SF2 for both exon 18 WT and U6 sequences (Fig. 3.1.3B, lane 1 vs. lanes 2-3). These results are in accordance with the data obtained with the classical pull down carried out with synthetic RNAs and suggest that the U6 mutation does not disrupt an ASF/SF2 binding site. Both sequences can bind to ASF/SF2 with lower efficiency in comparison to a classical ASF/SF2 enhancer.

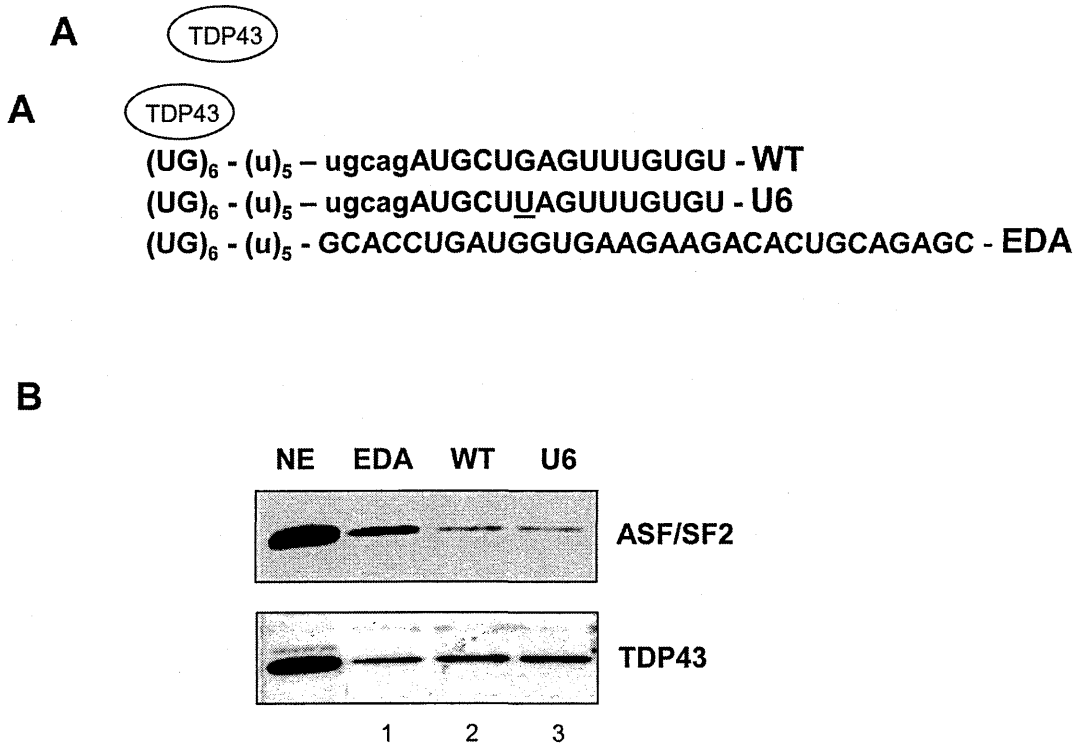

Figure 3.1.3: Western blot of pull-down analysis anti-ASF/SF2 normalized for TDP43 binding.

(A) Schematic representation of the *in vitro* transcribed RNA sequences that contain $(UG)_6$ repeats, $(U)_5$ spacer, part of intron 17 (lower case) and the first nucleotides of exon 18 WT and U6 mutant (upper case). The substituted nucleotide is underlined. The bottom lane shows the EDA sequence, linked to the UGs and Us repeats, used as positive control for ASF/SF2 binding. (B) Binding analysis of ASF/SF2 protein. After *in vitro* transcription, RNA was pulled down and Western blot analysis was carried out using ASF/SF2 and TDP43 antibodies. TDP43, in this case, was used to normalize the assay for the amount of RNA. Both WT and U6 RNAs were able to pulldown ASF/SF2 protein although with less efficiency if compared with the EDA positive control. The nuclear extract sample (NE) corresponds to 1/20th of the amount used for the pull down assay. The pulldown picture is representative of three independent experiments. As the picture shows, both WT and U6 RNA are able to pull down ASF/SF2 protein. The WT and U6 bind approximately only 10% of ASF/SF2, compared to the EDA sample and normalized for TDP43. The binding was determined by the evaluation of the relative intensity of the bands.

3.1.3 Evaluation of WT and T6 mutant BRCA1 exon 18 splicing efficiency by hybrid minigenes.

The G to T substitution at the exonic position +6 was previously reported to induce a significant BRCA1 exon 18 skipping in patients derived cells (Mazoyer, Puget et al. 1998). A functional splicing analysis was carried out in order to evaluate *in vivo* the effect of this mutation and to study in detail its disease-causing mechanism. For this purpose, I prepared two hybrid minigenes (Fig. 3.1.4). The hybrid minigene (pTB), previously used in other studies focused on aberrant splicing (Muro, Caputi et al. 1999; Pagani, Buratti et al. 2002; Pagani, Stuani et al. 2003), was the first system chosed to investigate the regulation of BRCA1 exon 18 splicing. The pTB hybrid minigene was generated by cloning the genomic DNA fragment of exon 18 (78 bp) including part of the flanking intron 17 (292 bp) and intron 18 (240 bp) into a construct consisting of the α-globin and fibronectin intron-exon sequences (Fig. 3.1.4A). In the second minigene developed, pBRA, exon 18 was placed in a more physiological context. The amplified region from BRCA1 gene also included introns 17 (the size was reduced from 3655 bp to 1857 bp to facilitate the cloning step), the entire intron 18 and parts of exon 17 and 19 were cloned in the context of α-globin three exons minigenes (Fig. 3.1.4B). Furthermore, to facilitate the subsequent mutagenesis analysis, a unique restriction EcoRI site was introduced within the intron 18 sequence of both minigenes (see materials and methods).

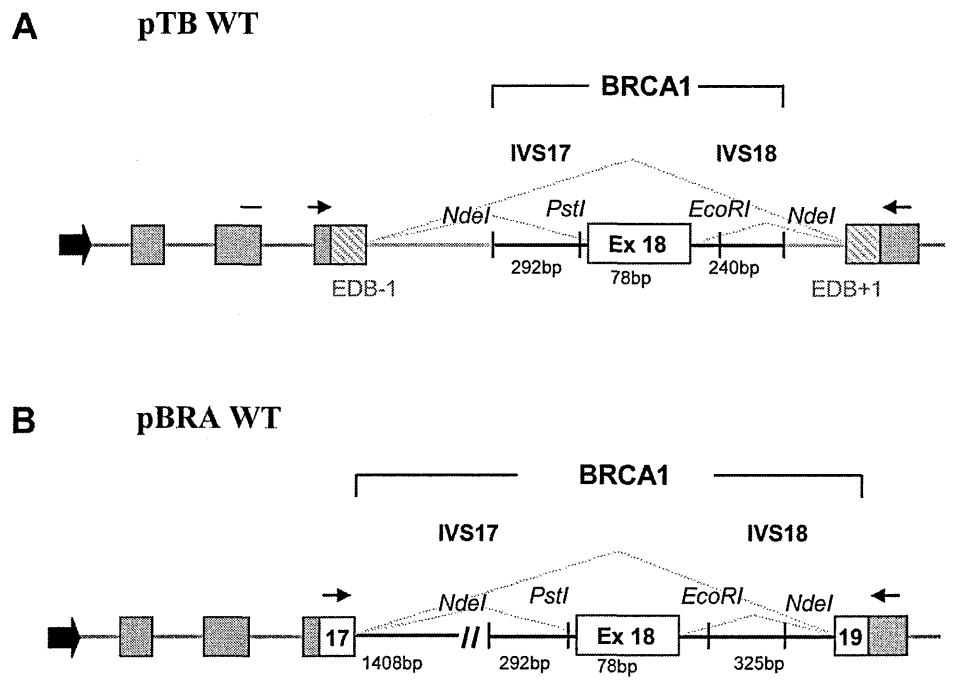

Figure 3.1.4: Schematic representation of the hybrid minigene systems.

(A) Schematic representation of the pTB hybrid minigene system. Dark grey, shaded and white boxes respectively represent α-globin, fibronectin EDB exons and human BRCA1 exon 18, with introns shown as thick lines. Exon 18 and parts of its flanking introns were cloned in the pTB plasmid using the restriction enzymes NdeI. Transcription is under the control of the minimal α-globin promoter and SV40 enhancer (big arrow at the 5'-end). The small black superimposed arrows represent the position of the primers used for the RT-PCR analysis after transient transfection of the minigene. The grey dotted lines show exon 18 splicing, inclusion and exclusion forms of mRNA. The position of unique PstI and EcoRI restriction sites used for subsequent cloning is shown.

(B) Schematic representation of the pBRA hybrid minigene system. Dark grey and white boxes represent α-globin and BRCA1 exons, respectively. In this case EDB exons where replaced by part of exon 17 and exon 19 with flanking intronic sequences (IVS). The length of BRCA1 sequences is shown; the size of intron 17 was reduced from 3655 nt to 1857 nt to facilitate the cloning procedure. The large black arrow at the 5' end indicates the minimal α-globin promoter and SV40 enhancer. And the two small arrows show the primers position used for RT–PCR. The two alternative splicing possibilities, exon 18 inclusion and exclusion, are indicated by dotted lines. The position of the unique PstI and EcoRI restriction sites used for subsequent cloning are reported.

As a first approach, both the pTB and pBRA systems were tested for the ability to reproduce the exon 18 WT and the T6 splicing pattern, as described in patients cells (Mazoyer, Puget et al. 1998). By site-directed mutagenesis a T substitution in position +6 of exon 18 was inserted in both pTB and pBRA minigenes (Fig. 3.1.5A). Transient transfection of WT and mutated constructs was performed in Hep3B cells and the transcrips derived were analysed by RT–PCR amplification with specific primers. The primers ALFA 2-3 and BRA 2 were used for the pTB while the BRC 90 and GLO 800 were specific for the pBRA transcript. The PCR products showed the presence of a unique band of 336 bp for the pTB WT minigene (Fig. 3.1.5B, lane 1) and a band of 244 bp for pBRA WT construct (Fig. 3.1.5C, lane 1), both corresponding to the fully included exon 18 in the mature transcript. This experiment was performed in triplicate and the subsequent sequence of the band confirmed the full BRCA1 exon 18 inclusion.

As expected, no transcript corresponding exon 18 skipping was detectable in the WT splicing profile and no endogenous-derived BRCA1 PCR products were observed.

In the case of both minigenes carrying the T6 substitution two amplified bands were generated. The sequence of the purified PCR products showed that the upper band corresponded to BRCA1 exon 18 full included and the lower band to the complete exon exclusion from the mRNA. In Hep3B the pTB T6 minigene showed 50 % of exon 18 inclusion (Fig. 3.1.5B, lane 2) whereas the T6 mutated exon 18 placed in the pBRA context was predominantly skipped, with only 20% of it being included (Fig. 3.1.5C, lane 2). Both minigenes were able to mimic the exon skipping effect of the T6 mutant described *in vivo*. Experiments were then performed on pBRA as it contains the more physiological context and showed slightly higher exon skipping.

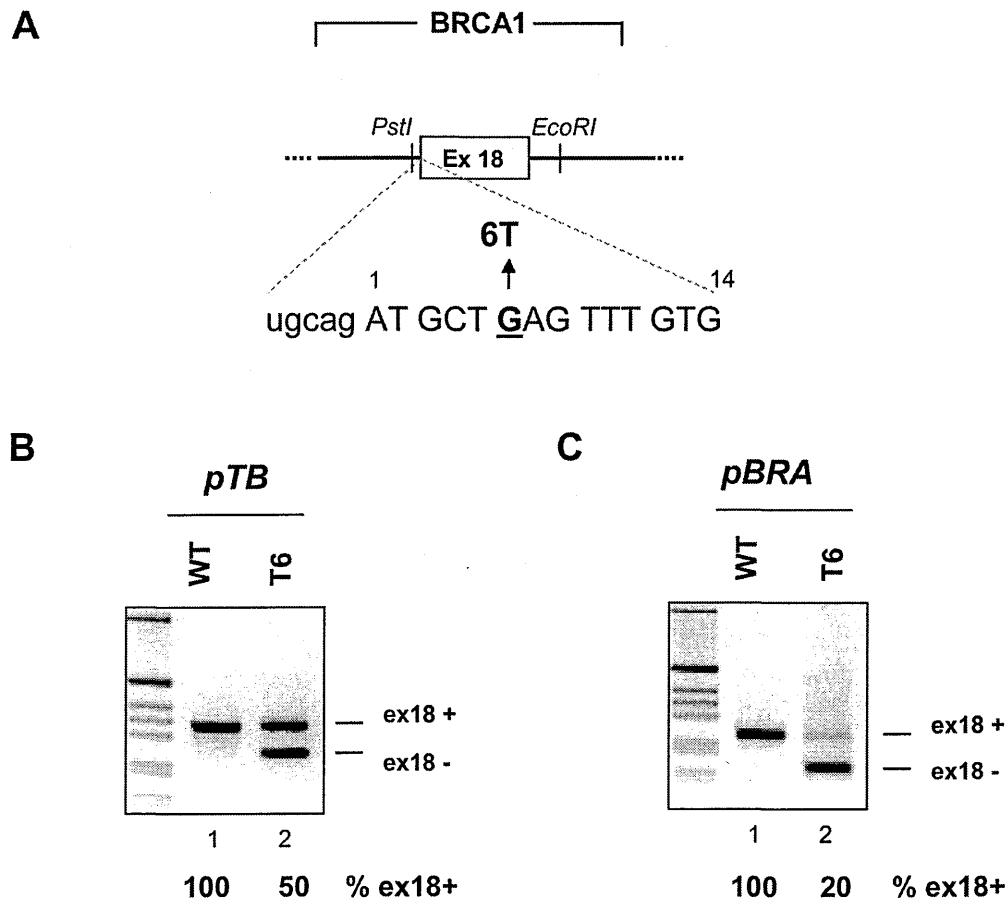

Figure 3.1.5: BRCA1 exon 18 +6 G>T is a splicing mutation.

(A) Schematic representation of the exon 18. In the lower part of the picture, a few bases from intron 17 (lower case) and exon 18 (from position +1 to +14) are shown, divided in codon triplets (upper case). The G to T natural mutation (E1694X) in position +6 is indicated (Mazoyer, Puget et al. 1998).

(B) (C) Splicing pattern analyses of the RT–PCR products derived from RNA of transfected cells, separated on a 2% agarose gel. The T6 mutation affects the exon 18 inclusion in both minigenes context. Two products are seen on agarose electrophoresis, with the lower band representing a product lacking exon 18 and the upper band representing a product including exon 18, as indicated schematically on the right of both gels. The WT exon 18 is fully included in each construct (lanes 1B and 1C). The T6 exon 18 is excluded in both minigenes in different amount depending on the surrounding minigene context (lanes 2B and 2C). The percentages of exon 18 inclusion, obtained from three independent transfections, are indicated below each lane number.

3.1.4 Effect of ASF/SF2 in vivo overexpression on BRCA1 exon 18 WT and T6 splicing.

The *in vitro* pulldown analysis showed that ASF/SF2 binding is not affected by the presence of +6 G to T mutation within BRCA1 exon 18 (Fig. 3.1.1 and 3.1.3). On the other hand, accordingly to the ESEfinder prediction, the presence of the T6 mutation abrogates the ASF/SF2 binding site within the BRCA 1 exon 18 sequence. Thus, we decided to investigate the effect of ASF/SF2 on the BRCA1 exon 18 splicing regulation in co-transfection experiments.

The ASF/SF2 coding plasmid (+), or the empty vector (-) as a control, were transfected together with the pBRA minigenes, WT or T6 mutant. Total RNA was extracted and a RT-PCR was carried out with a specific primer set and the amplified products were analysed by electrophoresis on agarose gel (Fig. 3.1.6). Surprisingly, overexpression of ASF/SF2 in cells transfected with the pBRA T6 plasmid showed an increase in the percentage of exon 18 inclusion. The experiment was repeated three times and the the percentage of exon 18 inclusion increased from 20% up to 55% in the presence of ASF/SF2 (Fig. 3.1.6, lanes 3-4). The cotransfection of the empty expression vector did not affect the splicing of pBRA T6 trascripts which showed the same profile of exon 18 skipping (Fig. 3.1.6, lane 3). A positive effect was not detectable for the exon 18 WT minigene that is already fully included in the absence of ASF/SF2 overexpression (Fig. 3.1.6, lanes 1-2). In conclusion, cotransfection of ASF/SF2 induced a significant increase in the percentage of exon 18 inclusion in the T6 mutant context, whereas no effect was detected for the WT.

A

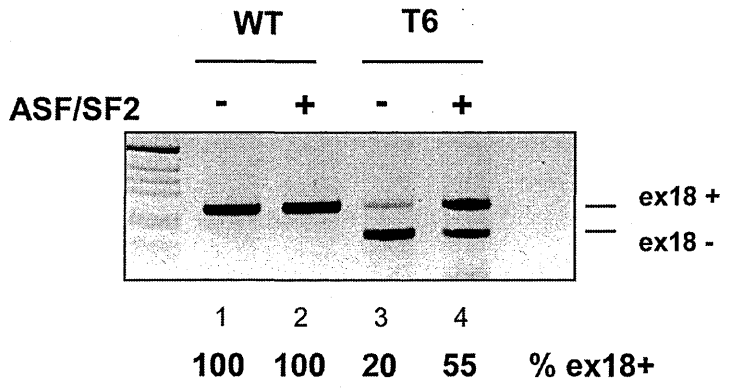

Figure 3.1.6: *In vivo* over-expression of ASF/SF2 promotes T6 exon 18 inclusions.

(A) HeLa cells were transiently transfected with WT or T6 pBRA minigenes as indicated, along with ASF/SF2-coding plasmid (+), or with the empty vector (-). Splicing patterns were analysed by separating the RT-PCR products on agarose gel. The identity of the transcripts, exon 18 inclusion (ex18+) and exon 18 exclusion (ex18-), is shown schematically at the right side of the gel. The percentages of exon 18 inclusion, obtained from three independent transfections, are reported below each lane. Although quantification of ASF/SF2 was not performed in this specific experiment, the ASF/SF2- coding plasmid was previously used in other similar over-expression assays (Caceres, Stamm et al. 1994; Pagani, Buratti et al. 2003; Pagani, Stuani et al. 2003).

3.1.5 Effect of in vivo siRNA mediated depletion of ASF/SF2 on BRCA1 exon 18 WT and T6 splicing

In the co-transfection experiment the T6 mutated exon 18 exclusion was partially recovered by ASF/SF2 overexpression, suggesting a possible role of this protein in the T6 splicing regulation. If this were the case, the specific knock down of ASF/SF2 would be expected to increase the ratio of the mRNA species that skip exon 18 carrying the T6 mutation.

The *in vivo* depletion of ASF/SF2 was obtained by transient siRNA treatment of HeLa cells. In order to test whether the siRNA effectively reduces the ASF/SF2 expression we checked the siRNA oligonucleotides effect on protein level. The experiments were performed in HeLa cells because they showed an efficient ASF/SF2 siRNA-mediated depletion. After this treatment, cell lysates were collected and analyzed by western blot assay for ASF/SF2 protein expression. In cells transfected with siRNA(+), but not cells treated with luciferase siRNA (-), we observed a strong reduction in the ASF/SF2 levels. The expression of a housekeeping gene (tubulin) was also checked by western blot, as control (Fig. 3.1.7A, lanes 1-2).

Once verified the effective depletion of ASF/SF2, the WT and T6 pBRA constructs were transfected in siRNA treated cells. After 48 h cells were collected and divided in two samples: one for detect the level of endogenous ASF/SF2 by western blot and the second for the analyses of the splicing pattern by RT-PCR. The RNA analysis showed that ASF/SF2 depletion induced a nearly complete exclusion of T6 exon 18 from the mature mRNA. The percentage of exon 18 inclusion decreased from 20% in the case of the control to quite total exon exclusion in ASF/SF2 depleted cells (Fig. 3.1.7B, lanes 3-4). On the contrary, WT BRCA1 exon 18 did not respond to ASF/SF2 depletion, showing complete exon inclusion (Fig. 3.1.7B, lanes 1-2). As

internal control for ASF/SF2 depletion, the endogenous RNA of Ron gene was analysed by RT-PCR. It was previously reported that the Ron exon 11 splicing is regulated by the ASF/SF2 factor. Overexpression of ASF/SF2 induces skipping of exon 11 promoting the production of the Delta Ron protein, a constitutive active isoform of the tyrosine kinase receptor for the macrophage stimulating protein. Vice versa, low levels of ASF/SF2 obtained by siRNA experiments are accompanied by the inclusion of exon 11, which leads to the production of the full-length Ron mRNA (Ghigna, Giordano et al. 2005).The RT-PCR performed as a control on the endogenous RON showed an increase in exon inclusion from 15% to 45% (Fig. 3.1.7C, lane 1-2), confirming that the removal of ASF/SF2 protein from HeLa cells affected the splicing of an endogenous transcript.

The *in vivo* analysis results, together with the pull down data, do not support the idea of ASF/SF2 binding disruption due to the T6 mutation within exon 18.

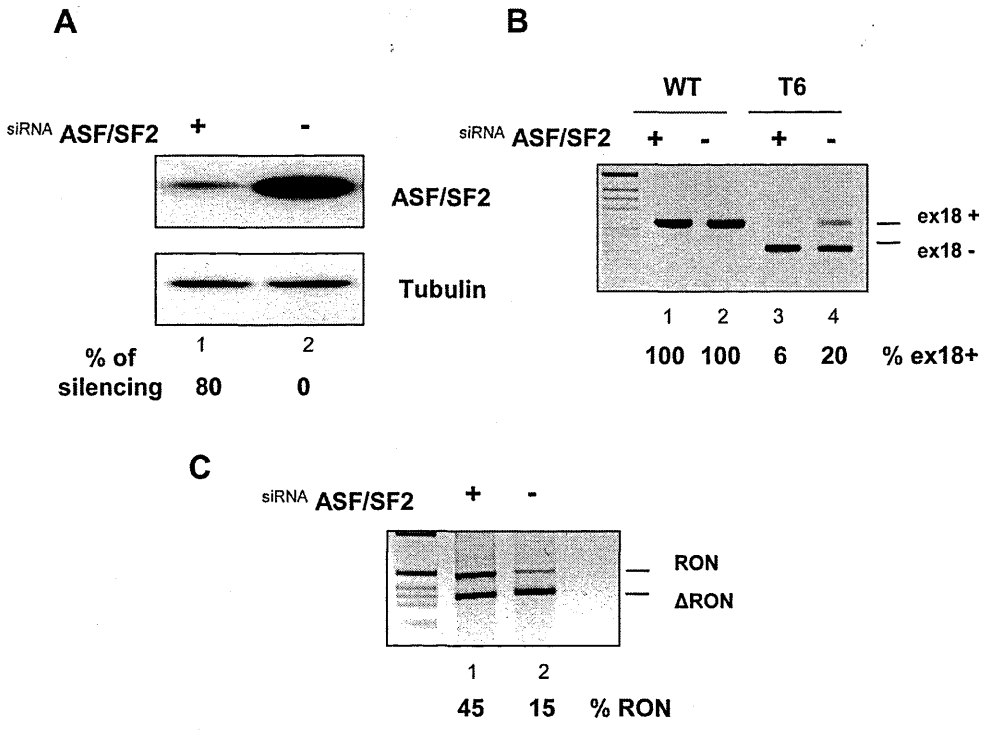

Figure 3.1.7: *In vivo* depletion of ASF/SF2 induces T6 exon 18 skipping.

(A) The upper panel shows the western blot analysis of protein extracts from HeLa cells treated with either ASF/SF2 siRNA (+) or a control siRNALuc (-). The normalization of the protein loaded was checked using an antibody against tubulin, as reported in the lower panel. The siRNA experiment is representative of at least three independent experiments. (B) The gel picture shows the RT–PCR products after transfection of the WT or T6 pBRA plasmid minigenes into ASF/SF2 siRNA-treated (+) and luciferase control-treated (-) HeLa cells. On the right of the gel is reported the mRNA processing outcomes. An increase of T6 exon 18 skipping was observed in the siRNA treated cells (lane 3). The percentage of BRCA1 exon 18 inclusion is indicated below each lane. (C) RT-PCR analysis performed on endogenous RON as internal control for the siRNA treatment. Two splicing products were seen after the amplification of the RNA: the upper band includes RON exon while the lower band ΔRON that lacks this exon. Note that the removal of ASF/SF2 rescues partially the RON inclusion (lane 1) as previously reported (Ghigna, Giordano et al. 2005).

3.2 Mapping of the splicing regulatory element within BRCA1 exon 18.

3.2.1 Single point mutation analysis of the exon 18 regulatory element.

To assess the relationship between the exonic regulatory regions and exon 18 skipping, I investigated the splicing contribution of the cis-acting elements surrounding the position +6. If the region analyzed is a pure enhancer element, the majority of the single substitution at this element is expected to destroy its function and leads to exon 18 skipping. For this purpose a large panel of pBRA minigenes, with a series of single point mutations from position +4 to +11, was constructed by PCR with specific primers on pBRA WT as template (Fig. 3.2.1A). All the substitutions were introduced within the region that matches with the putative ASF/SF2 enhancer binding site predicted by the ESEfinder program, from position +4 to +10 (Liu, Cartegni et al. 2001). A total number of 21 minigenes was generated substituting each nucleotide position with the three possible alternatives. The minigenes were transfected into Hep3B cells, analyzed by RT-PCR and the splicing pattern evaluated on agarose gel. As a control pBRA WT minigene was also included in the experiment (Fig. 3.2.1B).

Interestingly, out of 21 substitutions only the natural mutant T6 resulted in a very low level of exon 18 inclusion of about 25 % (Fig. 3.2.1B, lane 10).

Two additional single substitutions in position +4 (T4) and +6 (A6) partially affected the splicing pattern, causing 90 % and 80% of exon 18 inclusion, respectively (Fig. 3.2.1B, lanes 4, 8). The majority of the mutants (18 out of 21) showed the same pattern of the WT construct with complete exon inclusion.

Thus, the systematic site-direct mutagenesis showed a very few number of exon 18 skipping mutants, suggesting that the BRCA1 exon 18 region surrounding the +6 position does not behave like an exonic splicing enhancer.

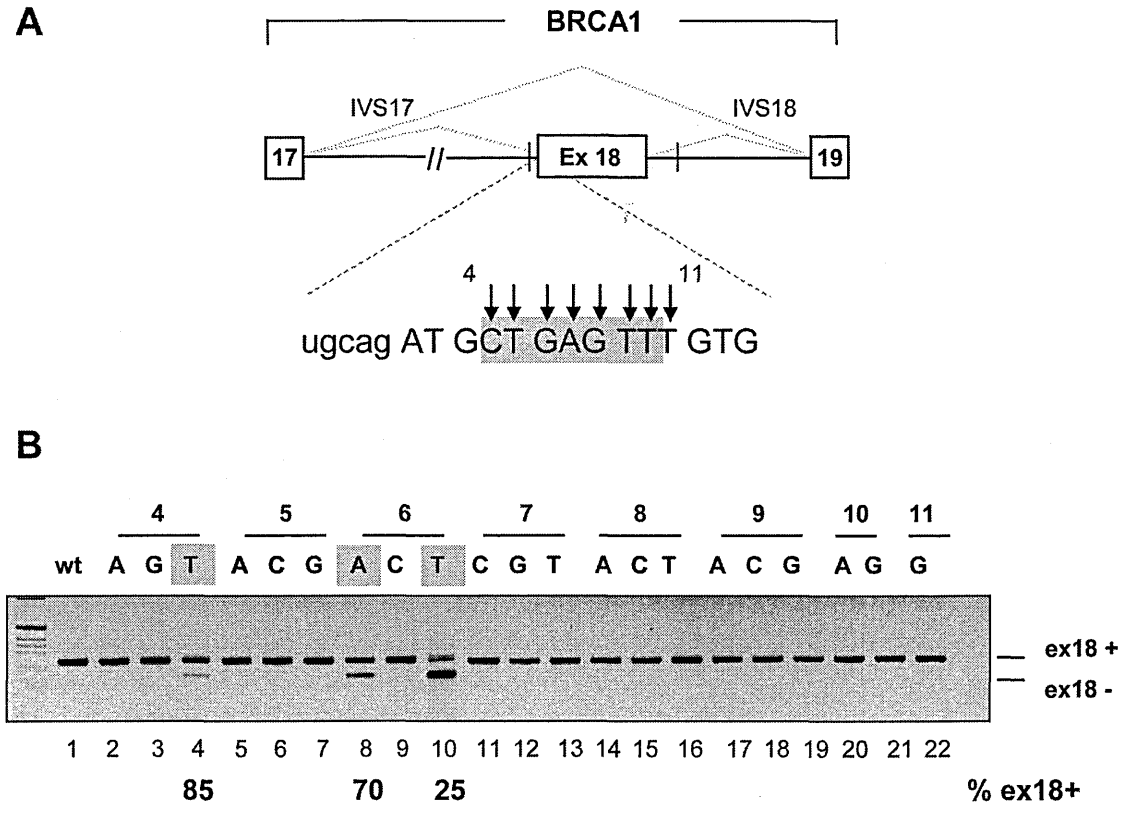

Figure 3.2.1: Single point mutation analysis of the exon 18 putative regulatory region.

(A) Schematic representation of the BRCA1 exons and introns cloned in the pBRA minigene. Part of intron 17 and BRCA1 exon 18 sequences are shown in lower and upper case, respectively. The predicted Exonic Splicing Enhancer (Liu, Cartegni et al. 2001), putative binding site for ASF/SF2, is boxed in light grey. The black arrows, from position +4 to +11, indicate the nucleotide where the all three possible single point mutations were introduced.

(B) Splicing pattern analysis. The agarose gel shows the RT-PCR products of WT and mutants minigenes after transient transfection into Hep3B cell line. The position of the two splicing forms corresponding to exon 18 inclusion (ex18+) and exclusion (ex18-) are indicated on the right of the picture. The grey boxes highlight the mutated nucleotides that induced exon skipping (lane 4, 8 and 10) and percentage of exon 18 inclusion is indicated below each lane. All the other mutants analysed that do not affect the proper exon 18 recognition are considered 100%.

3.2.2 Double point mutation analysis of the exon 18 splicing regulatory element.

The analysis with the single site-directed mutagenesis was more compatible with the creation of a negative element by the T6 substitution rather than the abrogation of a pre-existing enhancer sequence. Previous studies of another splicing mutation, described in the SMN2 exon 7 (Kashima and Manley 2003) and Dystrophin exon 31 (Disset, Bourgeois et al. 2006), demonstrated that the mechanism of exon skipping could be due to the creation of a silencer element. In the BRCA1 exon 18 model, the only mutation able to produce a high percentage of exon skipping was the T6 natural mutation, suggesting a negative effect of this substitution on the correct exon definition.

In order to map this putative silencer element within BRCA1 exon 18, a series of point mutations from position +4 to +9 were introduced in the context of T6 substitution (Fig. 3.2.2A). The nucleotides substitutions were inserted by site-directed mutagenesis by PCR on pBRA T6 as template. The resulting minigenes were transiently transfected into Hep3B cells, and analysed by RT-PCR. The splicing profile of each double mutant was compared to the pattern of the pBRA T6 plasmid. As shown in figure 3.2.2B, the double mutants' analysis resulted in a variable splicing pattern. The majority of the double site-directed variants increased the splicing efficiency compared to the single T6 minigene. In particular, all the substitutions in position +7 and +8 restored almost completely the exon 18 inclusion (Fig. 3.2.2B, lanes 9-14). Also the three mutants in position +5 and two in position +9 (T6A9 and T6C9) increased partially the splicing of exon 18, showing a percentage of exon inclusion of an average of 60% (Fig. 3.2.2B, lanes 6-8, 15 and 16). The only double mutant that did not show any effect was the T6G9, with a splicing pattern comparable with the single mutant T6, with only 20% of exon

inclusion (Fig. 3.2.2B, lane 17). On the contrary, the substitutions at positions +4 inserted in the context of the T6 natural variant resulted in a complete exon 18 exclusion (Fig. 3.2.2B, lanes 3-5). This systematic double point mutation analysis strongly suggested that the naturally occurring mutation T6 creates a sequence with splicing inhibitory properties. In particular, we noticed that the presence of a TAG sequence in position +6 to +8 is essential for splicing inhibition. All mutations within this sequence restored exon 18 inclusion, supporting the hypothesis that this triplet plays an inhibitory role. Thus, the TAG at position +6 therefore appears to be "the core" silencer element involved. The function of TAG motif in splicing silencing has been previously related to hnRNP A1 (Burd and Dreyfuss 1994; Del Gatto, Gesnel et al. 1996). The above results showed that the most of the tested mutations promote exon 18 inclusion suggesting that the BRCA1 exon 18 splicing, carrying the T6 mutation, is inhibited because of a silencer element.

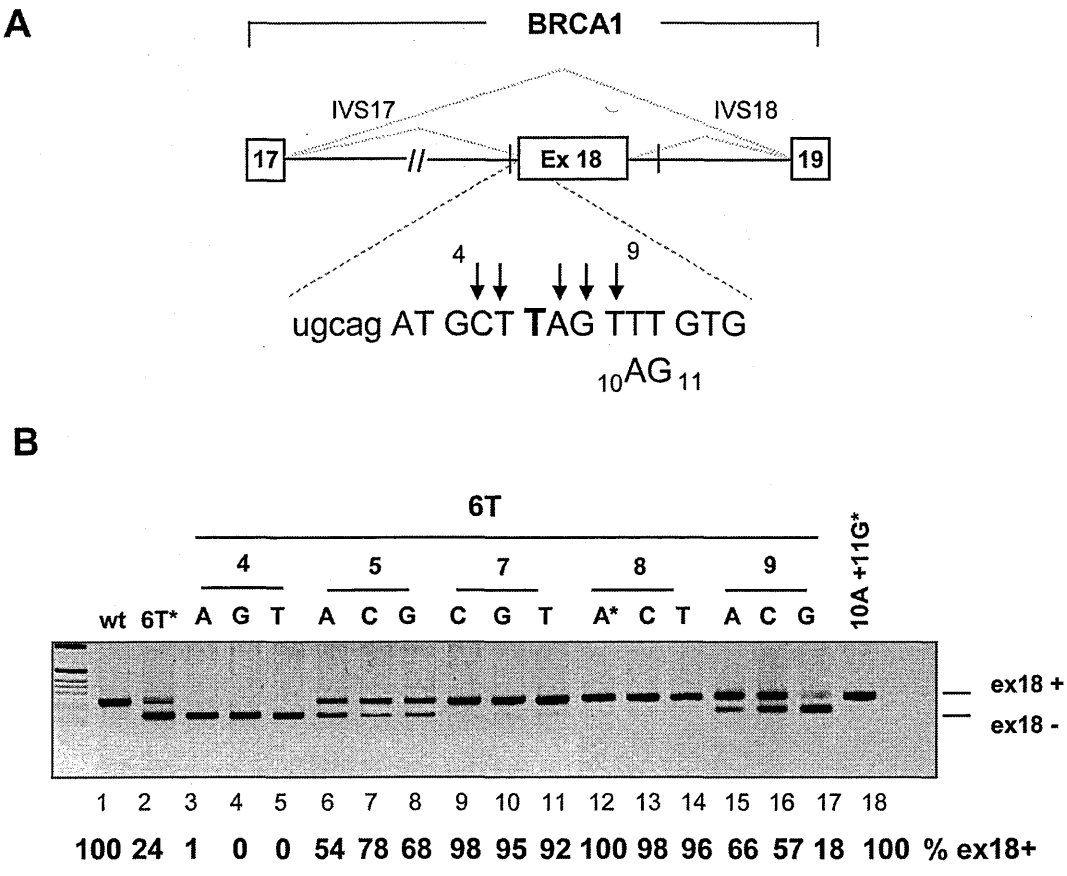

Figure 3.2.2: Double point mutation analysis of the exon 18 putative regulatory region.

(A) BRCA1 exons and introns cloned in the pBRA minigene are schematically reported. Part of intron 17 and BRCA1 exon 18 sequences are reported in lower and upper case, respectively. The positions of single point mutations, all introduced in the context of the 6T exon 18 sequence, are indicated by black arrows. The position of the double mutation +10A/+11G is also indicated below the sequence and it was the only one introduced in the exon 18 context.

(B) Splicing pattern analysis. The constructs were transiently transfected into Hep3B cells and the total RNA analysed by RT–PCR. The two splicing possibilities are indicated on the right of the gel. The mutations that create a stop codon are marked by an asterisk (lanes 2, 12 and 18). The majority of the mutants analysed restored partially or totally exon 18 inclusion compare to the T6 mutant. The mutants in position +4 further increase exon 18 exclusion. The percentages of BRCA1 exon 18 inclusion are reported below for every mutant analysed.

3.2.3 Analysis of silencer sequences in the BRCA1 exon 18 context.

Whereas single mutations had no significant effect in the exon 18 WT splicing, most of the double point mutations restored exon 18 inclusion in the presence of the T6 substitution. These results would not be expected if the mutated element were part of an ESE. Vice versa, these data are consistent with the hypothesis that the G to T transversion creates a negative-acting ESS within BRCA1 exon 18. This putative silencer sequence contains a TAG motif, previously identified in hnRNP A1–dependent ESS either by characterizing other gene sequences (Kashima and Manley 2003; Disset, Bourgeois et al. 2006; Kashima, Rao et al. 2007) either by performing hnRNP A1 SELEX analysis (Burd and Dreyfuss 1994). On the base of these previous data, we wanted to evaluate the effect of two known negative splicing regulatory sequences in the pBRA minigene context: the "TAGACA" from the SMN2 gene, exon 7 (Kashima and Manley 2003; Kashima, Rao et al. 2007) and the "TAGGGA" hnRNP A1 SELEX motif (Burd and Dreyfuss 1994). As a control, we also cloned the sequence of SMN1 (CAGACA) that should not have any inhibitory effect on splicing process. The sequences were inserted by PCR site-directed mutagenesis in the region from position +6 to +11 (Fig. 3.2.3A). The resulting hybrid minigenes were transiently transfected into Hep3B cells, and the splicing patterns analysed by RT-PCR. Exon 18 splicing was affected by the insertion of the selex A1 sequence at the same level as the T6 mutant, showing 20% of exon inclusion (Fig. 3.2.2B, lane 5). The SMN2 sequence in pBRA context showed a lower but still negative effect with 50% of exon 18 inclusion, whereas the control SMN1, which does not contain TAG, did not affect mRNA processing (Fig. 3.2.3B, lanes 4 and 3, respectively). The presence of these heterologous silencer sequences within BRCA1

exon 18 mimicked the effect of the original exonic regulatory element containing the T6 mutation.

A deletion from position +4 to +9 of exon 18 was also evaluated in the pBRA Δ4-9 construct. The analysis of the PCR product of the pBRA Δ4-9 did not show any variation in the splicing pattern in comparison with pBRA WT (Fig. 3.2.3B, lane 6) further ruling out the presence of a strong enhancer at the 5' end of exon 18.

As the G to T mutation in position +6 creates a premature termination codon within the exon coding sequence, we compared the results obtained with other mutants that also introduce a stop codon. This was done in order to exclude the potential involvement of the NMD pathway in the T6 exon 18 splicing. Thus, we compared the single T6 mutation with other mutations that created a stop codon within the exon 18 sequence. In particular two double mutations, T6A8 and 10A11G, and the Δ4-9 deletion create premature termination codons. In this case, and in contrast to the T6 mutant, the functional splicing analysis showed a complete inclusion of the exons (Fig. 3.2.2B, lanes 12 and 18; Fig. 3.2.3B lane 6). BRCA1 exon 18 stop codon and their role in triggering mRNA degradation was also recently studied (Perrin-Vidoz, Sinilnikova et al. 2002). In this analysis, performed by blocking translation with puromycin treatment, they showed that the steady-state level of mRNA carrying the T6 mutation in *vivo* was not affected (Perrin-Vidoz, Sinilnikova et al. 2002) indicating that this nonsense mutation did not trigger by NMD degradation.

Collecting the data obtained the single mutation G to T at position 6 of BRCA1 exon 18 is not involved in exon 18 splicing by creating a silencer element.

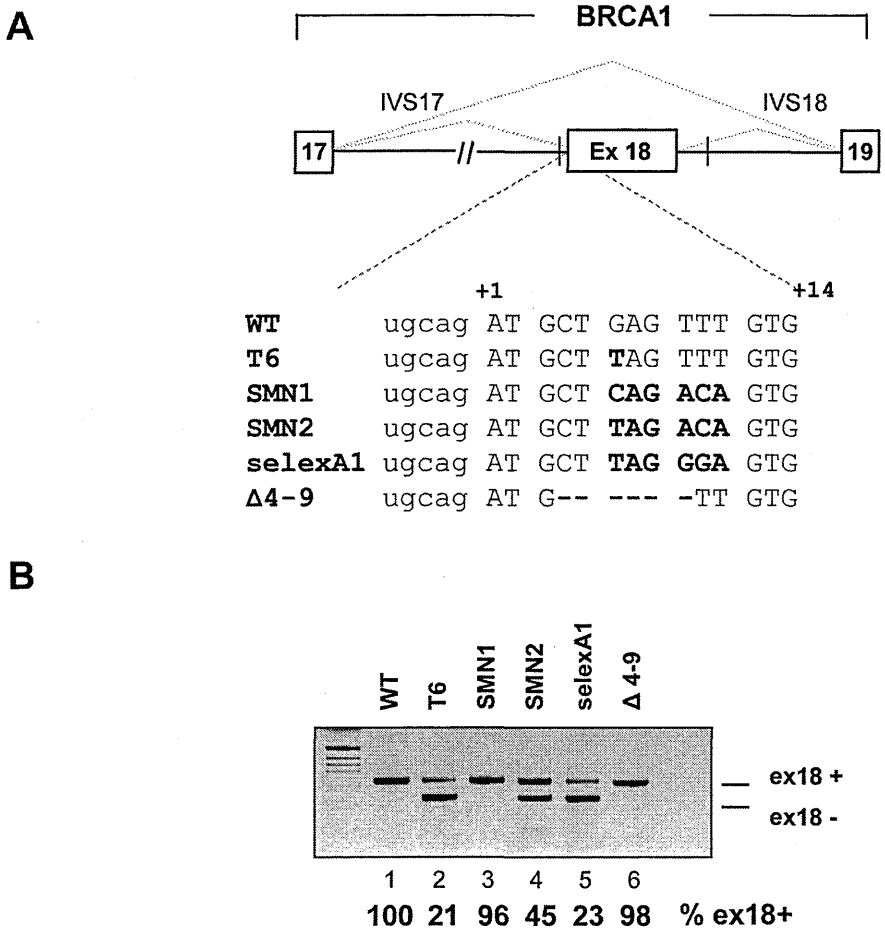

Figure 3.2.3: Analysis of silencer sequences in the BRCA1 exon 18 context.

(A) A scheme of BRCA1 exons and introns cloned in the pBRA minigene is reported. Part of BRCA1 intron 17 (lower case) and exon 18 (upper case) sequences are reported. The mutated nucleotides for each construct are highlighted in bold and the deleted region from position +4 to +9 is represented by dotted lines.

(B) Splicing pattern analysis. RT-PCR products of transfected Hep3B cells were separated on agarose gel. The position of the exon 18 inclusion and exclusion bands is indicated schematically on the right. The pBRA-SMN2 and selexA1 minigenes produced exon 18 skipping whereas SMN1 and Δ4-9 variants do not affect exon 18 recognition. The percentages of BRCA1 exon 18 inclusion are indicated below each lane.

3.2.4 Lack of Correlation between ASF/SF2 SR Protein Score Matrix and BRCA1 Exon 18 Splicing

In some exons, the SR protein score motif has been found to correlate with splicing efficiency, therefore raising the possibility of predicting the effect of single base substitutions on exon skipping (Liu, Cartegni et al. 2001; Cartegni, Chew et al. 2002).

The ASF/SF2, SC35, SRp40 and SRp55 motif-scoring matrices were used to analyse the entire BRCA1 exon 18 WT and T6 mutant sequences (Cartegni, Wang et al. 2003). The ESEfinder output showed multiple high-score motifs (3 matches for ASF/SF2, 2 for SC35, 1 for SRp40, and 4 for SRp55) distributed throughout the exon. The G to T substitution at position +6 specifically disrupts the first of three high-score ASF/SF2 motifs. reducing the score from 2.143 to 0.079 (below the threshold). ESEfinder program was also used to calculate ASF/SF2 dependent ESE motif scores of all the mutants analysed by functional splicing assay in order to test a possible correlation between the *in silico* approach versus the *in vivo* results (Table 3.2.1).

As reported in figure 3.2.1B, only three single mutants, T6, A6 and T4 affected the splicing of BRCA1 exon 18, leading to different percentage of exon skipping. All the three single substitutions disrupted the first high-score ASF/SF2 motif and, as predicted by the ESEfinder, produced exon 18 skipping. On the contrary, the program identified a higher number of single mutations able to disrupt the first predicted ASF/SF2 binding site, which did not correlates with the *in vivo* results. The majority of the mutants (18 out of 21) splicing pattern were not distinguishable from the WT construct showing complete exon 18 inclusion. The opposite situation was detected when the double mutants were searched for the ASF/SF2 high score motif.

In this case even though the substitutions in position +7 and +8, in the pBRA T6 context, restored almost completely the exon 18 inclusion the ESEfinder analysis predicted the disruption of the first ASF/SF2 binding site, in position +4. In addition three mutants in position +5 (T6A5, T6C5, T6G5) and two in position +9 (T6A9 and T6C9) that increased the correct splicing of exon 18, reaching a percentage of exon inclusion of an average of 60%, were not recognised by the program as putative ASF/SF2 binding site. On the other hand, the T6G9 double mutant that showed a splicing pattern comparable with the single mutant T6 (only 20% of exon inclusion), did not correlate with the predicted positive ASF/SF2 score value. The only evident correlation was obtained for the double mutants in position +4 (T6A4, T6G4, T6T4) in which the total skipping of exon 18 reported *in vivo* was in agreement with the under threshold ASF/SF2 score calculated by the program. SMN1 sequence had a good match for ASF/SF2 motif and gave total exon 18 inclusion. The other two hybrid sequences analyzed, SMN2 and selexA1, correlated with the prediction, showing low level of exon inclusion and lower score for ASF/SF2 binding. When the putative ESE was totally removed, in the Δ4-9 construct, as expected the ESEfinder program predicted the disruption of ASF/SF2 score value but *in vivo* the exon was fully recognized and included in the pre-mRNA.

All the results reported in the table 3.2.1 indicate that the currently available SR protein high score motifs does not explain the complex regulation of BRCA1 exon 18 and, in this context, any individual change in SR motif score cannot predict the splicing efficiency. We did not find any consistent correlation between *in vivo* splicing pattern and the ESEfinder prediction.

Table 3.2.1			
Minigene[a]	**Sequence[b]**	**% Exon 18 inclusion[c]**	**SF2/ASF[d]** Thr=1.956
WT	ATGCTGAGTTTGTG	100%	2.142(4)
A4	ATGATGAGTTTGTG	97%	-0.362(4)
G4	ATGGTGAGTTTGTG	100%	0.566(4)
T4	ATGTTGAGTTTGTG	90%	-0.807(4)
A5	ATGCAGAGTTTGTG	99%	3.265(4)
C5	ATGCCGAGTTTGTG	100%	1.540(4)
G5	ATGCGGAGTTTGTG	99%	2.818(4)
A6	ATGCTAAGTTTGTG	80%	0.079(4)
C6	ATGCTCAGTTTGTG	47%	2.389(4)
T6	ATGCTTAGTTTGTG	98%	0.079(4)
C7	ATGCTGCGTTTGTG	100%	1.149(4)
G7	ATGCTGGGTTTGTG	98%	-0.766(4)
T7	ATGCTGTGTTTGTG	97%	-0.308(4)
A8	ATGCTGAATTTGTG	96%	0.223(4)
C8	ATGCTGACTTTGTG	98%	2.747(4)
T8	ATGCTGATTTTGTG	100%	0.223(4)
A9	ATGCTGAGATTGTG	100%	1.685(4)
C9	ATGCTGAGCTTGTG	99%	1.685(4)
G9	ATGCTGAGGTTGTG	100%	4.263(4)
A10	ATGCTGAGTATGTG	96%	2.496(4)
G10	ATGCTGAGTGTGTG	95%	1.769(4)
G11	ATGCTGAGTTGGTG	98%	2.142(4)
T6A4	ATGATTAGTTTGTG	1%	-2.426(4)
T6G4	ATGGTTAGTTTGTG	0%	-1.498(4)
T6T4	ATGTTTAGTTTGTG	0%	-2.871(4)
T6A5	ATGCATAGTTTGTG	54%	1.202(4)
T6C5	ATGCCTAGTTTGTG	68%	-0.524(4)
T6G5	ATGCGTAGTTTGTG	78%	0.754(4)
T6C7	ATGCTTCGTTTGTG	98%	-0.914(4)
T6G7	ATGCTTGGTTTGTG	95%	-2.829(4)
T6T7	ATGCTTTGTTTGTG	92%	-2.371(4)
T6A8	ATGCTTAATTTGTG	100%	-1.840(4)
T6C8	ATGCTTACTTTGTG	98%	0.683(4)
T6T8	ATGCTTATTTTGTG	96%	-1.840(4)
T6A9	ATGCTTAGATTGTG	66%	-0.378(4)
T6C9	ATGCTTAGCTTGTG	57%	-0.378(4)
T6G9	ATGCTTAGGTTGTG	25%	2.199(4)
SMN1	ATGCTCAGACAGTG	96%	3.038(6)
SMN2	ATGCTTAGACAGTG	45%	-2.232(4)
selexA1	ATGCTTAGGGAGTG	23%	1.826(4)
Δ4-9	ATG------TTGTG	98%	-3.632(4)

SR High-score motifs (a) The 41 minigenes used in transfection experiments, in the same order as they are reported in the figures. (b) Nucleotide sequence from position +1 to +14 around the putative ESEmotifs. The mutated nucleotides are reported in bold.(c) Average of exon 18 inclusion percentages (d) ASF/SF2 score. High-score motifs are shown in black whereas the lower-scores are reported in red. The threshold value is 1.956. (e-f) SC35 and SRp55 new binding sites. The high-score values acquired by the presence of the mutations are reported.

3.3 Identification of nuclear proteins binding to the splicing regulatory element of the BRCA1 exon 18 T6 mutant.

3.3.1 Pull down assay and identification of splicing factors differentially binding to BRCA1 exon 18 WT and U6 RNAs.

Based on the systematic mutagenesis, the T6 mutation seems to create a sequence with silencer properties. This negative sequence may bind one or more splicing factors with an inhibitory role in the BRCA1 exon 18 splicing regulation. In order to identify specific trans-acting factor(s) binding to BRCA1 exon 18 T6 mutant I performed a RNA pull-down analysis. The pull down experiment was carried out using the two synthetic RNA oligonucleotides (20nt) corresponding to the first part of exon 18 and containing either the WT or the natural U6 mutation (Fig. 3.3.1A). These small RNAs were covalently linked to adipic acid dehydrazide agarose beads via 3' end and incubated with HeLa nuclear extract. After the washing steps the proteins, tightly bound to the RNA, were separated on SDS-PAGE gel and visualized by Coomassie Blue staining. The figure 3.3.1B shows the result of pull down assay after Coomassie staining in which most of the bands were present in both WT and U6 sequences. Interestingly, the U6 exon 18 RNA showed two bands not present in the WT lane. Vice versa, no additional bands were specifically detected to bind the WT sequence. The two specific bands with an apparent molecular weight of ~50 and 38 kDa were excised from the gel and sequenced using electrospray mass spectrometric analysis. Next, a database search revealed that the band of 35 kDa was a doublet corresponding to both hnRNP A1 and its close homologue hnRNP A2 splicing factor.

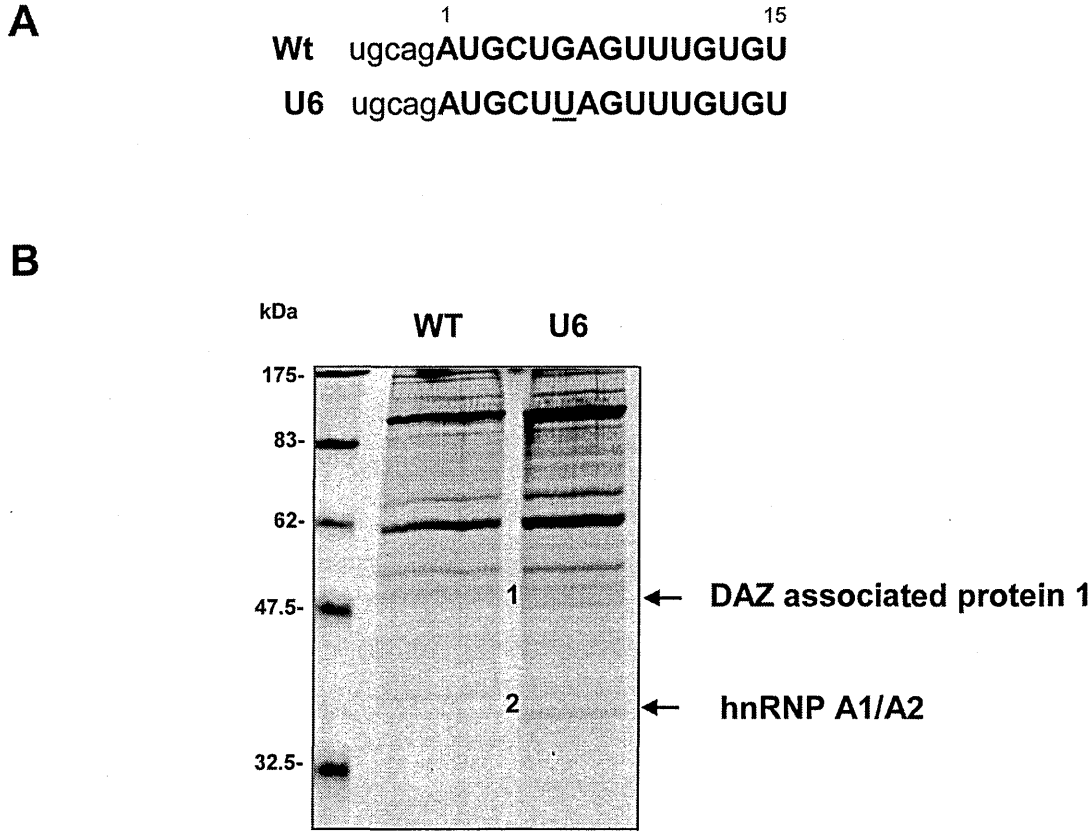

Figure 3.3.1: Identification of the proteins able to bind the BRCA1 exon 18 U6 RNA sequence.

(A) Sequences of the BRCA1 RNA oligonucleotides. Intronic and exonic sequences are shown in lower and upper case, respectively, and the mutated nucleotide U6 is underlined.

(B) Coomassie stained gel of a pull-down assay. The target RNAs were linked to adipic beads and incubated with HeLa nuclear extract. After several washing steps the proteins were loaded on SDS-Page gel and stained with Coomassie. The arrows indicate two bands of about 50 and 38 kDa present only in the U6 lane. The two bands were analysed by mass spectrometry analysis and the proteins were identified as DAZAP1 (1) and hnRNP A1/A2 (2), respectively.

The sequence of the band of 50 kDa matched with the Deleted in Azoospermia Associated Protein 1, named DAZAP1. HnRNP A1 and hnRNP A2 have been described as "negative" splicing regulators in different gene systems and in several cases showed a high binding affinity for RNA sequences containing the TAG motif (Burd and Dreyfuss 1994; Kashima and Manley 2003; Disset, Bourgeois et al. 2006; Kashima, Rao et al. 2007). In contrast to hnRNP factors, relatively little is known concerning the DAZAP1 and its role in splicing process. DAZAP1 shares a similar structure with the member of the hnRNP protein family (Dreyfuss, Matunis et al. 1993; Akindahunsi, Bandiera et al. 2005) and is present in the nucleus of somatic cells as a component of the hnRNP particles (Lin and Yen 2006). To date the role for DAZAP1 has been connected with the mRNA transport and stability (Lin and Yen 2006; Morton, Yang et al. 2006).

3.3.2 Pulldown analysis of BRCA1 exon 18 WT and U6 RNAs probed with hnRNP A1, -A2 and DAZAP1.

To study more in detail the differential binding efficiency of the splicing factors identified a pulldown assay including the "internal control" was performed. As described for ASF/SF2, the original pull down methodology was improved by linking the first part of BRCA1 exon 18 WT or U6 mutat to six UG dinucleotide repeats for TDP43 binding normalization (Fig. 3.3.2A).

The followed procedure is described in figure 3.1.2B: the target RNAs were T7-transcribed *in vitro*, treated with periodate in order to oxidize its 3' end making it able to bind the agarose beads. Next, incubation with nuclear extract was performed and numerous washes allowed the elimination of all the nuclear proteins with no affinity for our RNA.

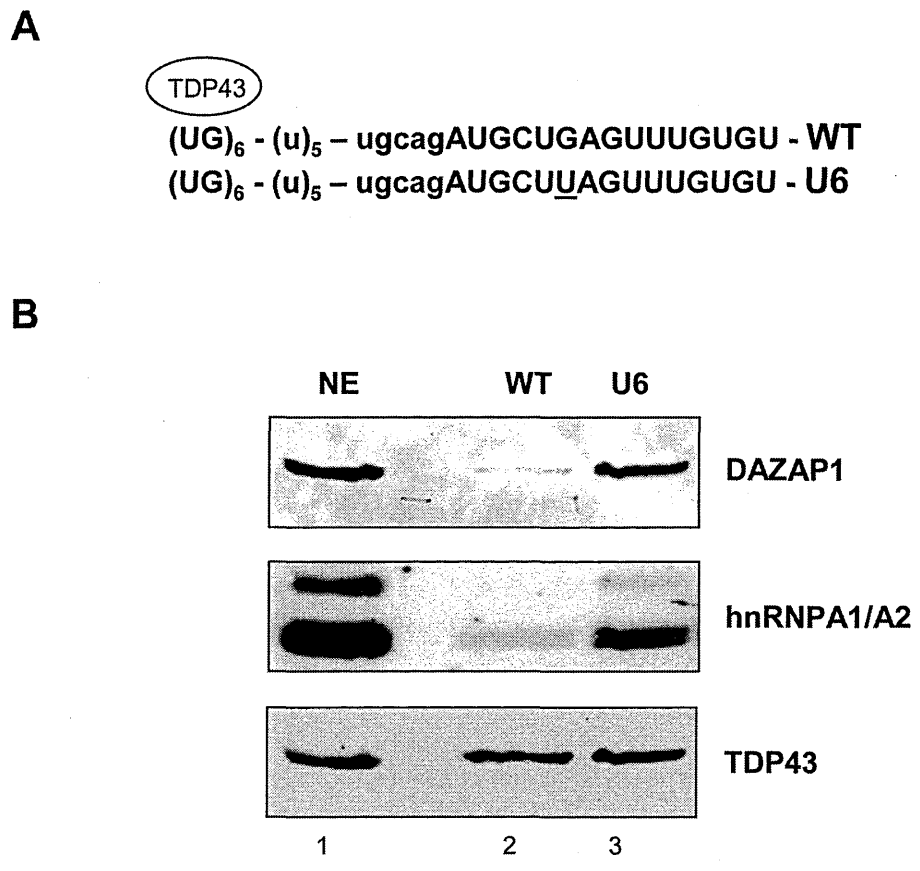

Figure 3.3.2: Western blot of pulldown analysis of BRCA1 exon 18 WT and U6 RNAs to determine the presence of hnRNP A1, -A2 and DAZAP1.

(A) Schematic representation of the RNA sequence transcribed *in vitro* (upper panel) that contains $(UG)_6$ repeats, a $(U)_5$ spacer and the target BRCA1 sequences. The mutagenized nucleotide is underlined. (B) Binding analysis of hnRNP A1, A2 and DAZAP1. The RNAs were *in vitro* transcribed, covalently linked to agarose beads and incubated with HeLa nuclear extract. The pulled down proteins were then analysed by western blotting with anti-hnRNP A1, hnRNP A2, DAZAP1 and TDP43 polyclonal antibodies. Immunoblottings were detected using the ECL method. The nuclear extract sample corresponds to 1/20th of the amount used for the pull-down assay. The pulldown picture is representative of three independent experiments. As the picture shows, the U6 RNA is able to pull down hnRNP A1, -A2 and DAZAP1 proteins. The WT binds approximately only 10 and 15% of DAZAP1 and hnRNP A1, -A2 proteins respectively, compared to the U6 RNA, normalized for TDP43. The percentage of binding was determined by the evaluation of the relative intensity of the bands.

Finally, complexes of proteins-RNA were separated in a SDS-page gel, blotted onto PVDF filters and evaluated using polyclonal antibodies against DAZAP1, hnRNP A1/A2. As shown in figure 3.3.2B the membrane was initially probed against TDP43, to normalize the experiment for the amount of bound RNA (Fig. 3.3.2B, lower panel) followed by incubation with DAZAP1 and hnRNP A1/A2 antibodies. In the presence of a comparable amount of pulled down TDP43, the U6 but not the WT mutant, bound very efficiently to DAZAP1 (Fig. 3.3.2B, upper panel lanes 2-3) and hnRNP A1/A2, whereas no signals were visible in the WT lane (Fig. 3.3.2B, middle panel, lanes 2-3). Thus, DAZAP1 together with hnRNP A1/A2 bind specifically to the BRCA1 exon 18 carrying the T6 natural mutation and can be responsible of the exon skipping phenotype.

3.3.3 Electro-mobility shift assay (EMSA) of BRCA1 exon 18 WT and mutants.

The RNA binding specificity of DAZAP1 and both hnRNP A1-A2 was also tested directly by electro-mobility shift assay (EMSA). In this case, the EMSA analysis was carried out with purified hnRNP A1, hnRNP A2 and DAZAP1 GST-tagged proteins (Fig. 3.3.3B). To perform the EMSA experiments we used the same target RNAs, WT and the U6 exon 18 sequences, utilized for the pulldown assay. An additional RNA oligonucleotide, which carries the double mutation G4U6, was chosen as target RNA (Fig. 3.3.3A). This mutant induced complete exon 18 skipping in the functional splicing assay analysis (Fig. 3.2.2B, lane 4) showing a more negative nature compared to the single T6 mutant. Labelled WT and mutants RNAs were incubated with each different recombinant protein to let the binding reaction formed. Then the RNA-protein complexes were resolved on a native acrylamide gel in order to check any binding differences between the RNAs and proteins interaction. The analysis

showed for each protein a similar binding pattern (Fig. 3.3.3C). It is visible a band of shifted material with the RNAs carrying the U6 and G4U6 mutations, whereas no complexes are formed in the presence of WT sequence. In the case of G4U6 RNA the complex signal resulted stronger than that one formed by the interaction between the proteins and the U6 RNA. These differences in binding are visible for all the three proteins analysed indicating a higher binding affinity for the double mutant sequence.

This EMSA study indicates that DAZAP1 and hnRNP A1-A2 proteins specifically recognize the U6 mutation as in the pulldown experiment. Furthermore, the binding efficiency showed by the WT, U6 and G4U6 RNAs against the three splicing factors was in agreement with the percentage of exon 18 skipping obtained by the functional splicing assay. Considering that the hnRNP A1, -A2 and DAZAP1 proteins could play an inhibitory role through exon T6 recognition, its partial exclusion from the pre-mRNA transcript could be considered due to the creation of a new binding site for the proteins identified. On the other hand, the exon 18 WT does not contain any binding site for these proteins; it is correctly recognised by the spliceosomal machinery and fully included in the mRNA.

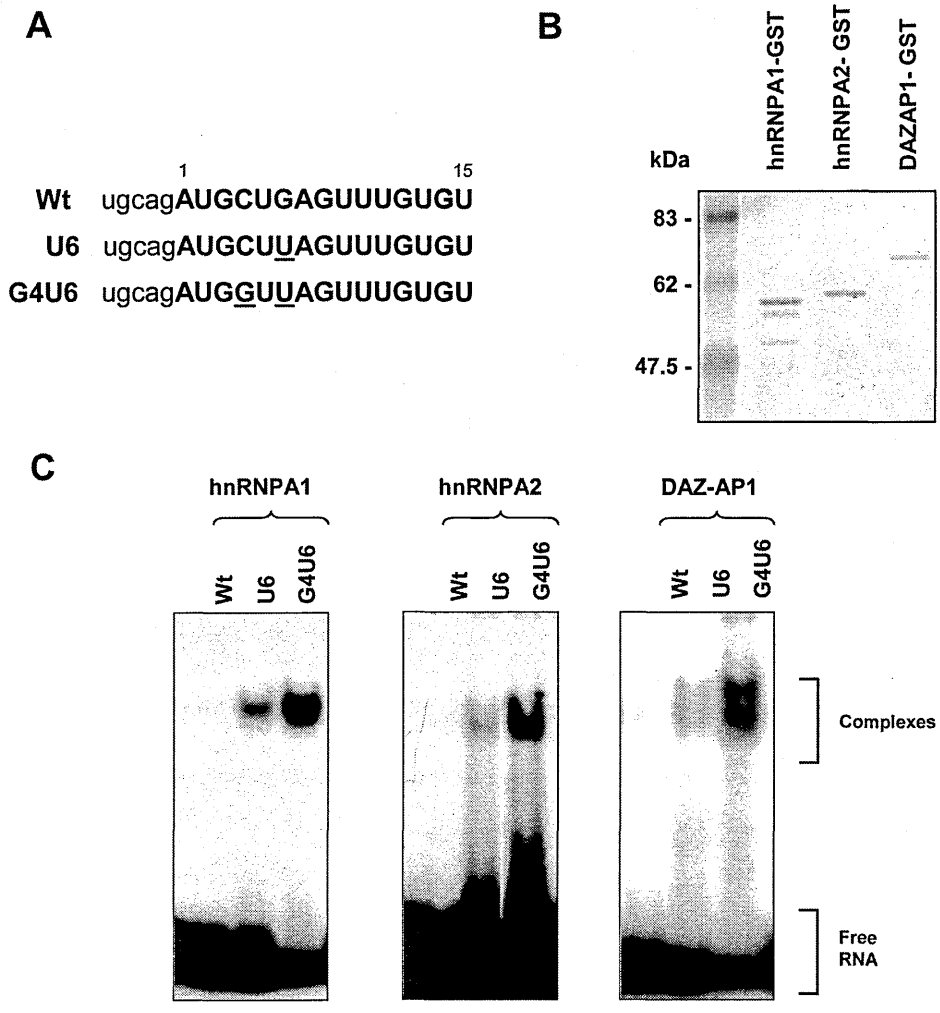

Figure 3.3.3: Electro mobility shift assay (EMSA) binding analysis of BRCA1 exon 18 WT and mutants.

(A) Oligonucleotide sequences used for EMSA analysis. The RNAs are 20 bp long, 5 nucleotides from intron 17 and 15 from exon 18 (capital letters) WT, U6 and G4U6. The mutated positions are underlined.

(B) Coomassie blue staining of the purified GST- proteins. The tagged proteins were expressed in BL21 (DE3) bacteria (Novagen) under the induction of IPTG. Purification of the proteins was carried out according to the manufacturer's instructions (see materials and methods).

(C) EMSA assay using purified human GST-hnRNP A1, A2 and DAZAP1 proteins with WT, U6 and G4U6 RNA oligonucleotides. Labelled RNAs were incubated with different GST-proteins as indicated. The RNA-protein complexes were resolved on 5% polyacrylamide gel. As a control only the WT RNA was loaded in the first lane of each gel. The position of the free RNA and of the shifted material complexes is indicated on the right of the gels.

3.4 Functional analysis of hnRNP A1, -A2 and DAZAP1 in BRCA1 splicing regulation.

3.4.1 Effect of hnRNP A1, -A2 and DAZAP1 overexpression on WT and T6 exon 18 splicing.

The EMSA analysis together with the RNA pulldown experiments strongly indicated that the BRCA1 exon 18 +6 G to T substitution creates a new binding site for hnRNP A1, A2 and DAZAP1 proteins. In order to highlight the functional role of these trans-acting factors in the BRCA1 exon 18 splicing regulation an *in vivo* approach was set up. A transfection experiment was performed in Hep3B cells in which pBRA T6 hybrid minigene together with the single or different combinations of hnRNP A1, -A2 and DAZAP1 coding plasmids (+), or the empty vector (-) as a control. The minigene pBRA WT was also analysed in the presence of the three coding plasmids. Total RNA was extracted and a RT-PCR was carried out with a specific primer set. Figure 3.4.1A reports the splicing pattern analysed by agarose gel electrophoresis. The pBRA T6 minigene splicing was not affected by the ovexpression of each single splicing factor, hnRNP A1, hnRNP A2 or DAZAP1, showing a percentage of exon inclusion equal to the control, of about 20% (Fig. 3.4.1A, lane 1-3 vs. lane 8). In order to test whether the different splicing factors could induce an effect on exon 18 splicing when co-expressed, we performed the overexpression experiment combining each splicing factor with a second one and moreover transfecting all the three factors together. Surprisingly, neither the overexpression of the double coding plasmids in all the three possible combinations (Fig. 3.4.1A, lanes 4-6) neither the triple co-transfection (Fig. 3.1.4A, lane 7) interfered with the splicing processing of pBRA T6 plasmid that showed the same percentage of inclusion observed for the control (Fig. 3.4.1, lane 8).

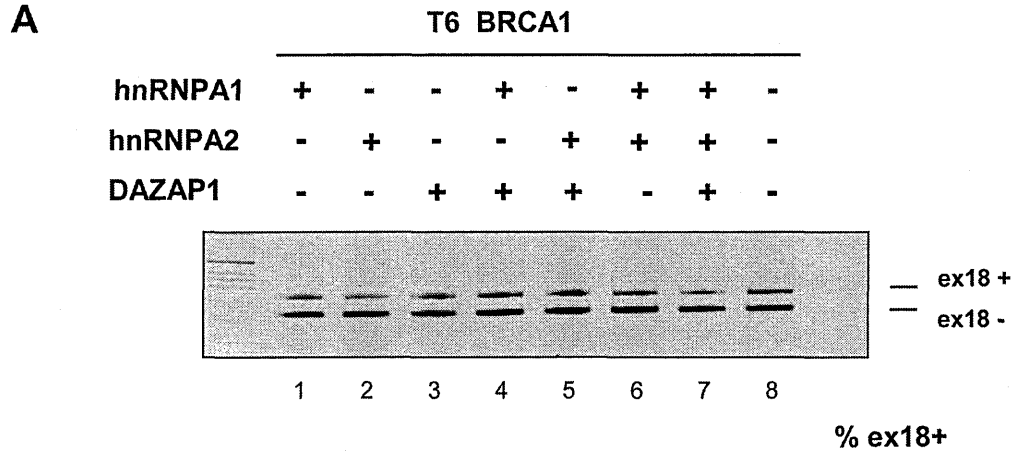

Figure 3.4.1: Effect of hnRNP A1, -A2 and DAZAP1 overexpression on T6 exon 18 splicing.

(A) Hep3B cells were transfected with the indicated exon 18 minigene pBRA T6 mutant either in presence of the empty vector (–) or with the splicing factor coding plasmid of interest (+). Total RNA was extracted and splicing variants detected by RT-PCR were resolved on 2% agarose gel. The identity of the transcripts, exon 18 inclusion (ex18+) and exon 18 exclusion (ex18-), is shown schematically at the right side of the gel.

No effect was observed in the case of the exon 18 WT minigene and overexpression of the three splicing factors together did not change the splicing profile of the exon 18 WT that remained 100% of inclusion (Fig. 3.4.1B, lane 1-2).

Co-transfection with hnRNP A1, A2 and DAZAP1 expression vectors should result in an inhibition of exon 18 splicing taking into account the putative negative role of these splicing factors. The absence of a negative effect on exon 18 T6 mutant could be explained by a possible saturation of the system due to the high expression level of these proteins within the cells. In addition, considering the fact that exon 18 T6 is nearly skipped (~80%) in the absence of any co-transfected factors, the detection of

any negative effect on splicing may be difficult to achieve. Thus, the complete absence of any response to hnRNP A1, A2 and DAZAP1 overexpression in the pBRA T6 minigene do not exclude their real involvement in exon 18 splicing regulation.

3.4.2 siRNA against hnRNP A1, A2 and DAZAP1 affects BRCA1 T6 exon 18 splicing.

In order to evaluate a possible functional role of hnRNP A1, A2 and DAZAP1 on BRCA1 exon 18 splicing I performed depletion of the target proteins by siRNA treatment. This represents a more suitable system to investigate whether these splicing factors are involved in the aberrant splicing of the mutated exon 18. If their binding, as supposed, promotes exon 18 skipping, the specific knock down of these proteins should increase the inclusion of the exon into the mature RNA.

To test whether the siRNA oligonucleotides against hnRNP A1, hnRNP A2 and DAZAP1 effectively reduces the target protein levels a transient transfection of these small RNAs in HeLa cell line was set up. Both single and different combinations (double and triple) of siRNA transfections were performed against these splicing factors. After two round of treatment the cells were collected, total proteins extracted and the silencing effect analysed by western blot using antibodies against each protein of interest. All the siRNAs induced a strong reduction in the target proteins level, whereas the cells transfected with an unrelated siRNA, against the Luciferase gene, did not show any variation in the proteins content. As internal control for the western blot analysis, to normalise the amount of loaded proteins, the expression level of tubulin gene was detected.

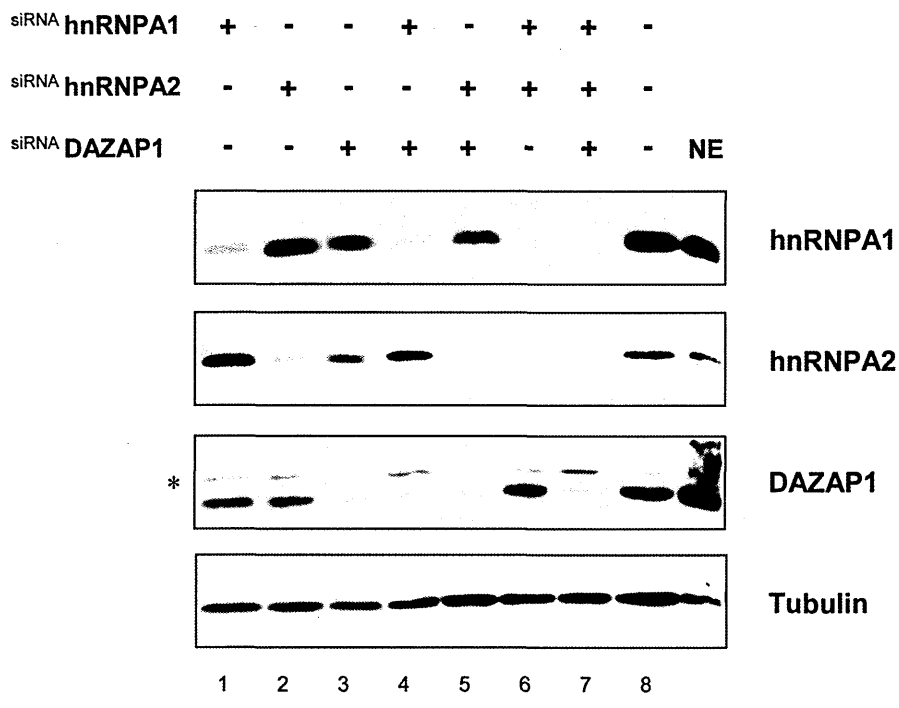

Figure 3.4.2: Knock down of hnRNP A1, -A2 and DAZAP1 proteins performed by siRNA treatment.

Western blot analysis of HeLa cells treated (+) with the indicated different siRNAs anti-hnRNP A1, -A2 and DAZAP1 proteins, alone or indifferent combinations. The control luciferase siRNA corresponds to lane 8. The analysis was performed using only one siRNA against each target. The western blot was carried out with polyclonal antibodies against each target protein and tubulin probing as a control of total protein loading. NE indicates the nuclear extract. The percentage of silencing of each protein reaches about the 80-90%. The siRNA experiment reported is representative of at least three independent experiments. The band indicated with an asterisk in the DAZAP1 blot corresponds to a unspecific protein detected by the polyclonal antibody against DAZAP1.

The cells treated with the different siRNAs were then transfected with the pBRA T6 plasmid and collected. Total RNA was extracted from half of the pellet whereas the second half served for total proteins analysis. In this way, splicing was evaluated by RT-PCR performed on total RNA extract and at the same time, the silencing effect was analysed by western blot using antibodies against each protein of interest. HeLa

cells transfected with the corresponding siRNAs showed, as described above, a strong reduction in hnRNP A1, A2 and DAZAP1 levels, whereas no effect was detected in the luciferase control (Fig. 3.4.2, lanes 1-7 vs. 8). The siRNA treatment against DAZAP1 induced also a small reduction in the amount of hnRNP A1 and hnRNP A2 (Fig. 3.4.2, lane 3) possibly related to a regulatory effect of DAZAP1 on hnRNP A1/A2 splicing and/or stability (Lin and Yen 2006). The combined siRNA-mediated knockdown of the three inhibitory splicing factors increased the percentage of exon inclusion of the T6 mutant from 20% to about 55% (Fig. 3.4.3A, lane 7 vs. 8). A slightly smaller effect on mutated exon 18 recovery was observed for the simultaneous siRNA-mediated depletion of hnRNP A1 and hnRNP A2 (Fig. 3.4.2B, lane 6), with about 45% of exon inclusion. In the case of the other two double combinations of hnRNP A1/DAZAP1 and hnRNP A2/DAZAP1 silencing, the splicing pattern of pBRA T6 minigene was not affected, resulting in 20% of exon inclusion (Fig. 3.4.2B, lanes 4-5 vs. 8). Depletion of each factor alone also showed no change in the percentage of the exon 18 inclusion in comparison to the not treated sample (Fig. 3.4.2B, 1-3 lanes vs. 8).

As a control for hnRNP A1/A2 silencing efficiency, pBRA T6 minigene was transfected together with the pSMN2 construct. Depletion of hnRNP A1 and –A2 was reported to rescue significantly of the SMN2 exon 7 skipping (Cartegni, Hastings et al. 2006; Kashima, Rao et al. 2007). Indeed, the pSMN2 splicing pattern responded to both hnRNP A1 and hnRNP A2 depletion increasing the exon 7 inclusion from 10% to 50% (Fig. 3.4.4A). As additional siRNA control, we tested whether the depletion of all the three splicing factors together could affect the proper exon 18 WT inclusion. The splicing pattern analysis, reported in figure 3.4.4B, showed that the siRNA treatment did not affect the 100% exon 18 WT inclusion. The depletion of hnRNP A1, -A2 and DAZAP1 factors affected exon 18 T6 splicing,

showing a functional role of the proteins identified and characterized by *in vitro* binding analysis. Overall, the single substitution G to T creates a silencer element, which is recognized by the hnRNP A1 and hnRNP A2 and the DAZAP1 splicing factors.

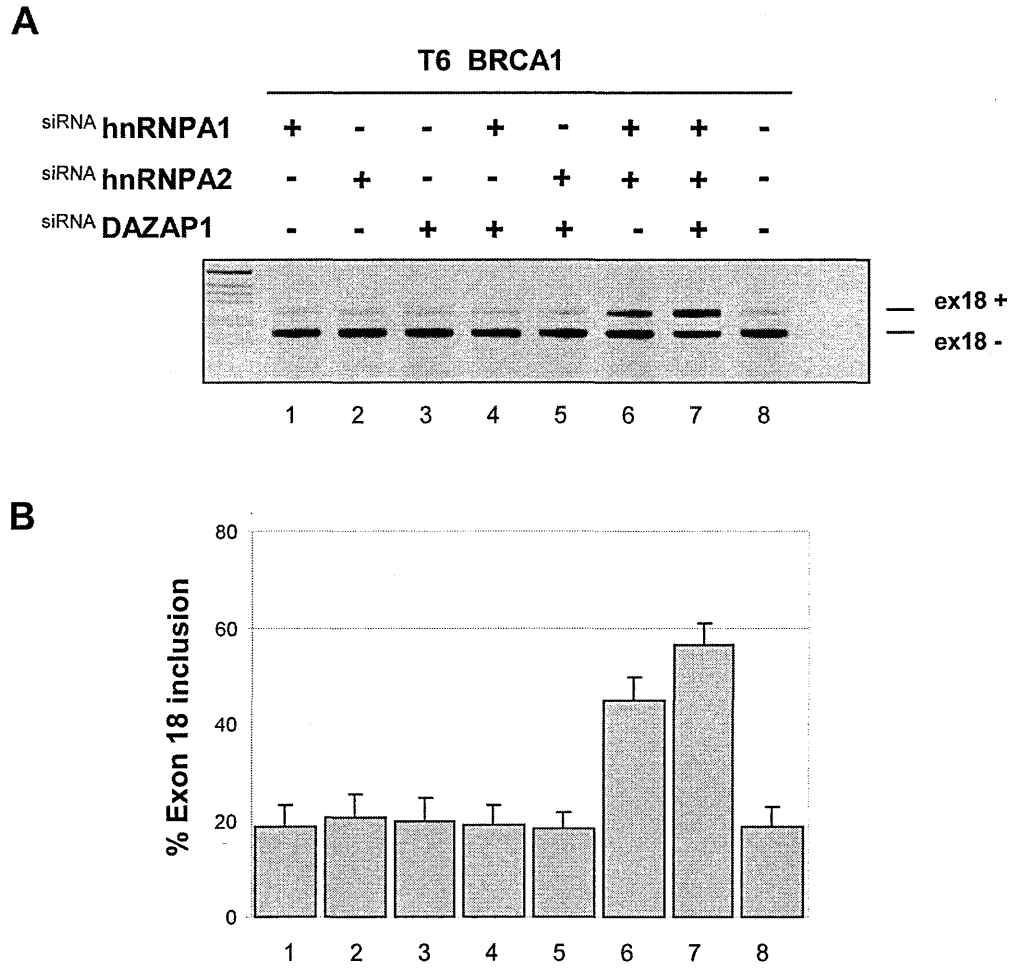

Figure 3.4.3: siRNA against hnRNP A1, A2 and DAZAP1 affects BRCA1 T6 exon 18 splicing.

(A) siRNA-treated (+) and untreated (-) cells were transfected with pBRA T6 minigene plasmid. Total RNA was prepared and analyzed by RT–PCR. The splicing products obtained were separated on 2% agarose gels and stained with ethidium bromide. The identity of the transcripts including exon 18 (ex18+) and lacking exon 18 (ex18-) is indicated. (B) The graph is the quantification of three independent experiments expressed as means ± SD.

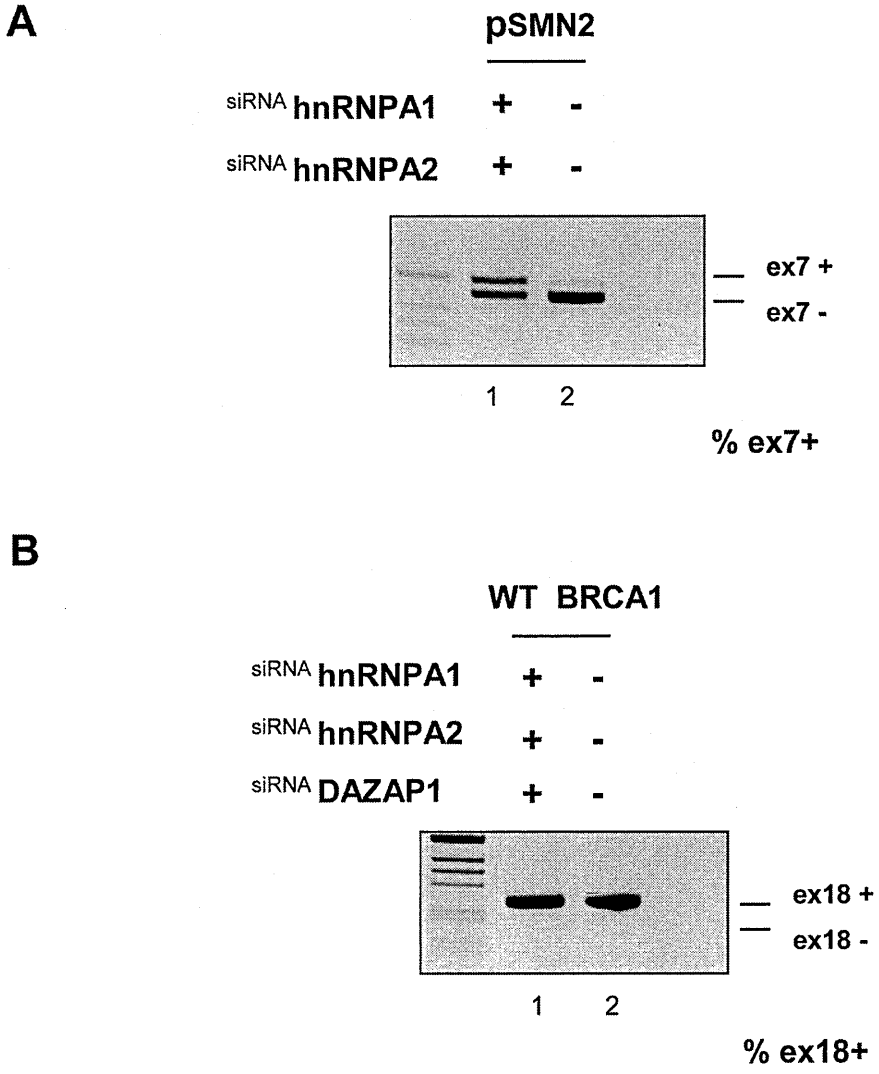

Figure 3.4.4: siRNA treatment on SMN2 exon 7 minigene system and pBRA WT as controls.

(A) Effect of double siRNA against hnRNP A1 and hnRNP A2 on the splicing pattern of the control pSMN2 minigene. pSMN2 and pBRCA1 T6 minigenes were cotransfected in the siRNA hnRNP A1 and hnRNP A2 treated cells and analysed with specific primers for pSMN2. The identity of the transcript with inclusion or exclusion of the SMN2 exon 7 is indicated.

(B) Effect of siRNA against hnRNP A1, -A2 and DAZAP1 on the WT pBRA minigene.

3.5 Mapping exonic regulatory elements within BRCA1 exon 18 sequence.

3.5.1 Serial deletions analysis within BRCA1 exon 18.

In order to identify additional BRCA1 exonic splicing regulatory element(s), I prepared several minigenes that contained a series of consecutive deletions along the BRCA1 exon 18 WT and carrying the T6 mutation (Fig. 3.5.1A, 3.5.2A). The deletions (~10nt) were produced in the exon sequence by overlapping PCR and the resulting pBRA minigenes were transiently transfected in Hep3B cells. Total RNA was extracted, RT-PCR was performed and the amplified products were analysed by electrophoresis on agarose gel.

All the deletions analysed in the context of the pBRA WT had no effect on the exon 18 splicing showing 100% exon inclusion (Fig. 3.5.1B) consistent with the absence of strong enhancer sequence in the WT exon 18. On the other hand, some of the deletions produced had an effect in the pBRA T6 context. Because each deletion was performed in the context of exon 18 T6 mutant, each splicing profile was compared with the pBRA T6 plasmid. As shown in figure 3.5.2B the majority of deletions within the mutated minigene context affected positively or negatively the percentage of exon inclusion, in comparison to the T6 mutant. Deletion of the sequences between positions 23-32 in T6 induced complete exon 18 skipping (Fig. 3.5.2B, lane 4), indicating that this sequence contains an enhancer, which becames relevant only in the context of the T6 mutant. Some of the deletions had a positive effect on the splicing pattern. In comparison to the T6 mutant, the T6Δ35-44 and T6Δ64-74 minigenes, increased the percentage of exon inclusion up to 90% and 100%, respectively (Fig. 3.5.2B, lane 5 and 7). These results suggested the presence of two silencer elements located at position 35-44 and 64-74. Interestingly the deletions disrupted a nucleotide sequences that include one or two "TAG" elements, which

maybe involved in the negative regulation of exon 18 together with the TAG located at position +6. Two deletions, the T6Δ12-21 and T6Δ47-57 minigenes, did not show any effect on the splicing pattern comparable with the T6 mutant (Fig. 3.5.1B, lanes 3 and 6). Furthermore, a possible connection between the two 4-9 and 23-32 deleted regions was tested by creating the Δ4-9/Δ23-32 minigene (Fig. 3.5.2A). These two deletions alone had no effect in the context of the exon 18 WT splicing whereas the 23-32 deletion induced exon skipping only when associated to the T6 mutant (Fig. 3.5.2B, lanes 4). The minigene carrying the double deletions (4-9 and Δ23-32) induced a partial exon 18 skipping (Fig. 3.5.2B, lane 8). This data suggested the presence of two nearby enhancer sequences, which may contribute in an additive manner to the definition of the WT exon. The extent of splicing inhibition mediated by the Δ23-32 deletion in the context of the single T6 substitution is more severe than the splicing inhibition mediated by the presence of the Δ4-9 deletion (Fig. 3.5.2B, lane 4 vs. 8). In fact, the T6Δ23-32 minigene induced complete exon exclusion, whereas the Δ4-9/Δ23-32 minigene showed partial exon skipping with only 50% of exon inclusion. This result is consistent with the strong inhibitory effect mediated by the binding of hnRNP A1/A2 and DAZAP1 on the T6 mutant region.

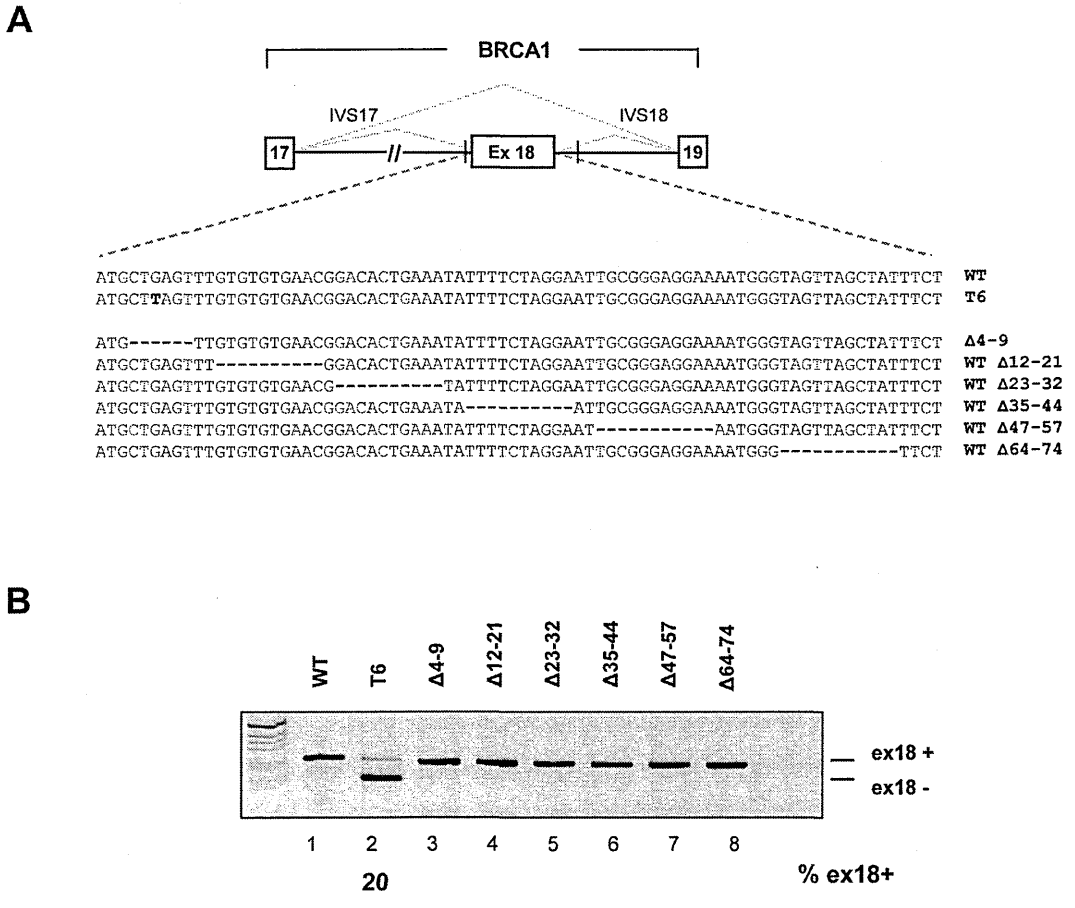

Figure 3.5.1: Splicing effect of single deletions performed on BRCA1 exon 18 WT sequence.

(A) The central part of pBRA minigene containing the BRCA1 exons and introns is indicated schematically. The entire sequence of exon 18 WT (78bp) is reported below. The G to T substitution in position +6 is highlighted in bold. The single deletions performed within BRCA1 exon 18 WT sequence are indicated by dashed lines. The exact position of each deletion within exon 18 is reported by a number on the right of the sequence.

(B) Splicing pattern analysis. The minigenes reported above were transfected in Hep3B cells and the total RNA analysed by RT–PCR. The two resulting splicing products corresponding to exon 18 inclusion (ex18+) and exclusion (ex18-) are indicated on the right. None of the deletions affected the exon 18 splicing. The only percentage of exon 18 inclusion reported is referred to the pBRA T6 minigene. All the constructs, a part T6, are considered 100% included.

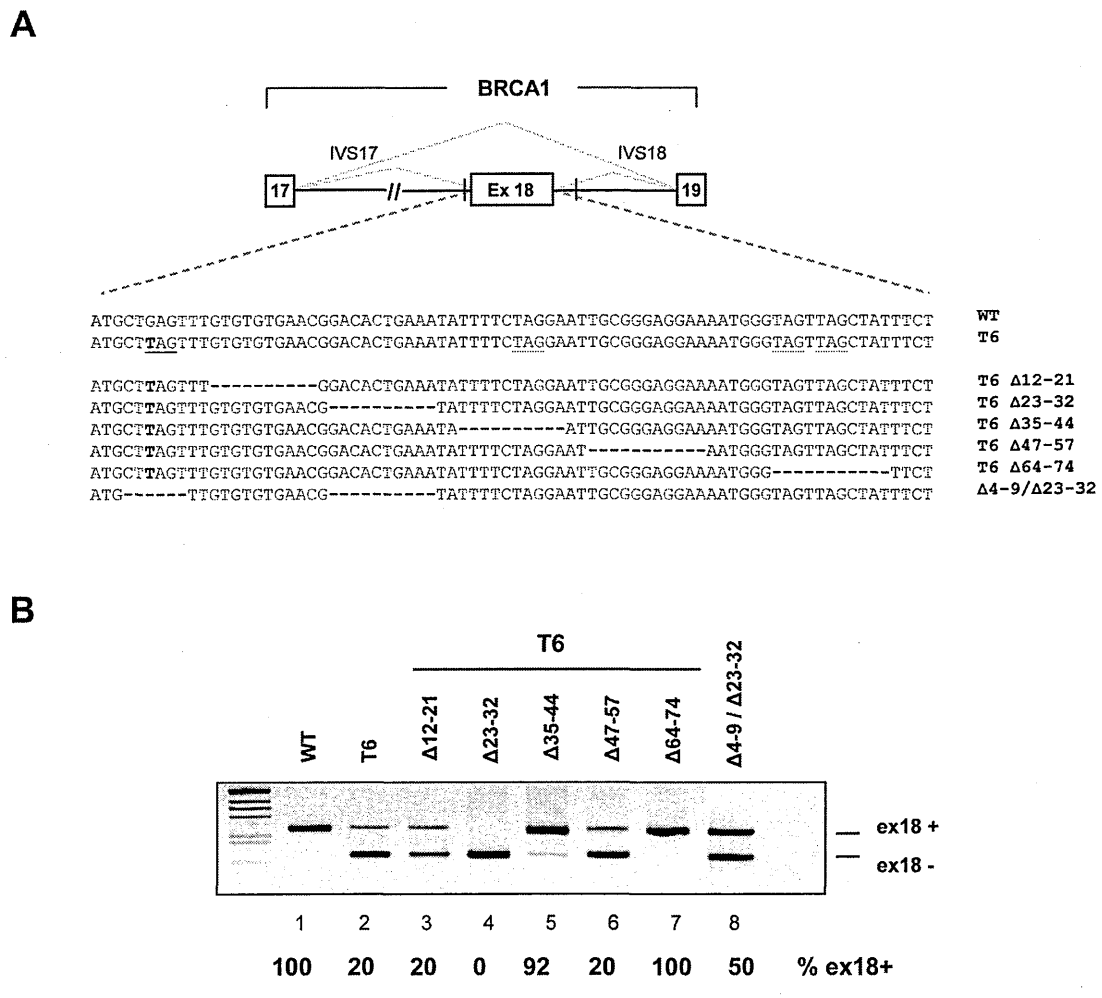

Figure 3.5.2: Splicing effect of single deletions performed on BRCA1 exon 18 T6 sequence.

(A) Scheme of pBRA WT construct that comprises BRCA1 exons and introns sequences. The 78nt of exon 18 WT is reported below. The G to T mutation in position +6 is highlighte in bold and the TAG motives are underlined. The single deletions performed within BRCA1 exon 18 T6 context and the double deletion (Δ4-9/Δ23-32) are indicated by dashed lines.

(B) Splicing pattern analysis. The minigenes indicated above were transfected in Hep3B cells and the total RNA analysed by RT–PCR. The two resulting splicing products corresponding to exon 18 inclusion (ex18+) and exclusion (ex18-) are indicated. Some of the deletions analysed affected negatively (Δ23-32) or positively (Δ35-44 and Δ64-74) the T6 exon 18 splicing. The percentage of exon inclusion in each case is shown.

3.5.2 Effect of ASF/SF2 overexpression in exon 18 T6 deleted minigenes.

In order to identify potential splicing regulatory sequences that could mediate the ASF/SF2 enhancing effect, observed on the T6 mutant in co-transfection experiments, (Fig. 3.1.6) the different minigenes carrying the deletions in the T6 context were co-trasfected with ASF/SF2.

Overexpression of ASF/SF2 increased the percentage of exon inclusion, from 20% to 55%, in two deleted constructs, T6Δ12-21, and T6Δ47-54 in a similar manner as detected for the single T6 mutant, (Fig. 3.5.3, lanes 4 and 10). The T6Δ35-44 minigene also responded to ASF/SF2 showing full inclusion of exon 18 (Fig. 3.5.3, lane 8). The already fully included T6Δ64-74 minigene was not affected by the increasing amount of ASF/SF2 protein (Fig. 3.5.3, lane 12). The minigene that did not respond to ASF/SF2 overexpresssion, showing complete exon 18 exclusion, was the construct carrying the T6 mutation plus the 23-32 deletion (Fig. 3.5.3, lane 6). This result suggested that the 23-32 element is an enhancer which stimulates BRCA1 exon 18 splicing through ASF/SF2 only in the context of the T6 mutant.

A

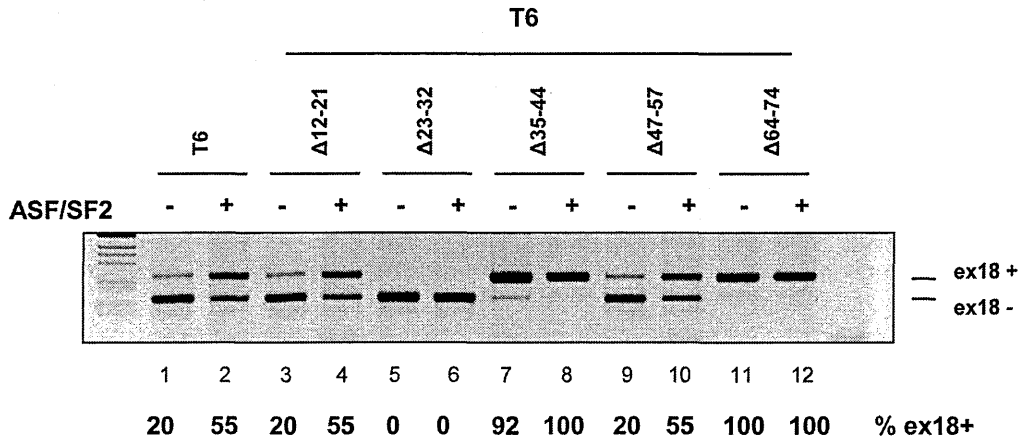

Figure 3.5.3: Effect of ASF/SF2 overexpression on exon 18 T6 deleted minigenes.

(A) The T6 deleted constructs, described before in Fig. 3.5.2, were cotransfected either with the ASF/SF2 coding plasmid (+) or with the corresponding empty vector (-). Splicing patterns were analysed by separating the RT-PCR products on agarose gel. The identity of the transcripts, exon 18 inclusion (ex18+) and exon 18 exclusion (ex18-), is shown schematically at the right side of the gel.

ASF/SF2 overexpression increased the percentage of exon 18 inclusion in all the constructs except for T6 Δ23-32 minigene that remaned totally excluded. The percentages of exon 18 inclusion are reported below each lane.

4 DISCUSSION

Several studies have identified a large number of exonic point mutations that, independently from their effect on the coding sequence, cause splicing defects inducing exon skipping. In several gene systems, like NF1, ATM, and CFTR, such exonic point mutations have been identified as the cause of disease inducing different degrees of exon skipping (Teraoka, Telatar et al. 1999; Ars, Serra et al. 2000; Pagani, Buratti et al. 2003; Pagani, Stuani et al. 2003). Notwithstanding the relative high frequency of this type of defects in human pathology and the growing interest in developing appropriate therapeutic strategies, the mechanism of exon skipping due to exonic mutation is not completely understood. There is evidence that indicates that these types of mutations affect functional exonic splicing regulatory elements (enhancer and silencer), but in some cases the underlying mechanism is still controversial (Cartegni, Chew et al. 2002; Pagani and Baralle 2004).

The natural G6T transversion in BRCA1 exon 18 was originally reported to induce exon skipping and was used as a model to explore the composition of splicing regulatory elements and the mechanism of defective splicing. Krainer and colleagues used this system, by evaluating a limited number of site-directed mutants by *in vitro* splicing assays along with the ESEfinder computer-assisted analysis, to propose a general model of defective splicing due to exonic mutations. In 2001 they suggested that the G6T mutation in BRCA1 exon 18, induces exon skipping because it disrupts an exonic splicing enhancer binding site for the ASF/SF2 protein (Liu, Cartegni et al. 2001). This model was supported by an ESE motif computer assisted analysis of 50 mutations previously reported to induce exon skipping in different human disease genes (more than half of the single-base substitutions reduced or eliminated at least

one ESE high-score motif) (Liu, Cartegni et al. 2001) and on a more detailed analysis of the SMN1/2 exon 7 genes (Cartegni and Krainer 2002). In the case of SMN genes, they originally reported a reduced binding to ASF/SF2 in SMN2 with respect to SMN1 and a series of site-directed mutants whose effect on pre mRNA splicing was compatible with the ESEfinder scores. However, other evidence in the SMN1/2 case later challenged this original observation. A site-directed mutagenesis analysis was more compatible with sequence with an Extended Inhibitory Context (Singh, Androphy et al. 2004). Most importantly, siRNA-mediated depletion of ASF/SF2 did not induce exon skipping of the normal SMN1 gene (Kashima, Rao et al. 2007). These results were followed by the opposite proposal that disease-causing mutations can induce exon skipping through the creation of a novel exonic splicing inhibitory sequences (Kashima and Manley 2003; Kashima, Rao et al. 2007).

To understand the mechanism of defective splicing induced by exonic mutations in this thesis we have extensively evaluated the BRCA1 exon 18 splicing model. Our data does not indicate a major role of ASF/SF2 in the aberrant splicing of this exon, but on the contrary indicate that the T6 mutation creates a silencer element that binds to the inhibitory splicing factors hnRNP A1/A2 and DAZAP1.

4.1 The T6 mutant does not disrupt an ASF/SF2-dependent splicing enhancer.

In this thesis we extensively explored the role of ASF/SF2 in the regulation of BRCA1 exon 18 splicing. Several lines of evidence indicate that this splicing factor is not directly involved in the BRCA1 exon skipping induced by the +6G to T mutation. We evaluated the role of ASF/SF2 by direct binding to the WT and mutant RNA sequences using a modified pull down assay, by performing a systematic site-directed mutagenesis using hybrid minigenes and we further tested the *in vivo* overexpression and depletion of ASF/SF2 on BRCA1 exon 18 splicing efficiency.

The first approach toward the characterization of the putative ASF/SF2 enhancer involved in BRCA1 exon 18 definition, was a classical pull down assay. Using this technique followed by western blot analysis, we found that both WT and T6 mutant seem to bind equally to ASF/SF2 (Fig. 3.1.1B, lanes 2, 3). To exclude technical limitation due to the difficult normalization of RNAs among the samples we set up a modified pull down assay. In this experiment the efficiency of ASF/SF2 binding was carefully evaluated by linking BRCA1 RNA targets to a (UG) repeats, recognized by the TDP43 factor, allowing an accurate normalization of the amount of proteins pulled down. The normalization confirmed the result obtained by the RNAs without the UG tail, detecting an equal signal of ASF/SF2 for both WT and T6 mutant (Fig. 3.1.3B, lanes 2, 3). In addition, the WT and mutant BRCA1 sequences showed a low level of interaction with ASF/SF2 compared with the classical GAA-rich fibronectin EDA enhancer signal (Fig. 3.1.3B lane 1). This strong difference with a classical ASF/SF2 enhancer suggests that the affinity of the splicing factor for WT and mutant BRCA1 sequences may be very low and have a secondary role in the exon 18 definition. To follow this observation we performed siRNA mediated depletion of ASF/SF2. As shown figure 3.1.7 the nearly complete depletion of this splicing factor did not change the splicing pattern in the WT minigene. Delta Ron was used as control which as expected, is unaffected by ASF/SF2 silencing. This result is consistent with a non-essential role for ASF/SF2 in the natural BRCA1 exon 18 definition. In addition, the lack of a strong ASF/SF2 enhancer at the 5' end of BRCA1, disrupted by the T6 mutant, was also shown by single site-directed mutagenesis analysis (Fig. 3.2.1B). This approach is particularly useful to test the composition of the splicing regulatory element in its original context. Site-direct mutagenesis analysis in hybrid minigenes has identified the composition of splicing regulatory elements in several cases: both in exons, (Pagani, Buratti et al. 2003;

Pagani, Stuani et al. 2003) and in introns (Pagani, Buratti et al. 2000). In the case of SMN2 exon 7, a similar analysis showed a sequence with an extended inhibitory complex (Singh, Androphy et al. 2004). Among 21 subsequent point mutations inserted along the 5' end of BRCA1 exon 18 only the natural T6 substitution induced a severe exon skipping and only two (A6 and T4), showed partial exon skipping (Fig. 3.2.1B, lanes 4, 8, 10). Most of the point substitutions had no effect on the normal exon 18 splicing pattern (Fig. 3.2.1B) as well as the complete deletion of the sequence between positions +4 to +9 (Δ 4-9 pBRA construct) (Fig. 3.2.3B, lane 6).

We also used the ESEfinder tool to calculate ASF/SF2 dependent ESE motif scores of all the mutants analysed in order to test a possible correlation between the *in silico* approach versus our *in vivo* results, as previously done for the SMN1/2 systems (Cartegni, Hastings et al. 2006). The comparative analysis performed showed that the ESEfinder ASF/SF2 score of all the substitution performed within exon 18 did not correlate with the splicing pattern obtained by the *in vivo* functional assay (Table 3.2.1). These data showed that simple prediction of an ASF/SF2 high score binding motif in the 5' region of exon 18 does not necessarily result in a strong enhancer element regulated by high binding affinity in the *in vivo* situation. Indeed, ESEfinder program predictions have shown only a partial correlation with the *in vivo* analysis of several gene systems such as dystrophin (Disset 2006), CFTR (Pagani, Buratti et al. 2003; Pagani, Stuani et al. 2003) and SMN2 (Kashima and Manley 2003; Kashima, Rao et al. 2007). Thus, the predictive power of this *in silico* approach has failed to establish a clear connection between splicing and abolishment/creation of a regulatory element, as shown for the BRCA1 exon 18 and the other reported cases. Collectively, these data strongly indicate that ASF/SF2 does not have a major role in BRCA1 exon 18 regulation and that the T6 mutant does not disrupt an ASF/SF2 binding site.

4.2 The T6 mutation creates a sequence with splicing inhibitory function.

As our *in vitro* and *in vivo* analysis indicated that ASF/SF2 is neither involved in the regulation of exon 18 WT nor responsible for the T6 mutant exon skipping. We therefore tested whether the natural mutation creates an exonic splicing silencer. The BRCA1 exon 18 silencer element was mapped by analyzing a series of double site directed mutants in the +4 to +9 region containing the T6 variant. Through functional splicing analysis of the minigenes carrying double mutations we showed that the T6 substitution creates a sequence with preferential silencer properties. The majority of point mutations showed a positive effect on T6 exon 18 splicing, partially or totally restoring its inclusion (Fig. 3.2.2B). In particular the substitutions that affected the central "TAG" triplet, between positions +6 and +8, completely recovered exon 18 inclusion (Fig. 3.2.2B, lanes 9-14) indicating the negative role of this motif in splicing regulation. This ESS sequence, containing the TAG motif, resembled a typical sequence recognized by hnRNP A1 splicing factor. This triplet was previously found in silencer elements described in other gene models (Kashima and Manley 2003; Disset, Bourgeois et al. 2006; Kashima, Rao et al. 2007), in the hnRNP A1 SELEX sequence (Burd and Dreyfuss 1994), and in the screening of a large number of putative silencer elements (Wang, Rolish et al. 2004). Due to the high similarity between the mutation described in the SMN2 exon 7 and the one reported for the BRCA1 exon 18 we have also evaluated the SMN1/2 exon 7 sequences (Kashima and Manley 2003; Kashima, Rao et al. 2007), and the SELEX hnRNP A1 motif (Burd and Dreyfuss 1994)(Fig. 3.2.3A). The SMN2 and SELEX motif sequences embedded in the BRCA1 exon 18 context produced exon skipping further reinforcing the idea that TAG motif plays a central role in the inhibitory activity of the silencer sequence (Fig. 3.2.3B, lanes 4, 5).

4.3 The T6 mutant creates a splicing silencer that binds to hnRNP A1, hnRNP A2 and DAZAP1.

Concerning the trans-acting factors involved in BRCA1 T6 exon 18 splicing, we have characterized three proteins assembled on the sequence carrying the T6 single point mutation. Pull down experiments followed by mass spectrometry analysis allowed us to identify three inhibitory splicing factors that interact individually and specifically with the exonic silencer element created by the T6 mutation: hnRNP A1, hnRNP A2 and DAZAP1 (Fig. 3.3.1). Subsequently, we confirmed the binding and the specific involvement of these proteins in T6 RNA recognition by western blot (Fig. 3.3.2B) and EMSA analysis in which the mutant sequence, differently from the WT, binds to all three different proteins (Fig. 3.3.3C). In addition, the G4U6 mutant RNA sequence showed a stronger complex signal than the one formed by the interaction between the proteins and the U6 RNA, indicating a higher binding affinity that was in agreement with the percentage of exon 18 skipping obtained by the functional splicing assay.

By means of siRNA experiments, we provide functional evidence that DAZAP1 together with the hnRNP A1/A2 is involved and probably have a redundant function in the splicing regulation of the T6 BRCA1 exon 18, reporting also a novel role for DAZAP1 in splicing regulation. The simultaneous treatment with hnRNP A1/A2 siRNA resulted in a significant recovery of the defective splicing of exon 18 carrying the T6 mutation (Fig. 3.4.3A, lane 6). The further depletion of DAZAP1 enhanced the T6 exon 18 recovery (Fig. 3.4.3A, lane 7) showing an involvement of DAZAP1 in the BRCA1 exon 18 splicing reaction. DAZAP1 could participate to the regulation of the BRCA1 T6 exon 18 splicing through the recognition of sequences similar to

the ARE motifs (CUUAGUUU), acting as a negative factor and promoting the exon 18 exclusion.

The requirement of multiple siRNAs for an efficient rescue of defective T6 splicing (Fig. 3.4.3A, lane 7) and the EMSA experiments with purified proteins (Fig. 3.3.3) suggested that both hnRNP A1/A2 and DAZAP1 specifically recognise the inhibitory sequence. This is further supported by the fact that DAZAP1 co-localizes and physically interacts with the splicing repressor proteins hnRNP A1 and hnRNP C1 (Lin and Yen 2006). However, triple treatment with siRNA for hnRNP A1/A2 and DAZAP1 did not completely restore BRCA1 exon inclusion. This could be explained by the presence of residual amounts of the splicing factors, not completely depleted by the siRNAs or by the binding of additional inhibitory splicing factor(s) at the silencer element not revealed by the pull-down analysis.

A similar study on BRCA1 T6 mutant was also recently performed by Kashima and Manley (Kashima, Rao et al. 2007). In agreement with our findings, they reported that ASF/SF2 role is dispensable for the regulation of exon 18 WT and that the depletion of ASF/SF2 has a small negative effect on T6 exon 18 inclusion (Kashima, Rao et al. 2007). Moreover, they performed UV-crossilinking and IP experiments detecting hnRNP A1 binding the T6 mutant but not the WT BRCA1 exon 18 sequence. In their study, DAZAP1 was not detected, maybe due to the different technique used (Kashima, Rao et al. 2007).

4.4 A novel splicing regulatory role of Deleted in Azoospermia Associated Protein 1 (DAZAP1).

The hnRNP A1 and hnRNP A2 proteins are well characterised inhibitory splicing factors involved in a variety of cellular and viral systems (Chabot, LeBel et al. 2003), and have been described as splicing repressors in the regulation of SMN2 exon 7 splicing (Kashima and Manley 2003; Kashima, Rao et al. 2007).

In contrast to hnRNP A1/A2, few data are available regarding the functional splicing properties of the Deleted in Azoospermia Associated Protein 1 (DAZAP1), which has been mainly described to be involved in mRNA transport and stability processes (Lin and Yen 2006) (Morton, Yang et al. 2006). DAZAP1, also named as proline-rich RNA-binding protein (PRRP), was originally isolated through its interaction with Deleted in Azoospermia (DAZ) protein in a yeast two-hybrid assay (Tsui, Dai et al. 2000). DAZAP1 is expressed at high level in testis, where it was identified together with other members of DAZ family, but it was also detected in many other human tissues (Tsui, Dai et al. 2000; Dai, Vera et al. 2001) and in several cells lines (Rousseau, Morrice et al. 2002; Lin and Yen 2006; Morton, Yang et al. 2006).

DAZAP1 shares a structure similar to the hnRNP protein family members with two RNA-binding domains (RBDs) at the N-terminus (Dreyfuss, Matunis et al. 1993; Akindahunsi, Bandiera et al. 2005). The C-terminal portion of DAZAP1 is rich in prolines instead of the glycines described for the hnRNP A/B proteins. Human DAZAP1 shows a high degree of evolutionary conservation with the orthologs isolated from mice and *Xenopus* (Dai, Vera et al. 2001; Zhao, Jiang et al. 2001). It was reported that DAZAP1 binds preferentially *in vitro* to poly(U) and poly(G) and to a lesser degree to poly(A) sequences (Tsui, Dai et al. 2000). Interestingly, SELEX experiments determined the RNA ligand for mouse PrrP that should contain two conserved sequences, AAUAG and $GU_{1-3}AG$ involved in a two-loop structure (Hori,

Taguchi et al. 2005). This motif contains the same UAG splicing sequence that we showed to be important in the silencing of the BRCA1 T6 mutant exon. Moreover, DAZAP1 was identified as one of the proteins able to specifically bind to the ARE element (AU-rich element) in the TNFα (Rousseau, Morrice et al. 2002). It was also shown that DAZAP1 is a physiological substrate for ERK1/ERK2 and might play a role in the regulation of mRNA stability/translation by the classical MAPK cascade (Morton, Yang et al. 2006).

DAZAP1 was detected in the nucleus of somatic cells as a component of the hnRNP particles and specifically co-localized with hnRNP A1 and hnRNP C1 (Lin and Yen 2006). Like other hnRNPs, DAZAP1 has the ability to shuttle between the nucleus and cytoplasm using a small conserved C-terminal segment (Lin and Yen 2006). Its nuclear localization seems to depend on active transcription, when RNA pol II is inhibited, DAZAP1 is retained in the cytoplasm (Lin and Yen 2006).

Taking together the data about DAZAP1, the more probable function of this factor might be connected with different steps of mRNA processing (Lin and Yen 2006; Morton, Yang et al. 2006) and here we provide the evidence of its inhibitory role in splicing regulation of BRCA1 when exon 18 carries a mutation.

4.5 The BRCA1 exon 18 contains additional weak splicing exonic regulatory elements.

To characterize additional splicing regulatory elements we performed a series of short consecutive exonic deletions within the BRCA1 minigene WT and mutant. We identified several splicing regulatory sequences, which interestingly are functionally relevant only in the presence of the T6 mutation (Fig. 3.5.2B) and not in the WT minigene (Fig. 3.5.1B.). Some deletions within the T6 exon 18 context affected the splicing pattern both reducing and increasing the splicing efficiency (Fig. 3.5.2B). We could identify two silencer elements at position 35-44 and 64-74 (Fig. 3.5.2B, lane 5, 7) and only one sequence with enhancer properties in position 23-32 (Fig. 3.5.2B, lane 4). Interestingly, the two regions with silencer properties contain the TAG sequence, which could be responsible along with the other TAG sequence created by the T6 mutant of a cooperative splicing inhibition (see below the model of regulation). To understand if the 23-32 enhancer might be involved in the ASF/SF2-dependent splicing enhancement, we performed overexpression experiments. Only the minigene containing the Δ23-32 deletion did not respond to ASF/SF2 (Fig. 3.5.3A, lane 6), suggesting the presence of a weak ASF/SF2-dependent splicing element in the 23-32 deleted region.

This ASF/SF2 dependent enhancer can explain why ASF/SF2 affected the splicing efficiency in the T6 BRCA1. In fact differently from the WT, the T6 BRCA1 minigene responded to changes in the intracellular concentration of ASF/SF2. The T6 BRCA1 exon 18 was affected both by siRNA-mediated depletion (Fig. 3.1.7B, lane 3) and overexpression of ASF/SF2 (Fig. 3.1.6A, lane 4) by reducing or increasing the percentage of exon inclusion, respectively.

4.6 Models for normal and defective BRCA1 exon 18 splicing.

Our results indicate that the primary determinant of the T6 BRCA1 exon 18 exclusion is not the disruption of an ASF/SF2 dependent enhancer but the creation of a new silencer element recognized specifically by the splicing inhibitory factors hnRNP A1/A2 and DAZAP1.

Based on our analysis we propose a model for WT and T6 BRCA1 exon 18 splicing regulation (figure 4.1). BRCA1 exon 18 WT may contain two weak ESE regions localized at position +4-9 and +23-32 and two downstream silencers at position at position 35-44 and 64-74 and its definition is largely dependent on the 3' and 5' ss (Fig. 4.1A). The two enhancer regulatory elements are dispensable for the definition of the BRCA1 WT exon, as their single deletions did not induce exon skipping (Fig. 3.5.1B). Only when deleted in combination can they induce exon skipping (Fig. 3.5.2B, lane 8) and the weak central enhancer element probably mediates its effect through its binding to ASF/SF2.

The G to T mutation in position +6 of exon 18 induces exon skipping through the creation of a strong silencer regulatory element located at the 5' end of the exon that binds to the inhibitory splicing factors hnRNP A1, hnRNP A2, and DAZAP1. Several mechanisms have been proposed to explain hnRNP A1 mediated splicing repression; our data extended these finding by proposing an involvement of DAZAP1 in these models. In the simplest scenario, silencer-bound hnRNP A1/A2 and DAZAP1 factors can directly antagonize another key splicing factor involved in exon definition. The hnRNP A1, hnRNP A2, and DAZAP1 binding is very close to the 3'ss and this could interfere with the early spliceosome assembly by blocking the access to general splicing factor such as U2AF65, U2AF35 and/or U2 snRNP to the acceptor site (Fig 4. 1b) (Tange, Damgaard et al. 2001). However, it is also possible that the initial binding of hnRNP A1, hnRNP A2 and DAZAP1 at the T6 sequence

could nucleate the formation of a chain of inhibitory proteins along the other two silencer elements located in the middle and at the end of the exon. These two silencers, as the T6 mutant sequence, contain TAG, which might represent the "core" sequences involved in this process. This model for cooperative repression involves the creation of a "zone of silencing" due to the propagation of the repressor proteins that will cover the entire exon 18, leading to its exclusion, as previously suggested for a hnRNP A1 silencer element described in HIV-1 *tat* exon 3 systems (Zhu, Mayeda et al. 2001).

Recently, a concomitant disruption of an exonic splicing enhancer not dependent on ASF/SF2 has been suggested to be involved in the exon skipping induced by the T6 mutant (Kashima, Rao et al. 2007). The presence of this putative enhancer was suggested by the fact that the siRNA-mediated silencing of hnRNP A1/A2 did not have a significant effect on T6 aberrant skipping (Kashima, Rao et al. 2007). We have carefully evaluated this point by performing co-transfection experiments with the SMN2 and T6 BRCA1 minigenes and in both cases we observed that silencing of hnRNP A1/A2 rescued both defective exons (Fig. 3.4.4). We do not have a clear explanation for these differences in the results. An inhibitory effect on the BRCA1 recognition may differ depending on the variable concentration of splicing factors, like DAZAP1, due to the use of different cell types. These factors might efficiently replace hnRNP A1/A2 at the T6 mutant silencer or can act on other inhibitory splicing regulatory elements, thus modulating the siRNA rescue splicing efficiency. However, even if our results clearly indicate that the T6 mutant creates a strong silencer, we cannot exclude the simultaneous disruption of a weak enhancer region present in the 5' end of the BRCA1, as suggested by our double delition analysis (Fig. 3.5.2B lane 8), in agreement with Kashima et al. (Kashima, Rao et al. 2007).

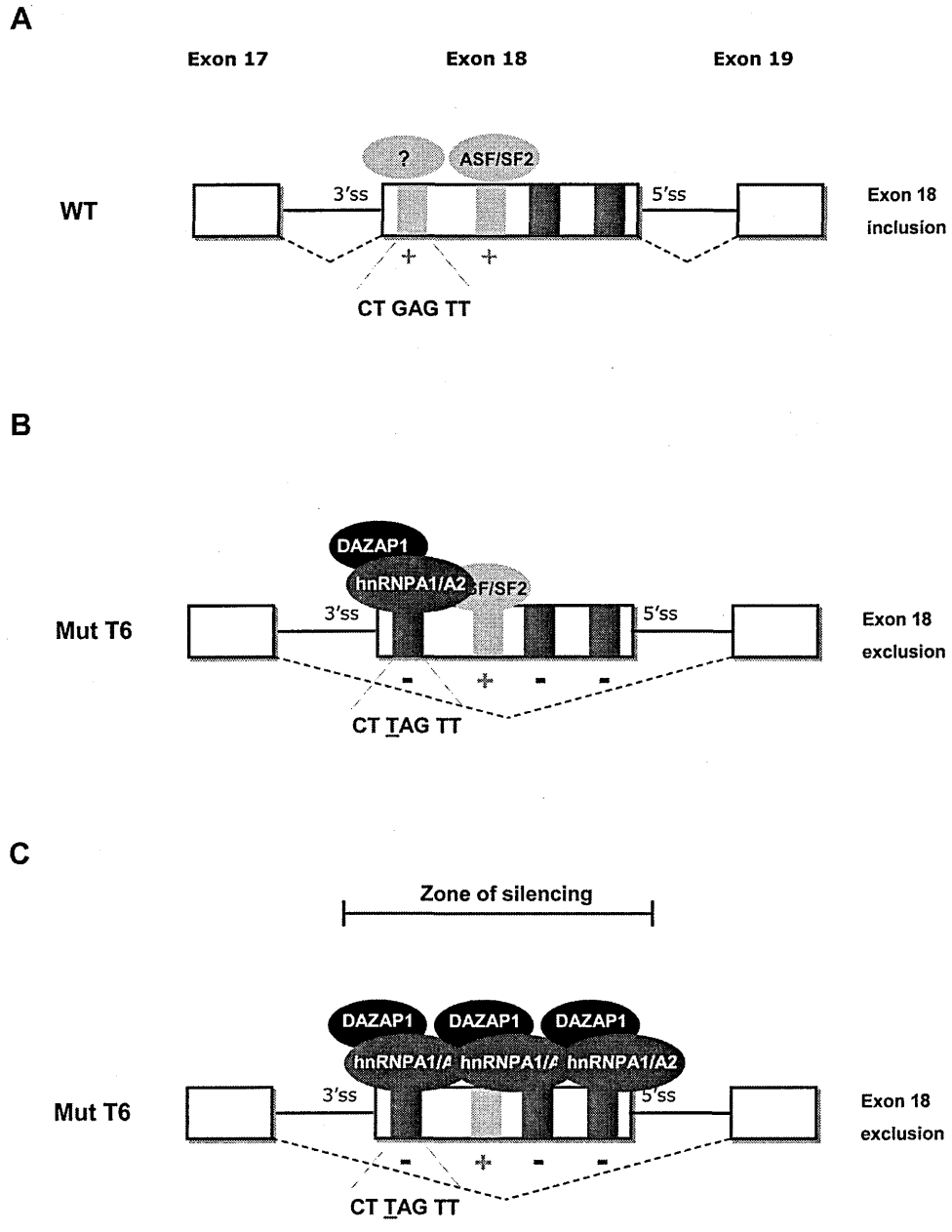

Figure 4.1: Splicing regulatory elements in BRCA1 exon 18 and creation of an ESS by the T6 mutant. (A) WT exon 18 contains two weak ESEs (light grey boxes) and two ESSs (black boxes) and its inclusion is largely dependent on the 3' and 5'ss definition. (B-C) The G to T mutation in position +6 creates a strong silencer element at the 5' end of the exon binding to hnRNP A1/A2 and DAZAP1 and may also disrupt a weak splicing enhancer. The ASF/SF2 ESE located in the middle of exon 18 stimulates BRCA1 exon 18 splicing only in the context of the T6 mutant. The mechanism of inhibition could be mediated by blocking 3'ss recognition (B) or by forming a zone of silencing that covers the entire exon (C).

In fact, the partial exon 18 skipping obtained by the double deletion of the 4-9 and 23-32 regions (Fig. 3.5.2B, lane 8) might be due to the presence of two nearby weak enhancer sequences, which may contribute in a synergic manner to the definition of the WT exon. Thus, with the limitation provided by the gross deletion analysis, which can interfere with RNA secondary structure and acceptor site accessibility or reduce the exon definition below a critical threshold, the 5' end on BRCA1 exon 18 might contain a weak enhancer, which becomes relevant only when the downstream enhancer is deleted.

However, the extent of splicing inhibition mediated by the Δ23-32 deletion in the context of the single T6 substitution was more severe than the splicing inhibition mediated by the presence of the Δ4-9 deletion (Fig. 3.5.2B, lane 4 vs. 8). This result indicated that the inhibitory effect of the T6 mutation is more repressive than the removal of the weak enhancer between position +4-9, further indicating the creation of a silencer element as the main cause of pathological exon 18 skipping.

FUTURE DIRECTIONS

In this thesis the role of exonic regulatory elements in the aberrant splicing of the human BRCA1 exon 18 have been investigated. This study brings some insights into the rather complex regulation of BRCA1 exon 18, whose T6 mutation provoke the substantial exon 18 skipping.

The data collected in this thesis clearly suggest that the T6 BRCA1 exon 18 splicing is modulated by several other cis-acting elements distributed along the exon. These elements were identified mainly by large deletion analysis and accordingly, it will be necessary to better characterize their specific contribution in the regulation of BRCA1 exon 18 splicing. Further experiments will characterize in more detail the nucleotide composition and the binding properties of these sequences. In addition, it would be important to explore the potential role of intronic regions, which, as described for the SMN case (Miyajima, Miyaso et al. 2002; Miyaso, Okumura et al. 2003; Singh, Singh et al. 2006; Kashima, Rao et al. 2007) might contribute to the splicing regulation.

In this thesis a novel role for DAZAP1 in the pre-mRNA splicing regulation has been described. An obvious question that needs to be addressed is the involvement of this splicing factor in the regulation of the similar SMN1/2 model. The behavior of BRCA1 T6 exon 18 mutation in fact, is analogous to that which promotes exon skipping in SMN2 exon 7. It will be interesting to compare the two models, in particular a possible role of DAZAP1 in the SMN2 exon 7 splicing regulation. DAZAP1 might be also involved in the control of other alternative splicing events. Using an exon array in DAZAP1 depleted cells it will be possible to study the role of this protein in general and/or in specific splicing events.

Co-transfection experiments suggested that the 23-32 region of the BRCA1 exon 18 is an enhancer element connected to the ASF/SF2 factor. This hypothesis will be further explored by performing binding experiment in order to confirm the presence of ASF/SF2 factor in the 23-32 exonic region. Moreover, the result obtained by the double deleted construct (Δ4-9/Δ23-32) supports the notion of a weak ESE in the vicinity of position +6, even if not related to ASF/SF2. Additional trans-acting factors could be involved in the recognition of this weak enhancer and their presence will be investigated by binding experiments.

A more general mechanistic problem will also be considered in order to understand the role of hnRNP A1, -A2 and DAZAP1 complex in the splicing inhibition. This binding may inhibit the recognition of the 3'ss by U2AF35 or in interfering with subsequent steps of spliceosome assembly. A set up of an *in vitro* system will allow the study of the different splicing steps. It is also possible, as indicated in the model (Fig. 4.1), that a cooperative binding of hnRNP A1/A2 and DAZAP1 covers the entire exon 18 and produces exon skipping. To this aim, the potential contribution of other silencers containing the TAG motifs, localized within the last part of exon 18, is going to be checked. The binding of hnRNP A1, -A2 and DAZAP1 to the first inhibitory element, surrounding the position +6, might promote a subsequent cooperative binding to the other putative ESSs localized at the end of the exon 18.

5 MATERIALS AND METHODS

Chemical Reagents.

General chemicals were purchased from Sigma-Aldrich, Merck, Gibco BRL and Boehringer.

Standard solutions.

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

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

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

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

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

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

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

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

DNA preparation.

Small scale preparation of plasmid DNA from bacterial cultures.

Single bacterial colony was selected and transferred into 3 ml of LB medium containing the ampicillin in a loosely capped 15-ml tube. The culture was incubated overnight at 37°C with vigorous shaking. Rapid purification of small amounts of recombinant plasmid DNA was performed using the method based on alkaline lysis of recombinant bacteria previously described by Sambrook (Sambrook 1989).

Briefly, the bacterial pellet was resuspend in 100 µl of solution I (10 mM EDTA pH 8.0, 25 mM Tris-HCl pH 8.0, 50 mM glucose). Then 200 µl of solution II for lysis (0.2 M NaOH, 1% SDS) was added and the contents mixed by inversion. Lysis was then stopped with 150 µl of solution III for neutralization (60 ml of 5M potassium acetate, 11.5 ml of glacial acetic acid, 28.5 ml H_20) and the content was mixed by inversion. The bacterial lysate was then centrifuged in an Eppendorf microcentrifuge at maximum speed for 5 minutes and the supernatant transferred to a new tube. An equal volume of 1:1 v/v phenol: chloroform solution was added to the supernatant. The tube was then vortexed and centrifuged as above. The step was repeated with aqueous phase and the equal volume of chloroform. The aqueus phase containing DNA was then recovered and the DNA was precipitated by adding 2 volumes of ethanol. The final pellet was resuspended in 47 µl of dH_20 with 3 µl of RNase A (Sigma Chemicals Ltd.) from a stock of 20 mg/ml prepared in sterile water and boiled for 10 minutes to destroy trace amounts of DNase activity. The DNA sample was then digested at 37°C for 1 hour in order to eliminate all the RNA from the sample. Routinely, 2-3 µl of such DNA preparation were taken for restriction enzyme digestion and sequence analysis.

Large scale preparation of plasmid DNA from bacterial cultures.

Large-scale and pure preparations of plasmid DNA were necessary in order to perform different kind of experiments like transfection, in vitro transcription and vectors/inserts preparation. JetStar purification kit (Genomed) was used according to the manufacturer's instructions. In order to get a good amount of plasmid, we used an overnight bacterial culture of 50 ml in LB medium.

Genomic DNA preparation.

Peripheral blood (5 ml) was treated with 20 ml of Lysis Buffer (Sucrose 0.32 M, Tris-HCl 10 mM pH 7.5, $MgCl_2$ 5 mM, Triton X-100 1%) and incubated on ice for 15 min. The mix was centrifuged 10 min at 1000 g at 4 °C. The pellet was washed three times with Fisio Buffer (NaCl 0.075 M, EDTA 0.025 M pH 8) and resuspended in 1 ml of Resuspension Buffer (Tris-HCl 10 mM pH 7.5-8, NaCl 0.4 M, EDTA 2 mM). 200 µl of SDS 10% were added and the mixture was incubated at 37 °C overnight to facilitate the extraction. The following day 600 µl of NaCl saturated water solution was added and the samples were mixed for 15 seconds to allow the protein precipitation. The mix was centrifuged 15 min at 1500g to pellet the proteins. The supernatant was recovered and the centrifugation repeated a second time. Finally, 1 volume of isopropanol was added to the supernatant to precipitate the DNA. Precipitated DNA was removed with a Pasteur, washed in 1 ml of 70% ethanol and resuspended in 500 µl of water. The DNA was then checked by electrophoresis on a 0.8% w/v agarose gel, quantified by spectrophotometer and stored at -20°C until used for PCRs.

RNA preparation from cultured cells.

Cultured cells were washed two times with 1X PBS and then RNA TRI reagent, purchased from Ambion, was added according to the manufacturer's instructions. Chloroform extraction was performed and supernatant precipitation was obtained by adding isopropanol. Then we rinsed the pellet in 70% ethanol. The final pellet was resuspended in H_20 and stored at -80°C. The RNA quality was checked by electrophoresis on a 1% agarose gel. If it was the case half of the resuspended RNA could be treated with DNase RNase-free (Roche Diagnostic) in 1X DNase buffer. The digest was incubated at 37°C for 30 minutes, and then the RNA was purified by

acid phenol extraction and precipitated by adding 2 volumes of ethanol and 1/10 volume of 3M NaOAc pH 5.2. The final pellet was resuspended in H_2O and frozen at –80 °C. The RNA quality was checked again by electrophoresis on 1% agarose gels.

Estimation of nucleic acid concentration.

The concentration of a DNA or RNA sample can be checked by the use of UV spectrophotometry. Both RNA and DNA absorb UV light very efficiently making it possible to detect and quantify either at low concentrations. The nitrogenous bases in nucleotides have an absorption maximum at about 260 nm. An optical density of 1.0 at 260 nm is usually taken to be equivalent to a concentration of 50 μg/ml for double stranded DNA, 40 μg/ml for single stranded DNA and RNA, and approximately 20 μg/ml for single-stranded oligonucleotide sample. The ratio of the absorbance at 260 nm/ absorbance at 280 nm is a measure of the purity of a sample; it should be 1.8 for pure sample of DNA and 2 for RNA. These values are reduced by protein contamination (Sambrook 1989).

Enzymatic modification of DNA.

Restriction enzymes

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

digestions of vectors and inserts were made of 5-10 μg DNA using the appropriate condition needed by the restriction enzyme in 100-200 μl reaction volume. Enzymatic activity was then removed either by heat inactivation or by phenol-chloroform extraction.

DNA Polymerase I, Large (Klenow) fragment.

Klenow enzyme was used to treat PCR products for blunt-end creation by 3' overhang removal and 3' recessed end fill-in. Briefly the DNA was dissolved in any 1X restriction enzyme NEBuffer or 1X EcoPol Reaction Buffer and supplemented with 33 μM each dNTP if "fill-in" was required (DNA fragments with protruding 3'ends). Then, 1U of Klenow per μg DNA was added and the mixture was incubated 15 minutes at RT. The reaction was inactivated by by heating at 75°C for 20 minutes.

T4 Polynucleotide Kinase.

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

Calf intestal phosphatase (CIP).

Calf intestal phosphatase (CIP), provided from New England Biolabs Inc, catalyzes the removal of 5´ phosphate groups from DNA and RNA. Since CIP-treated fragments lack the 5´ phosphoryl termini required by ligases, they cannot self-ligate.

This property was used to decrease the vector background in cloning strategies. The standard reaction was carried out in a final volume of 50-100 µl using 1U of enzyme per 0.5 µg DNA at 37°C for 1 hour.

T4 DNA ligase

T4 DNA ligase catalyzes the formation of phosphodiester bonds between neighbouring 3´-hydroxyl- and 5´-phosphate ends in double-stranded DNA, requiring ATP as a cofactor in this reaction. This enzyme was purchased from Roche Diagnostic and was used to join double stranded DNA fragments with compatible sticky or blunt ends. A standard reaction comprises 20 ng of linearised vector and 5-10 fold molar excess of insert in a total volume of 20 µl containing 1X ligase buffer and 1U of enzyme. Reactions were carried out at room temperature for at least 2-4 hours. In some ligations synthetic oligonucleotides were used as inserts for the reactions. In these cases, amounts added to each reaction to obtain ligation of the oligonucleotides in the resulting plasmids were about 50-100 fold molar excess over the DNA vector.

Agarose gel electrophoresis of nucleic acid.

DNA samples were size fractionated by electrophoresis in agarose gels ranging in concentrations from 0.8 % w/v (large fragments) to 2 % w/v (small fragments). The gels contained ethidium bromide (0.5 µg /ml) and 1X TBE solution. Horizontal gels were used for fast analysis of DNA restriction enzyme digests, estimation of DNA concentration, or DNA fragment separation prior to elution from the gel. A fast analysis of RNA samples was obtained by using a 1% agarose gel. The gels were electrophoresed at 50-80 mA in 1X TBE running buffer for a time depending on the

fragment length expected and gel concentration. DNA was visualized by UV transillumination and the result recorded by digital photography.

Elution and purification of DNA fragments from agarose gels

This protocol was used to purify both vectors and inserts of DNA for cloning purpose. The DNA samples were electrophoresed onto an agarose gel as described previously. The DNA was visualized with UV light and the required DNA fragment band was excised from the gel. This slab was cut into pieces, and the QIAquick Gel Extraction Kit was used according to the manufacturer's instructions. Briefly, 3 volumes of Buffer QX1 were added to 1 volume of gel for DNA fragments and then incubated at 50 °C for 10 min, vortexing every 5 min. After the gel slice was dissolved completely 1 gel volume of isopropanol was added to the sample. The mixture was loaded into a prepared column and centrifuged at maximum speed for 1 min. The flowthrough was discarded. One washing step was performed using 750 µl of PE buffer. The bound DNA was eluted with 30-50 µl of sterile water. The amount of the recovered DNA was approximately calculated by UV fluorescence of intercalated ethidium bromide in agarose gel electrophoresis.

Amplification of selected DNA fragments.

The polymerase chain reaction was performed using both genomic or plasmid DNA as templates and following the basic protocols of Roche Diagnostic Taq DNA polymerase. The volume of the reaction was 50 µl and comprised: 1x Taq buffer, dNTP mix (100 µM each), oligonucleotide primers (100 nM each), Taq DNA polymerase 2.5-5 U and 100 ng of genomic DNA or 0.1 ng of plasmid. All the synthetic DNA oligonucleotides used for PCR amplification were purchased from Sigma Aldrich (http://orders.sigma-genosys.eu.com).The standard amplification

conditions were the following: 94°C for 3 min for the initial denaturation, 94°C for 45 s, 54°C for 45 s, 72°C for 45 s for 30 cycles and 72°C for 10 min for the final extension. PCRs were optimized to be in the exponential phase of amplification and products were routinely fractionated in 2% (w/v) agarose gels.When different PCR conditions were used; they are described in each particular case. The amplifications were performed on Cetus DNA Thermal Cycler (Perkin Elmer) or on a Gene Amp PCR System (Applied Biosystem).

Sequence analysis for cloning purpose.

Sequence analysis of plasmid DNA was performed using CEQ 2000 sequencer (Beckman Coulter).The plasmid DNA of interest (approximately 500 ng) was purified through a MicroSpin S-400 HR Column (Amersham Pharmacia Biotech). The DNA was then amplified by asymmetric PCR (only one primer) using fluorescent labeled dideoxy nucleotide terminations according to the manufacturers' instructions. The samples were analyzed by loading them into the automatic sequencer. The first two primers listed were specific for the pBluscript KS plasmid whereas the other oligonucleotides were annealing in different position along the minigene sequence.

- **UNIVERSAL Dir**

5'-GTTTTCCCAGTCACGAC-3'

- **UNIVERSAL Rev**

5'-GGAAACAGCTATGACCATG-3'

- **BRC 281 Dir**

5'-TCAGCCTCTGATTCTGTCACCAGG-3'

- **BRC 350 Dir**

5'-TGAGGCTCTTTAGCTCCTTAGG-3'

- **BRC 474 Rev**

5'-CCTAGAAAATATTTCAGTGTCCGTTCAC-3'

- **BRC 616 Rev**

5'-GGTAACTCAGACTCAGCATCAGCAA-3'

Bacterial culture

The *E.coli* DH5α strain was used for ligation transformation, plasmid amplification and growth. The *E.coli* BL12 DE3 (Novagen) bacteria were used for protein production. Both bacteria strains were maintained in short term as single colonies on agar plates at 4°C and for long term storage on glycerol stocks (sterile glycerol to a final 30 % v/v concentration to liquid bacterial cultures) stored at -80°C. When necessary, from glycerol stocks the bacteria were plated and amplified by an overnight incubation in Luria-Bertani medium (per liter: 10 g Difco Bactotryptone, 5 g Oxid yeast extract, 10 g NaCl, pH 7.5). Bacterial growth media were sterilized before use by autoclaving. Ampicillin (Sigma-Aldrich) was added to media to a final concentration of 100 μg/ml.

Preparation of bacterial competent cells.

Bacterial competent cells of both strains were prepared following the method described by Chung (Chung, Niemela et al. 1989). *E.coli* strains were grown overnight in 10 ml of LB at 37°C. The following day, 200 ml of fresh LB were added and the cells were grown at room temperature for 2-3 h until the OD600 was 0.3 – 0.4. Then the cells were put in ice and centrifuged at 4°C and 1000 g for 15 minutes. The pellet was resuspended in 1/10 volume of cold 1X TSS solution (10% v/v PEG, 5% v/v DMSO, 35 mM $MgCl_2$, pH 6.5 in LB medium). The cells were divided into aliquot, rapidly frozen in liquid nitrogen and stored at -80°C. Competent cells

efficiency was tested using 0.1 ng of pUC19 control DNA plasmid. Transformation efficiency should be greater than 1 x 10^6 transformants/μg pUC19 DNA. So the competence was deemed satisfactory if this procedure resulted in more than 100 colonies.

Transformation of bacteria

Transformations of ligation reactions were performed using half of the initial reaction volume. Transformation of positive clones was carried out using 1ng of the DNA plasmid. Briefly the transformation protocol consisted in 30 min of incubation on ice of DNA with 50 μl of competent cells followed by a heat shock at 42°C for 2 minutes. The cells were placed again on ice for 2 minutes and then spread onto agarose plates containing the appropriate antibiotic concentration (100 μg/ml of ampicillin). The plates were incubated for 12-15 hours at 37°C in order to allow the colonies formation. When the DNA inserts were cloned into β-galactosidase-based virgin plasmid (pUC19, Bluscript KS), 30 μl of IPTG 100 mM and 30 μl of X-gal (4% v/v in dimethylformamide) were prepared onto the surface of the agarose before plating to facilitate the screening of positive clones (white colonies) through identification of β-galactosidase activity (blue colonies).

Hybrid minigene constructs.

Transient transfection of minigenes is a commonly used *in vivo* technique to identify the different features of exon regulation during splicing process. It is a good method to study specific cis-acting elements that control constitutive and alternative exons, the cell-specific splicing pattern and also to identify trans-acting factors that recognize these elements and regulate splicing (Cooper 2005).

A minigene is a "simplified" version of a gene and usually contains a genomic region from the gene of interest that includes the alternatively spliced part and its flanking genomic seqences. The genomic segment was generated by PCR amplification directly from target DNA as template and the oligonucleotides used has to contain the restriction enzyme sites at their ends that match restriction sites in the recipient plasmid (Cooper 2005). This system, after transient transfection and RNA analysis, allowed us to study the splicing outcome.

The minigenes used derived from a basic hybrid construct containing exons from α-globin and fibronectin, under the control of the α-globin promoter. The intronic region between the two fibronectin exons contains a unique NdeI site in which the target genomic segment under study can be inserted. Moreover the two junctions between the α-globin exon 3 and the fibronectin exons have a BstEII site. This region in the middle of the two restriction sites can be removed and substituted with a large genomic region of interest. In this thesis two minigenes have been created: the pTB and pBRA. This has been done through PCR amplification of the exon under study with its intronic flanking regions, treatment with Klenow-Kinase enzymes as described previously, purification from agarose gel and cloning in pBluscript KS vector into the restriction site SmaI. Following sequencing, in order to check the nucleotide sequence to be correct, the subcloning in the minigenes was performed using the enzyme restriction site NdeI or BstEII, depending on the case.

The minigene contains at the 5' end the α-globin gene promoter and SV40 enhancer sequences to allow polymerase II transcription in the transfected cell lines. The reporter minigene is composed of α-globin (black boxes) and fibronectin exons (shaded boxes), whiles at the 3' end contains a functionally competent polyadenylation site, derived from the α-globin gene. The unique NdeI and the BstEII sites are indicated.

pTB minigene.

The BRCA1 genomic region encompassing exon 18 and adjacent intronic sequences (289 bp of intron 17, full length exon 18 and 255 bp of intron 18) was amplified with BRC138 Dir and BRC760 Rev primers and blunt ligated in the SmaI site of pBluscript KS to generate pBS BRA18. Following sequencing in order to check the nucleotide sequence to be correct, the subcloning in the pTB minigene was performed using the enzyme restriction site NdeI. The the α-globin, the fibronectin and BRCA1 18 exons and are in frame.

- **BRC 138 Dir**

5'-GGCATATGGAGATCTATAGCTAGCCTTGGCGTCTAGAAGATGG-3

- **BRC 760 Rev**

5'-AATCCCTAGCTCATATGCTAACATTGCTAGG-3'

pBRAminigene.

The BRCA1 genomic region containing the last 85 bp of exon 17 and first 1408 bp of intron 17 was amplified from normal DNA using BRC90BstEII Dir and BRC1566 Rev primers and blunt ligated in the SmaI site of pBluscript KS to generate pBS BRA17. Similarly, the last 334 bp of intron 18 along with 58 bp at the beginning of exon 19 were amplified with BRC735 Dir and BRC4451 Rev primers and blunt ligated in SmaI site of pBluscript KS to generate pBS BRA19.

The two plasmids, pBS BRA 17 and 19, were sequenced in order to check the identity of the DNA sequence. Subsequently, pBS BRA19 was digested with NdeI and KpnI and cloned in the corresponding sites of pBS BRA17 to generate pBS BRA17-19, with NdeI as unique site in the middle. The BstEII- BstEII fragment contained in pBS BRA17-19 was then inserted in the unique BstEII site of a modified pTB α-globin minigene to obtain pBRAint. pBS BRA18, described

previously, was digested and subcloned in the pBRAint minigene in the unique NdeI site creating the final minigene named pBRA. The final minigene lacks the central part of intron 17 (1960 bp) and the α-globin and BRCA1 17, 18 and 19 exons are in frame.

- **BRC 90 Dir**

5'- CTGGTACCAAGTTTGCCAGAAAACACCACATCACTTTAACTAATC -3'

- **BRC 1566 Rev**

5'- AACACCCAGAGGTCTCC -3'

- **BRC 735 Dir**

5'- TAGCAATGTAGCATATGAGCTAGGGATTTA -3'

- **BRC 4451 Rev**

5'-AACATCAAGTACTTACCTCATTCAGC-3'

PstI – EcoRI cassette.

On both minigenes an EcoRI site was inserted in intron 18 (49 bases downstream the 5'ss of exon 18) by overlapping PCR using as template pBS BRA18. As external primers were used the Universal Dir and Rev and as internal the following oligonucleotides:

- **BRC 560 EcoRI Dir**

5'-AACACCTCAGAATTCCATTTTTACACCTAACG-3'

- **BRC 560 EcoRI Rev**

5'-GGTGTAAAAATGGAATTCTGAGGTGTT-3'

The final amplified fragment, obtained after a second round of PCR performed using the external oligos, was NdeI digested and cloned into the two minigenes described above, pTBWT or pBRAWT minigenes. The natural PstI (next to the 3'ss of exon

18) and the artificial EcoRI unique sites create a cassette system, which was used to facilitate the subsequent cloning procedures.

pBRA minigenes variants.

All pTB and pBRA constructs carrying point, double or complex mutations were generated by site-directed mutagenesis using standard PCR conditions. All the nucleotide substitutions were produced using as PCR template pBRA-WT hybrid minigene, as primer sense different degenerate oligonucleotides and as antisense primer the following:

- **BRC 616 REV**

5'-GGTAACTCAGACTCAGCATCAGCAA-3'

Most of the different PCR products were gel purified and normally PstI-EcoRI digested and directly cloned in pBRA-WT hybrid minigene or in pTB WT plasmid. In some cases digestion was carried out by using different restriction enzymes mentioned below. The presence of the mutations in all the constructs was checked by a CEQ2000 sequencer (Beckman Coulter, Fullerton, CA) according to the manufacturer's instructions and using the primers described before.

Generation of point mutations within the BRCA1 exon 18 regulatory region.

The sense oligos used to introduce point mutations in BRCA1 exon 18 regulatory region were the following:

- **BRC CD4 Dir (D=A, G, T).**

5'-TCATTCTGCAGATGDTGAGTTTGTGTGTGAA-3'

- **BRC T5V Dir (V=A, C, G).**

5'-TCATTCTGCAGATGCVGAGTTTGTGTGTGAA-3'

- **BRC G6H Dir (H=A, C, T).**

5'-TCATTCTGCAG ATGCTHAGTTTGTGTGTGAA-3'

- **BRC A7B Dir (B=C, G, T).**

5'-TCATTCTGCAGATGCTGBGTTTGTGTGTGAA-3'

- **BRC G8H Dir (H=A, C, T).**

5'-TCATTCTGCAGATGCTGAHTTTGTGTGTGAA-3'

- **BRC T9V Dir (V=A, C, G).**

5'-TCATTCTGCAGATGCTGAGVTTGTGTGTGAA-3'

- **BRC T10R Dir (R=A, G).**

5'-TCATTCTGCAGATGCTGAGTRTGTGTGTGAA-3'

- **BRC 11G Dir**

5'-TCATTCTGCAG ATGCTGAGTTGGTGTGTGAA-3'

Generation of double point mutations within the BRCA1 exon 18 regulatory region.

The sense oligos used to introduce double point mutations, in the BRCA1 exon 18 regulatory region were the following:

- **BRC 4D6T Dir (=A, G, T).**

5'-TCATTCTGCAGATGRTTAGTTTGTGTGTGAA-3'

- **BRC 5V6T Dir (V=A, C, G).**

5'-TCATTCTGCAGATGCVTAGTTTGTGTGTGAA-3'

- **BRC 6T7B Dir (B=C, G, T).**

5'-TCATTCTGCAGATGCTTBGTTTGTGTGTGAA-3'

- **BRC 6T8H Dir (H=A, C, T).**

5'-TCATTCTGCAGATGCTTAHTTTGTGTGTGAA-3'

- **BRC 6T9V Dir (V=A, C, G).**

5'-TCATTCTGCAGATGCTTAGVTTGTGTGTGAA-3'

- **BRC 10A11G Dir**

5'-TCATTCTGCAGATGCTGAGTAGGTGTGTGAA-3'

Generation of complex mutations within the BRCA1 exon 18 regulatory region.

The sense oligos used to introduce complex point mutations, in the BRCA1 exon 18 regulatory region were the following:

- **BRC SMN1/2 Dir**

5'-TCATTCTGCAGATGCTYAGACAGTGTGTGAACGGACACTGAAATATT3'

- **BRC 6hnRNP A1selex Dir**

5'-TCATTCTGCAGATGCTTAGGGAGTGTGTGAACGGACACTGAAATATT3'

Generation of small overlapping deletions within the BRCA1 exon 18.

Most of the small overlapping deletions introduced within exon 18 sequence were achieved by a unique PCR amplification using the Dir primers listed below and the BRC 616 Rev oligo.

- **BRC Delta 4-9 Dir**

5'-TCATTCTGCAGATGTTGTGTGTGAACGGACACTGAAATATTTTCTAG-3'

- **BRC Delta 12-21 Dir (K=G, T).**

5'-TCATTCTGCAGATGCTKAGTTTGGACACTGAAATATTTTCTAG-3'

- **BRC Delta 23-32 Dir (K=G, T).**

5'-TCATTCTGCAGATGCTKAGTTTGTGTGTGAACGTATTTTCTAGGAATT GCG-3'

- **BRC Delta 35-44 Dir (K=G, T).**

5'-TCATTCTGCAGATGCTKAGTTTGTGTGTGAACGGACACTGAAATAATT GCGGGAGGAAAATG-3'

- **BRC Delta 4-9/23-32 Dir.**

5'-TCATTCTGCAGATGTTGTGTGTGAACGTATTTTCTAGGAATTGCG-3'

The two delta constructs containing the deletion (~10 bp) within the exon 18 between 47-57 and 64-74 regions were created by overlapping PCR using suitable primers on pBRA WT or T6 minigenes as templates. The first round of PCRs was performed using two sets of primers, combining an external primer with an internal one. The external oligos were BRC218 Dir and BRC 616 Rev described before. The internal primers used were the following:

- **Delta 47-57 Dir**

5'-TATTTTCTAGGAATAATGGGTAGTTAGCTATTTC-3'

- **Delta 47-57 Rev**

5'-ACTACCCATTATTCCTAGAAAATATTTCAGTGTCCG-3'

- **Delta 64-74 Dir**

5'-GGAGGAAAATGGGTTCTGTAAGTATAATACTAT-3'

- **Delta 64-74 Rev**

5'-TATACTTACAGAAACCCATTTTCCTCCCGCAATTCCTAG-3'

The second round of PCR was carried out using the external primers and as template the two amplified products, obtained from the previous two separated PCRs, mixed together. All the inserts obtained were cloned by the standard PstI-EcoRI cassette procedure.

Eukaryotic cell lines.

The cell line used for transfection and cotransfection experiments was Hep3B cells: human hepatocellular carcinoma fibroblast-like. The cell line used for siRNA experiments was HeLa cells derived from cervical cancer cells.

Maintenance and analysis of cells in culture.

Both Hep3B and HeLa cells were grown in Dulbecco's Modiefied Eagle Medium (with glutamine, sodium pyruvate, pyriodoxine and 4.5 g/l glucose) supplemented with 10% fetal calf serum (Euro Clone) and antibiotic Antimycotic (Sigma) according to the manufacturer's instruction. A standard 100mm dish containing a confluent monolayer of cells was washed with 1XPBS solution, treated with 1-2 ml Trypsin (PBS containing 0.045 mM EDTA and 0.1% trypsin) and incubated at 37°C for 2 minutes or until cells were dislodged. After adding 5 ml of media, cells were precipitated by centrifugation and resuspended in pre-warmed medium. A subcultivation ratio of 1:4 to 1:6 was used for Hep3B and a ratio of 1:6 to 1:8 was performed for HeLa cells. 1 ml of this cell dilution was added to 10 ml fresh medium and plated in a new 100 mm dish.

Transfection of recombinant DNA.

Hep 3B 3 x 10^5 cells were plated as described above into 35 mm tissue culture dishes in order to reach a final confluence of 40-70%. The DNA used for transfection was prepared with JetStar purification kit (Genomed) as previously described. Liposome-mediated transfections human hepatocarcinoma cells were performed using DOTAP Liposomal Transfection Reagent (Alexis Biochemicals). For 35mm dish we prepared separately 1.5 µg of plasmid in final 50 µl of H_20 volume and 3 µl of lipofectin in 47µl H_20 for each sample. Then we mixed together the DNA and the lipofectin reagent to reach a final volume of 100µl. This mixture was incubated at room

temperature for 15-20 minutes to allow the formation of DNA-liposome complexes and then added to the cells in 2 ml of serum free culture medium. After 12-18 hours the medium was replaced with fresh medium supplemented by 10% FBS, 24 hours later, the cells were washed with 1X PBS and harvested. Each transfection experiments were repeated at least 3 times.

Co-transfection.

For co-transfection experiments, 1.5 μg of the minigene plasmid was co-transfected with 1 μg of the plasmid carrying cDNA coding for different proteins. In detailed ASF/SF2 and hnRNP A1 coding sequence were cloned into a pCG vector (a kind gift from Dr J. Caceres) wherease hnRNP A2 and DAZAP1 cDNA were inseterted into pFlag vector.When necessary 0.5 μg of the pCG or pFlag empty vector was used as a control.

mRNA functional splicing analysis.

cDNA synthesis.

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

PCR analysis.

PCR was carried out as described previously using standard cycles in a final reaction volume of 30 µl. Oligonucleotides used for that PCR analysis were those specific for the minigene system used or for the endogenous RNA of interest. To amplify only the messengers derived from the pTB plasmid, the primers Alfa2-3 Dir and Bra2 Rev were used because specific for the junction between the alfa-globin exons 2-3 and the EDB +1 exon respectively. For the pBRA minigene the PCR was performed with BRC90BstDir and Glo800Rev which hybridize to BRCA1 exon 17 and α-globin exon 3 sequences, respectively.

- **ALFA 2-3 Dir**

5'-CAACTTCAAGCTCCTAAGCCACTGC-3'

- **BRA 2 Rev**

5'-TAGGATCCGGTCACCAGGAAGTTGGTTAAATCA-3'

- **BRC 90 DIR**

5'-CTGGTACCAAGTTTGCCAGAAAACACCACATCACTTTAACTAATC-3'

- **GLO 800 REV**

5'-GCTCACAGAAGCCAGGAACTTGTCCAGG-3'

In the case of SMN2 plasmid the standard PCR conditions were changed according to Cartegni work (Cartegni, Hastings et al. 2006). Briefly, the amplification conditions were the following: 94°C for 3 min for the initial denaturation, 94°C for 30 s, 58°C for 1 min, 72°C for 1min for 30 cycles and 72°C for 10 min for the final extension. The oligonucleotides sequences were taken from the same published work (Cartegni, Hastings et al. 2006).

- **pCIFwdB**

5'-GACTCACTATAGGCTAGCCTCG-3'

- **SMN8-300+5'R**

5'-AAGTACTTACCTGAAATCTAATCCACATTCAAATTTTCTCAACTG-3'

In the case of the endogenous RON amplification, the standard PCR was modified according to the previous published work (Ghigna, Giordano et al. 2005). Briefly, the number of cycles was increased up to 40 using these conditions: 94°C for 3 min for the initial denaturation, 94°C for 45 s, 56°C for 45 s, 72°C for 45 s and 72°C for 10 min for the final extension. The oligonucleotides sequences were described previously (Ghigna, Giordano et al. 2005).

- **DELTA Ron 2507 Dir**

5'-CCTGAATATGTGGTCCGAGACCCCCAG-3'

- **DELTA Ron 2991 Rev**

5'-CTAGCTGCTTCCTCCGCCACCAGTA-3'

The analysis of the amplified products, 5 ul of each PCR reaction, was performed by gel electrophoresis on a 1.5% or 2% agarose gel. The identity of the resulting splicing products was verified by direct sequence analysis. ImageJ 1.38 software (http://rsb.info.nih.gov/ij/) was used in order to quantify the ratio between exon inclusion/skipping. The variability among different experiments was always <20%.

In silico predictions.

In silico analysis was performed using the free access ESEfinder (http://rulai.cshl.edu/tools/ESE/)

RNA binding protein analysis.

Denaturing polyacrylamide gel electrophoresis (SDS-PAGE).

Protein samples were added to protein sample buffer (2X final). Conventional slab gel SDS PAGE (Laemmli 1970) was performed in vertical gels with the required

percentage of polyacrylamide (37.5:1 acrylamide:bis-acrylamide, ProtoGel, National Diagnostics), depending on each case. The gels were run at 40 mA in 1X SDS-PAGE running buffer. After running, gels were either stained with coomassie Blue R250 in methanol-water-acetic acid 45:45:10 (v/v/v) or electroblotted onto PVDF membrane (Amersham Biosciences) for Western Blot analysis.

Affinity purification of RNA binding proteins.

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

- **Exon 18 WT**

5'- UGCAGAUGCUGAGUUUGUGU -3'

- **Exon 18 U6**

5'- UGCAGAUGCUUAGUUUGUGU -3'

Briefly 12 µg of target BRCA1 RNA oligos were placed in 400 µl of reaction mixture (100 mM NaOAC, pH 5.0 and 5mM sodium m-periodate), incubated for 1 h in the dark at room temperature, ethanol-precipitated and finally resuspended in 100 µl of 100 mM NaOAC (pH 5.0). Approx. 400 µl of adipic acid dehydrazide agarose beads 50% slurry (Sigma) previously equilibrated with 100 mM NaOAC (pH 5.2) were added to each periodate-treated RNA and the mix was incubated for 12 h at 4°C on a rotator. The beads with the bound RNA were then washed two times with 1 ml of 2 M NaCl, and equilibrated in 1X washing buffer (5.2 mM HEPES at pH 7.5, 1 mM $MgCl_2$, 0.8 mM MgAcetate). Then the beads were incubated, in a final volume of 500 µl, with 0.5 mg of HeLa cell nuclear extract (C4, Biotech), 1X binding buffer (5.2 mM HEPES, pH 7.9, 1 mM $MgCl_2$, 0.8 mM MgAcetate, 0.52 mM DTT, 3.8% glycerol, 0.75 mM ATP, 1mM GTP) and Heparin (final concentration 1 µg/µl), for 30 min on a rotator, at room temperature. The beads were then washed four times

with 1.5 ml of washing buffer before addition of SDS sample buffer and loading on SDS-10% polyacrylamide gels. Proteins were visualized by Coomassie brilliant blue staining as described above. Protein sequence analysis of the bands excised from the gel was performed using an electrospray ionization mass spectrometer (LCQ DECA XP-ThermoFinnigam). Protein bands were digested with trypsin and the resulting peptides were extracted with water and 60% acetonitrile/1% trifluoroacetic acid. Fragments were then analyzed by mass spectrometry and proteins were identified by analysis of the peptide MS/MS data with Turbo SEQUEST (ThermoFinnigam) and MASCOT (Matrix Science).

In vitro transcription for pull down analysis.

Plasmids for *in vitro* transcription were generated by annealing the sense and antisense oligos purchased from Sigma, for both WT and T6 mutant exon 18 sequences.

- **$(UG)_6$-$(U)_5$ - Exon 18 WT/U6 Dir**

5'-CTGTGTGTGTGTTTTTTGCAGATGCTKAGTTTGTGTG -3'

- **$(UG)_6$-$(U)_5$ - Exon 18 WT/U6 Rev**

5'-GATCCACACAAACTMAGCATCTGCAAAAAACACACACACACAGAGCT-3'

This was followed by direct cloning into SacI-BamHI digested pBluescript KS vector, under the T7 RNA promoter.

Part of the fibronectin EDA sequence was also cloned following the same procedure and used for pull-down analyses, acting as internal control for ASF/SF2 binding.

- **$(UG)_6$-$(U)_5$ - EDA**

5'-CTGTGTGTGTGTTTTTTGCACCTGATGGTGAAGAAGACACTGCAGAGC-3'

Plasmids were first linearized by HindIII digestion and then *in vitro* transcribed with T7 RNA Polymerase (Promega) according to standard procedures. Briefly 2-5 μg of DNA were transcribed using 40 U of T7 RNA polymerase, in the presence of 1X transcription buffer, and 10 mM DTT, both supplied from Promega, 50 U of RNAsi inhibitor and 5 mM each of the four NTPs to a final volume of 100 μl. Following incubation for 2 h at 37°C, the RNA was purified using NICK columns (Amersham Bioscences), precipitated and resuspended in RNAse-free water. Approx 12-15 μg of transcribed and purified RNA were placed in a 400 μl reaction mixture of 100 mM NaOAc pH 5.0 and 5 mM sodium m-periodate (Sigma) and processed as described above for pull-down assays. After the final centrifugation, 60 μl of SDS–PAGE sample buffer were added to the samples, followed by heating for 5 min at 90°C before loading on 12% SDS–PAGE. The gel was electroblotted on a PVDF membrane according to standard protocols (Amersham Biosciences) and the membrane was blocked with 1X PBS-5% skimmed milk. Proteins were identified using different antibodies and Western blots signals were detected with a chemiluminescence kit (ECL, Pierce Biotechnology). Rabbit anti-hnRNP A1 serum was generously provided by R. Klima and polyclonal antibodies against hnRNP A2 and TDP43 by E. Buratti (ICGEB, Trieste, Italy). Purified GST-DAZAP1 protein was used to immunize a rabbit (New Zeland strain) according to standard protocols, to obtain polyclonal antibodies anti DAZAP1. Anti-ASF/SF2 monoclonal antibody (mAb 96) was purchased from Zymed Laboratories Inc. and anti-tubulin mAb was kindly provided by F. Porro (ICGEB, Trieste, Italy).

Electromobility Shift Assay (EMSA).

Three synthetic RNA oligos exon 18 WT and U6, described previously, and the double mutant G4U6 were used as targets for EMSA experiments.

- **Exon 18 G4U6**

5'-UGCAGAUGGUUAGUUUGUGU-3'

RNA synthetic oligonucleotides (200 ng) were labelled by phosphorylation with [γ–32P] ATP and T4 polynucleotide kinase for 1 h at 37°C, precipitated and resuspended in 200 µl of water. Each binding reaction was made mixing the purified protein of interest with the labeled RNA oligo in 1x binding buffer (5.2 mM HEPES, pH 7.9, 1 mM MgCl2, 0.8 mM MgAcetate, 0.52 mM DTT, 3.8% glycerol, 0.75 mM ATP, 1mM GTP) to a 20 µl final volume. The reaction was left at room temperature for 15–20 minutes before loading the sample on a native polyacrylamide gel (5%) in 0.5X TBE buffers which was run at 100–120 V at 4 °C for 3-4 h. The protein–nucleic acid complexes were visualized using biomax MS films (Kodak). The cDNA of each target protein was amplified and cloned in pGEX-3X plasmid (Pharmacia) and then expressed in BL21 (DE3) bacteria under the induction of 0.5 mM IPTG. The resulting recombinant proteins were purified with glutathione S-Sepharose 4B beads (Pharmacia) according to the manufacturer's instructions using imidazole buffers for protein elution.

Small interfering RNA (siRNA) transfection.

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

human hnRNP A1, 5'-CAGCUGAGGAAGCUCUUCA-3', (Kashima and Manley 2003);

human hnRNP A2, 5'-GGAACAGUUCCGUAAGCUC-3' (Kashima and Manley 2003);

human DAZ-AP1, 5'-GAGACUCUGCGCAGCUACU-3';

human ASF/SF2, 5'-ACGAUUGCCGCAUCUACGU-3' (Cartegni, Hastings et al. 2006);

luciferase #2 gene control, 5'-GCCAUUCUAUCCUCUAGAGGAUG-3'.

HeLa were plated at 2.5 x 10^5 cells per well in 60 mm plates to achieve 40–50% confluency. The next day, 6 µl Oligofectamine was combined with 24 µl of Opti-MEM medium (Invitrogen) and 5-10 µl of 40µM siRNA duplex oligos were diluted in a final volume of 400 µl of Opti-MEM medium. The two mixtures were combined and left for 20 min at RT. Finally, this mix was added to the cells, which were maintained in 1.6 ml of Opti-MEM. After 24h a second round of siRNA transfection was performed as described above. Six to eight hours later Opti-MEM was exchanged with DMEM medium without antibiotic and the cells were transfected with the minigene of interest using Qiagen Effectene transfection reagents. One µg of DNA was mixed with 8 µl of Enhancer for each transfection and the mixture was incubated at room temperature for 5 minutes to allow the condensation of the DNA. Then, 10 µl of Effectene were added to the mixture and an incubation of 10 minutes has been performed. After the addition of 500 µl of complete culture DMEM medium, the mixture was added to the cells in 4 ml of the same medium and incubated at 37°C. After 12 h, HeLa cells were harvested and divided in two parts for protein and RNA extractions. RT-PCR from total RNA was performed as described above in the transfection protocol. Whole protein extracts were obtained by cell sonication in lysis buffer (15 mM HEPES, pH 7.5, 250 mM NaCl, 0.5% NP40, 10% glycerol and 1 mM PMSF) and analyzed for hnRNP A1, A2, DAZAP1 and ASF/SF2 endogenous protein expression by immunoblot analysis using the antibodies described above. Tubulin was used as protein loading control. Each siRNA treatment experiment was repeated at least three times.

6 REFERENCES

Adams, M. D., S. E. Celniker, et al. (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185-95.

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): 237-46.

Akindahunsi, A. A., A. Bandiera, et al. (2005). "Vertebrate 2xRBD hnRNP proteins: a comparative analysis of genome, mRNA and protein sequences." Comput Biol Chem **29**(1): 13-23.

Amendt, B. A., D. Hesslein, et al. (1994). "Presence of negative and positive cis-acting RNA splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1." Mol Cell Biol **14**(6): 3960-70.

Amendt, B. A., Z. H. Si, et al. (1995). "Presence of exon splicing silencers within human immunodeficiency virus type 1 tat exon 2 and tat-rev exon 3: evidence for inhibition mediated by cellular factors." Mol Cell Biol **15**(8): 4606-15.

Anderson, S. F., B. P. Schlegel, et al. (1998). "BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A." Nat Genet **19**(3): 254-6.

Ars, E., E. Serra, et al. (2000). "Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1." Hum Mol Genet **9**(2): 237-47.

Bai, Y., D. Lee, et al. (1999). "Control of 3' splice site choice in vivo by ASF/SF2 and hnRNP A1." Nucleic Acids Res **27**(4): 1126-34.

Baralle, M., N. Skoko, et al. (2006). "NF1 mRNA biogenesis: effect of the genomic milieu in splicing regulation of the NF1 exon 37 region." FEBS Lett **580**(18): 4449-56.

Barnard, D. C., J. Li, et al. (2002). "Regulation of alternative splicing by SRrp86 through coactivation and repression of specific SR proteins." Rna **8**(4): 526-33.

Batsche, E., M. Yaniv, et al. (2006). "The human SWI/SNF subunit Brm is a regulator of alternative splicing." Nat Struct Mol Biol **13**(1): 22-9.

Ben-Dov, C., B. Hartmann, et al. (2008). "Genome-wide analysis of alternative pre-mRNA splicing." J Biol Chem **283**(3): 1229-33.

Bentley, D. (2002). "The mRNA assembly line: transcription and processing machines in the same factory." Curr Opin Cell Biol **14**(3): 336-42.

Bentley, D. L. (2005). "Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors." Curr Opin Cell Biol **17**(3): 251-6.

Berget, S. M. (1995). "Exon recognition in vertebrate splicing." J Biol Chem **270**(6): 2411-4.

Berget, S. M., C. Moore, et al. (1977). "Spliced segments at the 5' terminus of adenovirus 2 late mRNA." Proc Natl Acad Sci U S A **74**(8): 3171-5.

Berglund, J. A., K. Chua, et al. (1997). "The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC." Cell **89**(5): 781-7.

Birney, E., S. Kumar, et al. (1993). "Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors." Nucleic Acids Res **21**(25): 5803-16.

Black, D. L. (1991). "Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells?" Genes Dev **5**(3): 389-402.

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

Blanchette, M. and B. Chabot (1999). "Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization." Embo J **18**(7): 1939-52.

Bochar, D. A., L. Wang, et al. (2000). "BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer." Cell **102**(2): 257-65.

Boehringer, D., E. M. Makarov, et al. (2004). "Three-dimensional structure of a pre-catalytic human spliceosomal complex B." Nat Struct Mol Biol **11**(5): 463-8.

Bonnal, S., F. Pileur, et al. (2005). "Heterogeneous nuclear ribonucleoprotein A1 is a novel internal ribosome entry site trans-acting factor that modulates alternative initiation of translation of the fibroblast growth factor 2 mRNA." J Biol Chem **280**(6): 4144-53.

Bork, P., K. Hofmann, et al. (1997). "A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins." Faseb J **11**(1): 68-76.

Boucher, L., C. A. Ouzounis, et al. (2001). "A genome-wide survey of RS domain proteins." Rna **7**(12): 1693-701.

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

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

Breathnach, R., C. Benoist, et al. (1978). "Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries." Proc Natl Acad Sci U S A **75**(10): 4853-7.

Brosi, R., H. P. Hauri, et al. (1993). "Separation of splicing factor SF3 into two components and purification of SF3a activity." J Biol Chem **268**(23): 17640-6.

Buratti, E. and F. E. Baralle (2001). "Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9." J Biol Chem **276**(39): 36337-43.

Buratti, E. and F. E. Baralle (2004). "Influence of RNA secondary structure on the pre-mRNA splicing process." Mol Cell Biol **24**(24): 10505-14.

Buratti, E., M. Baralle, et al. (2006). "Defective splicing, disease and therapy: searching for master checkpoints in exon definition." Nucleic Acids Res **34**(12): 3494-510.

Buratti, E., M. Baralle, et al. (2004). "hnRNP H binding at the 5' splice site correlates with the pathological effect of two intronic mutations in the NF-1 and TSHbeta genes." Nucleic Acids Res **32**(14): 4224-36.

Buratti, E., T. Dork, et al. (2001). "Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping." Embo J **20**(7): 1774-84.

Buratti, E., C. Stuani, et al. (2007). "SR protein-mediated inhibition of CFTR exon 9 inclusion: molecular characterization of the intronic splicing silencer." Nucleic Acids Res **35**(13): 4359-68.

Burd, C. G. and G. Dreyfuss (1994). "RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing." Embo J **13**(5): 1197-204.

Burge, C. B., R. A. Padgett, et al. (1998). "Evolutionary fates and origins of U12-type introns." Mol Cell **2**(6): 773-85.

Caceres, J. F. and A. R. Kornblihtt (2002). "Alternative splicing: multiple control mechanisms and involvement in human disease." Trends Genet **18**(4): 186-93.

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

Caceres, J. F., S. Stamm, et al. (1994). "Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors." Science **265**(5179): 1706-9.

Callebaut, I. and J. P. Mornon (1997). "From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair." FEBS Lett **400**(1): 25-30.

Caputi, M., G. Casari, et al. (1994). "A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon." Nucleic Acids Res **22**(6): 1018-22.

Caputi, M. and A. M. Zahler (2001). "Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family." J Biol Chem **276**(47): 43850-9.

Caputi, M. and A. M. Zahler (2002). "SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D." Embo J **21**(4): 845-55.

Cartegni, L., S. L. Chew, et al. (2002). "Listening to silence and understanding nonsense: exonic mutations that affect splicing." Nat Rev Genet **3**(4): 285-98.

Cartegni, L., M. L. Hastings, et al. (2006). "Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2." Am J Hum Genet **78**(1): 63-77.

Cartegni, L. and A. R. Krainer (2002). "Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1." Nat Genet **30**(4): 377-84.

Cartegni, L., J. Wang, et al. (2003). "ESEfinder: A web resource to identify exonic splicing enhancers." Nucleic Acids Res **31**(13): 3568-71.

Chabot, B., M. Blanchette, et al. (1997). "An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1." Mol Cell Biol **17**(4): 1776-86.

Chabot, B., C. LeBel, et al. (2003). "Heterogeneous nuclear ribonucleoprotein particle A/B proteins and the control of alternative splicing of the mammalian heterogeneous nuclear ribonucleoprotein particle A1 pre-mRNA." Prog Mol Subcell Biol **31**: 59-88.

Chen, C. D., R. Kobayashi, et al. (1999). "Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene." Genes Dev **13**(5): 593-606.

Chen, C. F., S. Li, et al. (1996). "The nuclear localization sequences of the BRCA1 protein interact with the importin-alpha subunit of the nuclear transport signal receptor." J Biol Chem **271**(51): 32863-8.

Chen, I. T. and L. A. Chasin (1994). "Large exon size does not limit splicing in vivo." Mol Cell Biol **14**(3): 2140-6.

Chou, M. Y., N. Rooke, et al. (1999). "hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells." Mol Cell Biol **19**(1): 69-77.

Chow, L. T., R. E. Gelinas, et al. (1977). "An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA." Cell **12**(1): 1-8.

Chung, C. T., S. L. Niemela, et al. (1989). "One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution." Proc Natl Acad Sci U S A **86**(7): 2172-5.

Collins, C. A. and C. Guthrie (1999). "Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome." Genes Dev **13**(15): 1970-82.

Coolidge, C. J., R. J. Seely, et al. (1997). "Functional analysis of the polypyrimidine tract in pre-mRNA splicing." Nucleic Acids Res **25**(4): 888-96.

Cooper, T. A. (2005). "Use of minigene systems to dissect alternative splicing elements." Methods **37**(4): 331-40.

Coulter, L. R., M. A. Landree, et al. (1997). "Identification of a new class of exonic splicing enhancers by in vivo selection." Mol Cell Biol **17**(4): 2143-50.

Cramer, P., J. F. Caceres, et al. (1999). "Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer." Mol Cell **4**(2): 251-8.

Cramer, P., C. G. Pesce, et al. (1997). "Functional association between promoter structure and transcript alternative splicing." Proc Natl Acad Sci U S A **94**(21): 11456-60.

Dai, T., Y. Vera, et al. (2001). "Characterization of the mouse Dazap1 gene encoding an RNA-binding protein that interacts with infertility factors DAZ and DAZL." BMC Genomics **2**(1): 6.

Dallaire, F., S. Dupuis, et al. (2000). "Heterogeneous nuclear ribonucleoprotein A1 and UP1 protect mammalian telomeric repeats and modulate telomere replication in vitro." J Biol Chem **275**(19): 14509-16.

Das, R., Z. Zhou, et al. (2000). "Functional association of U2 snRNP with the ATP-independent spliceosomal complex E." Mol Cell **5**(5): 779-87.

David, C. J. and J. L. Manley (2008). "The search for alternative splicing regulators: new approaches offer a path to a splicing code." Genes Dev **22**(3): 279-85.

de la Mata, M., C. R. Alonso, et al. (2003). "A slow RNA polymerase II affects alternative splicing in vivo." Mol Cell **12**(2): 525-32.

de la Mata, M. and A. R. Kornblihtt (2006). "RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20." Nat Struct Mol Biol **13**(11): 973-80.

Del Gatto-Konczak, F., M. Olive, et al. (1999). "hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer." Mol Cell Biol **19**(1): 251-60.

Del Gatto, F., M. C. Gesnel, et al. (1996). "The exon sequence TAGG can inhibit splicing." Nucleic Acids Res **24**(11): 2017-21.

Delaney, S. J., D. P. Rich, et al. (1993). "Cystic fibrosis transmembrane conductance regulator splice variants are not conserved and fail to produce chloride channels." Nat Genet **4**(4): 426-31.

Deng, C. X. and S. G. Brodie (2000). "Roles of BRCA1 and its interacting proteins." Bioessays **22**(8): 728-37.

Denis, M. M., N. D. Tolley, et al. (2005). "Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets." Cell **122**(3): 379-91.

Disset, A., C. F. Bourgeois, et al. (2006). "An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements." Hum Mol Genet **15**(6): 999-1013.

Dominguez, C. and F. H. Allain (2006). "NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition." Nucleic Acids Res **34**(13): 3634-45.

Dominski, Z. and R. Kole (1991). "Selection of splice sites in pre-mRNAs with short internal exons." Mol Cell Biol **11**(12): 6075-83.

Donahue, C. P., C. Muratore, et al. (2006). "Stabilization of the tau exon 10 stem loop alters pre-mRNA splicing." J Biol Chem **281**(33): 23302-6.

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

Dreyfuss, G., M. J. Matunis, et al. (1993). "hnRNP proteins and the biogenesis of mRNA." Annu Rev Biochem **62**: 289-321.

Fairbrother, W. G. and L. A. Chasin (2000). "Human genomic sequences that inhibit splicing." Mol Cell Biol **20**(18): 6816-25.

Fairbrother, W. G., R. F. Yeh, et al. (2002). "Predictive identification of exonic splicing enhancers in human genes." Science **297**(5583): 1007-13.

Fairbrother, W. G., G. W. Yeo, et al. (2004). "RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons." Nucleic Acids Res **32**(Web Server issue): W187-90.

Fan, S., Y. X. Ma, et al. (2001). "Role of direct interaction in BRCA1 inhibition of estrogen receptor activity." Oncogene **20**(1): 77-87.

Faustino, N. A. and T. A. Cooper (2003). "Pre-mRNA splicing and human disease." Genes Dev **17**(4): 419-37.

Fischer, U., V. Sumpter, et al. (1993). "Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap." Embo J **12**(2): 573-83.

Fiset, S. and B. Chabot (2001). "hnRNP A1 may interact simultaneously with telomeric DNA and the human telomerase RNA in vitro." Nucleic Acids Res **29**(11): 2268-75.

Fong, Y. W. and Q. Zhou (2001). "Stimulatory effect of splicing factors on transcriptional elongation." Nature **414**(6866): 929-33.

Furger, A., J. M. O'Sullivan, et al. (2002). "Promoter proximal splice sites enhance transcription." Genes Dev **16**(21): 2792-9.

Futreal, P. A., Q. Liu, et al. (1994). "BRCA1 mutations in primary breast and ovarian carcinomas." Science **266**(5182): 120-2.

Gallego, M. E., R. Gattoni, et al. (1997). "The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the beta-tropomyosin alternative exon 6A." Embo J **16**(7): 1772-84.

Garcia-Blanco, M. A., A. P. Baraniak, et al. (2004). "Alternative splicing in disease and therapy." Nat Biotechnol **22**(5): 535-46.

Ge, H. and J. L. Manley (1990). "A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro." Cell **62**(1): 25-34.

Ghigna, C., S. Giordano, et al. (2005). "Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene." Mol Cell **20**(6): 881-90.

Glanzer, J., K. Y. Miyashiro, et al. (2005). "RNA splicing capability of live neuronal dendrites." Proc Natl Acad Sci U S A **102**(46): 16859-64.

Goguel, V. and M. Rosbash (1993). "Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast." Cell **72**(6): 893-901.

Golling, G., A. Amsterdam, et al. (2002). "Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development." Nat Genet **31**(2): 135-40.

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): 233-43.

Graveley, B. R. (2000). "Sorting out the complexity of SR protein functions." Rna **6**(9): 1197-211.

Graveley, B. R. (2001). "Alternative splicing: increasing diversity in the proteomic world." Trends Genet **17**(2): 100-7.

Graveley, B. R., K. J. Hertel, et al. (2001). "The role of U2AF35 and U2AF65 in enhancer-dependent splicing." Rna **7**(6): 806-18.

Grover, A., H. Houlden, et al. (1999). "5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10." J Biol Chem **274**(21): 15134-43.

Guil, S. and J. F. Caceres (2007). "The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a." Nat Struct Mol Biol **14**(7): 591-6.

Guthrie, C. (1991). "Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein." Science **253**(5016): 157-63.

Haile, D. T. and J. D. Parvin (1999). "Activation of transcription in vitro by the BRCA1 carboxyl-terminal domain." J Biol Chem **274**(4): 2113-7.

Hall, S. L. and R. A. Padgett (1994). "Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites." J Mol Biol **239**(3): 357-65.

Hall, S. L. and R. A. Padgett (1996). "Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns." Science **271**(5256): 1716-8.

Hamilton, B. J., C. M. Burns, et al. (1997). "Modulation of AUUUA response element binding by heterogeneous nuclear ribonucleoprotein A1 in human T lymphocytes. The roles of cytoplasmic location, transcription, and phosphorylation." J Biol Chem **272**(45): 28732-41.

Han, K., G. Yeo, et al. (2005). "A combinatorial code for splicing silencing: UAGG and GGGG motifs." PLoS Biol **3**(5): e158.

Hanamura, A., J. F. Caceres, et al. (1998). "Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors." Rna **4**(4): 430-44.

Hashizume, R., M. Fukuda, et al. (2001). "The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation." J Biol Chem **276**(18): 14537-40.

Hawkins, J. D. (1988). "A survey on intron and exon lengths." Nucleic Acids Res **16**(21): 9893-908.

Hefferon, T. W., J. D. Groman, et al. (2004). "A variable dinucleotide repeat in the CFTR gene contributes to phenotype diversity by forming RNA secondary structures that alter splicing." Proc Natl Acad Sci U S A **101**(10): 3504-9.

Hiller, M., Z. Zhang, et al. (2007). "Pre-mRNA secondary structures influence exon recognition." PLoS Genet **3**(11): e204.

Hoffmeyer, S., P. Nurnberg, et al. (1998). "Nearby stop codons in exons of the neurofibromatosis type 1 gene are disparate splice effectors." Am J Hum Genet **62**(2): 269-77.

Hong, W., M. Bennett, et al. (1997). "Association of U2 snRNP with the spliceosomal complex E." Nucleic Acids Res **25**(2): 354-61.

Honore, B., H. H. Rasmussen, et al. (1995). "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromosomes." J Biol Chem **270**(48): 28780-9.

Hori, T., Y. Taguchi, et al. (2005). "The RNA ligands for mouse proline-rich RNA-binding protein (mouse Prrp) contain two consensus sequences in separate loop structure." Nucleic Acids Res **33**(1): 190-200.

Horowitz, D. S. and A. R. Krainer (1994). "Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing." Trends Genet **10**(3): 100-6.

House, A. E. and K. W. Lynch (2006). "An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition." Nat Struct Mol Biol **13**(10): 937-44.

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

Huh, G. S. and R. O. Hynes (1993). "Elements regulating an alternatively spliced exon of the rat fibronectin gene." Mol Cell Biol **13**(9): 5301-14.

Hui, J., L. H. Hung, et al. (2005). "Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing." Embo J **24**(11): 1988-98.

Hutchison, S., C. LeBel, et al. (2002). "Distinct sets of adjacent heterogeneous nuclear ribonucleoprotein (hnRNP) A1/A2 binding sites control 5' splice site selection in the hnRNP A1 mRNA precursor." J Biol Chem **277**(33): 29745-52.

Ibrahim el, C., T. D. Schaal, et al. (2005). "Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers." Proc Natl Acad Sci U S A **102**(14): 5002-7.

Iervolino, A., G. Santilli, et al. (2002). "hnRNP A1 nucleocytoplasmic shuttling activity is required for normal myelopoiesis and BCR/ABL leukemogenesis." Mol Cell Biol **22**(7): 2255-66.

Ikeda, Y., A. Takagi, et al. (2001). "Novel compound heterozygous mutations for lipoprotein lipase deficiency. A G-to-T transversion at the first position of exon 5 causing G154V missense mutation and a 5' splice site mutation of intron 8." J Lipid Res **42**(7): 1072-81.

Izaurralde, E., A. Jarmolowski, et al. (1997). "A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export." J Cell Biol **137**(1): 27-35.

Izaurralde, E., J. Lewis, et al. (1994). "A nuclear cap binding protein complex involved in pre-mRNA splicing." Cell **78**(4): 657-68.

Izaurralde, E. and I. W. Mattaj (1995). "RNA export." Cell **81**(2): 153-9.

Jensen, K. B., B. K. Dredge, et al. (2000). "Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability." Neuron **25**(2): 359-71.

Jiang, Z., J. Cote, et al. (2000). "Aberrant splicing of tau pre-mRNA caused by intronic mutations associated with the inherited dementia frontotemporal dementia with parkinsonism linked to chromosome 17." Mol Cell Biol **20**(11): 4036-48.

Johnson, J. M., J. Castle, et al. (2003). "Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays." Science **302**(5653): 2141-4.

Jumaa, H., G. Wei, et al. (1999). "Blastocyst formation is blocked in mouse embryos lacking the splicing factor SRp20." Curr Biol **9**(16): 899-902.

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

Kadener, S., J. P. Fededa, et al. (2002). "Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation." Proc Natl Acad Sci U S A **99**(12): 8185-90.

Kambach, C., S. Walke, et al. (1999). "Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles." Curr Opin Struct Biol **9**(2): 222-30.

Kamma, H., D. S. Portman, et al. (1995). "Cell type-specific expression of hnRNP proteins." Exp Cell Res **221**(1): 187-96.

Karni, R., E. de Stanchina, et al. (2007). "The gene encoding the splicing factor SF2/ASF is a proto-oncogene." Nat Struct Mol Biol **14**(3): 185-93.

Kashima, T. and J. L. Manley (2003). "A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy." Nat Genet **34**(4): 460-3.

Kashima, T., N. Rao, et al. (2007). "hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing." Hum Mol Genet **16**(24): 3149-59.

Kashima, T., N. Rao, et al. (2007). "An intronic element contributes to splicing repression in spinal muscular atrophy." Proc Natl Acad Sci U S A **104**(9): 3426-31.

Kaufmann, D., W. Leistner, et al. (2002). "Aberrant splicing in several human tumors in the tumor suppressor genes neurofibromatosis type 1, neurofibromatosis type 2, and tuberous sclerosis 2." Cancer Res **62**(5): 1503-9.

Kellenberger, E., G. Stier, et al. (2002). "Induced folding of the U2AF35 RRM upon binding to U2AF65." FEBS Lett **528**(1-3): 171-6.

Kielkopf, C. L., S. Lucke, et al. (2004). "U2AF homology motifs: protein recognition in the RRM world." Genes Dev **18**(13): 1513-26.

Kielkopf, C. L., N. A. Rodionova, et al. (2001). "A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer." Cell **106**(5): 595-605.

Kim, S., B. M. Merrill, et al. (1997). "Identification of N(G)-methylarginine residues in human heterogeneous RNP protein A1: Phe/Gly-Gly-Gly-Arg-Gly-Gly-Gly/Phe is a preferred recognition motif." Biochemistry **36**(17): 5185-92.

Kohtz, J. D., S. F. Jamison, et al. (1994). "Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors." Nature **368**(6467): 119-24.

Konarska, M. M. and P. A. Sharp (1987). "Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes." Cell **49**(6): 763-74.

Kornblihtt, A. R., M. de la Mata, et al. (2004). "Multiple links between transcription and splicing." Rna **10**(10): 1489-98.

Krainer, A. R., G. C. Conway, et al. (1990). "The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites." Cell **62**(1): 35-42.

Kramer, A. (1996). "The structure and function of proteins involved in mammalian pre-mRNA splicing." Annu Rev Biochem **65**: 367-409.

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.

Krum, S. A., G. A. Miranda, et al. (2003). "BRCA1 associates with processive RNA polymerase II." J Biol Chem **278**(52): 52012-20.

LaBranche, H., S. Dupuis, et al. (1998). "Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase." Nat Genet **19**(2): 199-202.

Ladd, A. N., N. Charlet, et al. (2001). "The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing." Mol Cell Biol **21**(4): 1285-96.

Ladd, A. N. and T. A. Cooper (2002). "Finding signals that regulate alternative splicing in the post-genomic era." Genome Biol **3**(11): reviews0008.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-5.

Lamond, A. I. (1993). "The spliceosome." Bioessays **15**(9): 595-603.

Lamond, A. I. and D. L. Spector (2003). "Nuclear speckles: a model for nuclear organelles." Nat Rev Mol Cell Biol **4**(8): 605-12.

Langford, C. J., F. J. Klinz, et al. (1984). "Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast." Cell **36**(3): 645-53.

Lavigueur, A., H. La Branche, et al. (1993). "A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding." Genes Dev **7**(12A): 2405-17.

Lefebvre, S., L. Burglen, et al. (1995). "Identification and characterization of a spinal muscular atrophy-determining gene." Cell **80**(1): 155-65.

Lemaire, R., J. Prasad, et al. (2002). "Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins." Genes Dev **16**(5): 594-607.

Lewandowska, M. A., C. Stuani, et al. (2005). "Functional studies on the ATM intronic splicing processing element." Nucleic Acids Res **33**(13): 4007-15.

Li, S., P. L. Chen, et al. (1999). "Binding of CtIP to the BRCT repeats of BRCA1 involved in the transcription regulation of p21 is disrupted upon DNA damage." J Biol Chem **274**(16): 11334-8.

Li, X. and J. L. Manley (2005). "Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability." Cell **122**(3): 365-78.

Li, X. and J. L. Manley (2006). "Cotranscriptional processes and their influence on genome stability." Genes Dev **20**(14): 1838-47.

Lin, S., G. Coutinho-Mansfield, et al. (2008). "The splicing factor SC35 has an active role in transcriptional elongation." Nat Struct Mol Biol **15**(8): 819-26.

Lin, Y. T. and P. H. Yen (2006). "A novel nucleocytoplasmic shuttling sequence of DAZAP1, a testis-abundant RNA-binding protein." Rna **12**(8): 1486-93.

Liu, H. X., L. Cartegni, et al. (2001). "A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes." Nat Genet **27**(1): 55-8.

Liu, H. X., S. L. Chew, et al. (2000). "Exonic splicing enhancer motif recognized by human SC35 under splicing conditions." Mol Cell Biol **20**(3): 1063-71.

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

Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. E. (2000). Molecular cell biology (4th edition).

Longman, D., I. L. Johnstone, et al. (2000). "Functional characterization of SR and SR-related genes in Caenorhabditis elegans." Embo J **19**(7): 1625-37.

Lopez, A. J. (1998). "Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation." Annu Rev Genet **32**: 279-305.

Lorson, C. L., E. Hahnen, et al. (1999). "A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy." Proc Natl Acad Sci U S A **96**(11): 6307-11.

Mabon, S. A. and T. Misteli (2005). "Differential recruitment of pre-mRNA splicing factors to alternatively spliced transcripts in vivo." PLoS Biol **3**(11): e374.

MacLachlan, T. K., K. Somasundaram, et al. (2000). "BRCA1 effects on the cell cycle and the DNA damage response are linked to altered gene expression." J Biol Chem **275**(4): 2777-85.

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

Maniatis, T. and B. Tasic (2002). "Alternative pre-mRNA splicing and proteome expansion in metazoans." Nature **418**(6894): 236-43.

Manke, I. A., D. M. Lowery, et al. (2003). "BRCT repeats as phosphopeptide-binding modules involved in protein targeting." Science **302**(5645): 636-9.

Mardon, H. J., G. Sebastio, et al. (1987). "A role for exon sequences in alternative splicing of the human fibronectin gene." Nucleic Acids Res **15**(19): 7725-33.

Markovtsov, V., J. M. Nikolic, et al. (2000). "Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein." Mol Cell Biol **20**(20): 7463-79.

Maroney, P. A., C. M. Romfo, et al. (2000). "Functional recognition of 5' splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly." Mol Cell **6**(2): 317-28.

Martinez-Contreras, R., P. Cloutier, et al. (2007). "hnRNP proteins and splicing control." Adv Exp Med Biol **623**: 123-47.

Martinez-Contreras, R., J. F. Fisette, et al. (2006). "Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing." PLoS Biol **4**(2): e21.

Matlin, A. J., F. Clark, et al. (2005). "Understanding alternative splicing: towards a cellular code." Nat Rev Mol Cell Biol **6**(5): 386-98.

Matsuo, M., H. Nishio, et al. (1992). "Partial deletion of a dystrophin gene leads to exon skipping and to loss of an intra-exon hairpin structure from the predicted mRNA precursor." Biochem Biophys Res Commun **182**(2): 495-500.

Mayeda, A. and A. R. Krainer (1992). "Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2." Cell **68**(2): 365-75.

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): 4321-31.

Mazoyer, S., N. Puget, et al. (1998). "A BRCA1 nonsense mutation causes exon skipping." Am J Hum Genet **62**(3): 713-5.

McCullough, A. J. and S. M. Berget (1997). "G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection." Mol Cell Biol **17**(8): 4562-71.

Merendino, L., S. Guth, et al. (1999). "Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG." Nature **402**(6763): 838-41.

Messiaen, L., T. Callens, et al. (1997). "Characterisation of two different nonsense mutations, C6792A and C6792G, causing skipping of exon 37 in the NF1 gene." Hum Genet **101**(1): 75-80.

Miki, Y., J. Swensen, et al. (1994). "A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1." Science **266**(5182): 66-71.

Misteli, T. and D. L. Spector (1997). "Protein phosphorylation and the nuclear organization of pre-mRNA splicing." Trends Cell Biol **7**(4): 135-8.

Miyajima, H., H. Miyaso, et al. (2002). "Identification of a cis-acting element for the regulation of SMN exon 7 splicing." J Biol Chem **277**(26): 23271-7.

Miyaso, H., M. Okumura, et al. (2003). "An intronic splicing enhancer element in survival motor neuron (SMN) pre-mRNA." J Biol Chem **278**(18): 15825-31.

Moisan, A., C. Larochelle, et al. (2004). "BRCA1 can modulate RNA polymerase II carboxy-terminal domain phosphorylation levels." Mol Cell Biol **24**(16): 6947-56.

Monani, U. R., C. L. Lorson, et al. (1999). "A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2." Hum Mol Genet **8**(7): 1177-83.

Moore, M. J. and P. A. Sharp (1993). "Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing." Nature **365**(6444): 364-8.

Morton, S., H. T. Yang, et al. (2006). "Phosphorylation of the ARE-binding protein DAZAP1 by ERK2 induces its dissociation from DAZ." Biochem J **399**(2): 265-73.

Mullan, P. B., J. E. Quinn, et al. (2006). "The role of BRCA1 in transcriptional regulation and cell cycle control." Oncogene **25**(43): 5854-63.

Muro, A. F., M. Caputi, et al. (1999). "Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display." Mol Cell Biol **19**(4): 2657-71.

Muro, A. F., A. Iaconcig, et al. (1998). "Regulation of the fibronectin EDA exon alternative splicing. Cooperative role of the exonic enhancer element and the 5' splicing site." FEBS Lett **437**(1-2): 137-41.

Narita, T., Y. Yamaguchi, et al. (2003). "Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex." Mol Cell Biol **23**(6): 1863-73.

Narod, S. A. and W. D. Foulkes (2004). "BRCA1 and BRCA2: 1994 and beyond." Nat Rev Cancer **4**(9): 665-76.

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): 4113-25.

Neuhausen, S. L. and C. J. Marshall (1994). "Loss of heterozygosity in familial tumors from three BRCA1-linked kindreds." Cancer Res **54**(23): 6069-72.

Niwa, M. and S. M. Berget (1991). "Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns." Genes Dev **5**(11): 2086-95.

Nogues, G., S. Kadener, et al. (2002). "Transcriptional activators differ in their abilities to control alternative splicing." J Biol Chem **277**(45): 43110-4.

O'Keefe, R. T., C. Norman, et al. (1996). "The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast." Cell **86**(4): 679-89.

Ouchi, T., S. W. Lee, et al. (2000). "Collaboration of signal transducer and activator of transcription 1 (STAT1) and BRCA1 in differential regulation of IFN-gamma target genes." Proc Natl Acad Sci U S A **97**(10): 5208-13.

Padgett, R. A., P. J. Grabowski, et al. (1986). "Splicing of messenger RNA precursors." Annu Rev Biochem **55**: 1119-50.

Pagani, F. and F. E. Baralle (2004). "Genomic variants in exons and introns: identifying the splicing spoilers." Nat Rev Genet **5**(5): 389-96.

Pagani, F., E. Buratti, et al. (2003). "Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9." J Biol Chem **278**(29): 26580-8.

Pagani, F., E. Buratti, et al. (2002). "A new type of mutation causes a splicing defect in ATM." Nat Genet **30**(4): 426-9.

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): 21041-7.

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): 1111-20.

Pagani, F., C. Stuani, et al. (2003). "Promoter architecture modulates CFTR exon 9 skipping." J Biol Chem **278**(3): 1511-7.

Parvin, J. D. (2001). "BRCA1 at a branch point." Proc Natl Acad Sci U S A **98**(11): 5952-4.

Patel, A. A. and J. A. Steitz (2003). "Splicing double: insights from the second spliceosome." Nat Rev Mol Cell Biol **4**(12): 960-70.

Perrin-Vidoz, L., O. M. Sinilnikova, et al. (2002). "The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons." Hum Mol Genet **11**(23): 2805-14.

Pinol-Roma, S. and G. Dreyfuss (1992). "Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm." Nature **355**(6362): 730-2.

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): 336-9.

Pozzoli, U. and M. Sironi (2005). "Silencers regulate both constitutive and alternative splicing events in mammals." Cell Mol Life Sci **62**(14): 1579-604.

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): 2256-69.

Ram, O. and G. Ast (2007). "SR proteins: a foot on the exon before the transition from intron to exon definition." Trends Genet **23**(1): 5-7.

Rappsilber, J., U. Ryder, et al. (2002). "Large-scale proteomic analysis of the human spliceosome." Genome Res **12**(8): 1231-45.

Reed, R. (1989). "The organization of 3' splice-site sequences in mammalian introns." Genes Dev **3**(12B): 2113-23.

Reed, R. (2000). "Mechanisms of fidelity in pre-mRNA splicing." Curr Opin Cell Biol **12**(3): 340-5.

Reed, R. and T. Maniatis (1988). "The role of the mammalian branchpoint sequence in pre-mRNA splicing." Genes Dev **2**(10): 1268-76.

Ring, H. Z. and J. T. Lis (1994). "The SR protein B52/SRp55 is essential for Drosophila development." Mol Cell Biol **14**(11): 7499-506.

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.

Roberts, G. C., C. Gooding, et al. (1998). "Co-transcriptional commitment to alternative splice site selection." Nucleic Acids Res **26**(24): 5568-72.

Roscigno, R. F. and M. A. Garcia-Blanco (1995). "SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome." Rna **1**(7): 692-706.

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): 11222-9.

Rosen, E. M., S. Fan, et al. (2006). "BRCA1 regulation of transcription." Cancer Lett **236**(2): 175-85.

Rousseau, S., N. Morrice, et al. (2002). "Inhibition of SAPK2a/p38 prevents hnRNP A0 phosphorylation by MAPKAP-K2 and its interaction with cytokine mRNAs." Embo J **21**(23): 6505-14.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Sanford, J. R., J. Ellis, et al. (2005). "Multiple roles of arginine/serine-rich splicing factors in RNA processing." Biochem Soc Trans **33**(Pt 3): 443-6.

Sanford, J. R., N. K. Gray, et al. (2004). "A novel role for shuttling SR proteins in mRNA translation." Genes Dev **18**(7): 755-68.

Sauliere, J., A. Sureau, et al. (2006). "The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit." Mol Cell Biol **26**(23): 8755-69.

Schaal, T. D. and T. Maniatis (1999). "Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA." Mol Cell Biol **19**(1): 261-73.

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

Schmucker, D., J. C. Clemens, et al. (2000). "Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity." Cell **101**(6): 671-84.

Scully, R., S. F. Anderson, et al. (1997). "BRCA1 is a component of the RNA polymerase II holoenzyme." Proc Natl Acad Sci U S A **94**(11): 5605-10.

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): 2533-8.

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): 7155-74.

Sharp, P. A. (1994). "Split genes and RNA splicing." Cell **77**(6): 805-15.

Sharp, P. A. and C. B. Burge (1997). "Classification of introns: U2-type or U12-type." Cell **91**(7): 875-9.

Shelley, C. S., C. R. Sharpe, et al. (1985). "Comparison of the human apolipoprotein genes. Apo AII presents a unique functional intron-exon junction." J Mol Biol **186**(1): 43-51.

Shen, H. and M. R. Green (2004). "A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly." Mol Cell **16**(3): 363-73.

Shepard, J., M. Reick, et al. (2002). "Characterization of U2AF(6), a splicing factor related to U2AF(35)." Mol Cell Biol **22**(1): 221-30.

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): 1258-67.

Singh, N. K., N. N. Singh, et al. (2006). "Splicing of a critical exon of human Survival Motor Neuron is regulated by a unique silencer element located in the last intron." Mol Cell Biol **26**(4): 1333-46.

Singh, N. N., E. J. Androphy, et al. (2004). "An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy." Biochem Biophys Res Commun **315**(2): 381-8.

Singh, N. N., R. N. Singh, et al. (2007). "Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes." Nucleic Acids Res **35**(2): 371-89.

Singh, R. N. (2007). "Evolving concepts on human SMN pre-mRNA splicing." RNA Biol **4**(1): 7-10.

Sironi, M., G. Menozzi, et al. (2004). "Silencer elements as possible inhibitors of pseudoexon splicing." Nucleic Acids Res **32**(5): 1783-91.

Smith, C. W. and B. Nadal-Ginard (1989). "Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing." Cell **56**(5): 749-58.

Smith, C. W. and J. Valcarcel (2000). "Alternative pre-mRNA splicing: the logic of combinatorial control." Trends Biochem Sci **25**(8): 381-8.

Staley, J. P. and C. Guthrie (1998). "Mechanical devices of the spliceosome: motors, clocks, springs, and things." Cell **92**(3): 315-26.

Stamm, S. (2008). "Regulation of alternative splicing by reversible protein phosphorylation." J Biol Chem **283**(3): 1223-7.

Stark, H., P. Dube, et al. (2001). "Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle." Nature **409**(6819): 539-42.

Sterner, D. A. and S. M. Berget (1993). "In vivo recognition of a vertebrate mini-exon as an exon-intron-exon unit." Mol Cell Biol **13**(5): 2677-87.

Sterner, D. A., T. Carlo, et al. (1996). "Architectural limits on split genes." Proc Natl Acad Sci U S A **93**(26): 15081-5.

Sun, H. and L. A. Chasin (2000). "Multiple splicing defects in an intronic false exon." Mol Cell Biol **20**(17): 6414-25.

Sun, Q., A. Mayeda, et al. (1993). "General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer." Genes Dev **7**(12B): 2598-608.

Svitkin, Y. V., L. P. Ovchinnikov, et al. (1996). "General RNA binding proteins render translation cap dependent." Embo J **15**(24): 7147-55.

Tacke, R., Y. Chen, et al. (1997). "Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer." Proc Natl Acad Sci U S A **94**(4): 1148-53.

Talerico, M. and S. M. Berget (1994). "Intron definition in splicing of small Drosophila introns." Mol Cell Biol **14**(5): 3434-45.

Tange, T. O., C. K. Damgaard, et al. (2001). "The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element." Embo J **20**(20): 5748-58.

Tazi, J., S. Durand, et al. (2005). "The spliceosome: a novel multi-faceted target for therapy." Trends Biochem Sci **30**(8): 469-78.

Teigelkamp, S., T. Achsel, et al. (1998). "The 20kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60kD and 90kD proteins." Rna **4**(2): 127-41.

Teraoka, S. N., M. Telatar, et al. (1999). "Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences." Am J Hum Genet **64**(6): 1617-31.

Tian, H. and R. Kole (1995). "Selection of novel exon recognition elements from a pool of random sequences." Mol Cell Biol **15**(11): 6291-8.

Ting, N. S. and W. H. Lee (2004). "The DNA double-strand break response pathway: becoming more BRCAish than ever." DNA Repair (Amst) **3**(8-9): 935-44.

Tsui, S., T. Dai, et al. (2000). "Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1." Genomics **65**(3): 266-73.

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): 928-31.

Umen, J. G. and C. Guthrie (1995). "The second catalytic step of pre-mRNA splicing." Rna **1**(9): 869-85.

Valadkhan, S. and J. L. Manley (2001). "Splicing-related catalysis by protein-free snRNAs." Nature **413**(6857): 701-7.

Valadkhan, S., A. Mohammadi, et al. (2007). "Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing." Rna **13**(12): 2300-11.

Valcarcel, J., R. K. Gaur, et al. (1996). "Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]." Science **273**(5282): 1706-9.

van der Houven van Oordt, W., M. T. Diaz-Meco, et al. (2000). "The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation." J Cell Biol **149**(2): 307-16.

Varani, L., M. Hasegawa, et al. (1999). "Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17." Proc Natl Acad Sci U S A **96**(14): 8229-34.

Varani, L., M. G. Spillantini, et al. (2000). "Structural basis for recognition of the RNA major groove in the tau exon 10 splicing regulatory element by aminoglycoside antibiotics." Nucleic Acids Res **28**(3): 710-9.

Venables, J. P. (2007). "Downstream intronic splicing enhancers." FEBS Lett **581**(22): 4127-31.

Wagner, E. J. and M. A. Garcia-Blanco (2001). "Polypyrimidine tract binding protein antagonizes exon definition." Mol Cell Biol **21**(10): 3281-8.

Wang, H. Y., X. Xu, et al. (2001). "SC35 plays a role in T cell development and alternative splicing of CD45." Mol Cell **7**(2): 331-42.

Wang, J., Y. Takagaki, et al. (1996). "Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability." Genes Dev **10**(20): 2588-99.

Wang, Y., D. Cortez, et al. (2000). "BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures." Genes Dev **14**(8): 927-39.

Wang, Z. and C. B. Burge (2008). "Splicing regulation: from a parts list of regulatory elements to an integrated splicing code." Rna **14**(5): 802-13.

Wang, Z., M. E. Rolish, et al. (2004). "Systematic identification and analysis of exonic splicing silencers." Cell **119**(6): 831-45.

Wang, Z., X. Xiao, et al. (2006). "General and specific functions of exonic splicing silencers in splicing control." Mol Cell **23**(1): 61-70.

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

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): 4614-23.

Williams, R. S. and J. N. Glover (2003). "Structural consequences of a cancer-causing BRCA1-BRCT missense mutation." J Biol Chem **278**(4): 2630-5.

Wirth, B. (2000). "An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA)." Hum Mutat **15**(3): 228-37.

Wirth, B., L. Brichta, et al. (2006). "Spinal muscular atrophy and therapeutic prospects." Prog Mol Subcell Biol **44**: 109-32.

Wu, J. Y. and T. Maniatis (1993). "Specific interactions between proteins implicated in splice site selection and regulated alternative splicing." Cell **75**(6): 1061-70.

Wu, L. C., Z. W. Wang, et al. (1996). "Identification of a RING protein that can interact in vivo with the BRCA1 gene product." Nat Genet **14**(4): 430-40.

Wu, S., C. M. Romfo, et al. (1999). "Functional recognition of the 3' splice site AG by the splicing factor U2AF35." Nature **402**(6763): 832-5.

Xu, B., S. Kim, et al. (2001). "Involvement of Brca1 in S-phase and G(2)-phase checkpoints after ionizing irradiation." Mol Cell Biol **21**(10): 3445-50.

Xu, C. F., J. A. Chambers, et al. (1997). "Mutations and alternative splicing of the BRCA1 gene in UK breast/ovarian cancer families." Genes Chromosomes Cancer **18**(2): 102-10.

Xu, X., D. Yang, et al. (2005). "ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle." Cell **120**(1): 59-72.

Yarden, R. I., S. Pardo-Reoyo, et al. (2002). "BRCA1 regulates the G2/M checkpoint by activating Chk1 kinase upon DNA damage." Nat Genet **30**(3): 285-9.

Yoshida, K. and Y. Miki (2004). "Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage." Cancer Sci **95**(11): 866-71.

Yu, X., C. C. Chini, et al. (2003). "The BRCT domain is a phospho-protein binding domain." Science **302**(5645): 639-42.

Zamore, P. D., J. G. Patton, et al. (1992). "Cloning and domain structure of the mammalian splicing factor U2AF." Nature **355**(6361): 609-14.

Zhang, H., K. Somasundaram, et al. (1998). "BRCA1 physically associates with p53 and stimulates its transcriptional activity." Oncogene **16**(13): 1713-21.

Zhang, M., P. D. Zamore, et al. (1992). "Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit." Proc Natl Acad Sci U S A **89**(18): 8769-73.

Zhang, Q. S., L. Manche, et al. (2006). "hnRNP A1 associates with telomere ends and stimulates telomerase activity." Rna **12**(6): 1116-28.

Zhang, X. H. and L. A. Chasin (2004). "Computational definition of sequence motifs governing constitutive exon splicing." Genes Dev **18**(11): 1241-50.

Zhang, X. H., T. Kangsamaksin, et al. (2005). "Exon inclusion is dependent on predictable exonic splicing enhancers." Mol Cell Biol **25**(16): 7323-32.

Zhao, W. M., C. Jiang, et al. (2001). "A proline-rich protein binds to the localization element of Xenopus Vg1 mRNA and to ligands involved in actin polymerization." Embo J **20**(9): 2315-25.

Zheng, Z. M. (2004). "Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression." J Biomed Sci **11**(3): 278-94.

Zhu, J., A. Mayeda, et al. (2001). "Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins." Mol Cell **8**(6): 1351-61.

Zhuang, Y. and A. M. Weiner (1986). "A compensatory base change in U1 snRNA suppresses a 5' splice site mutation." Cell **46**(6): 827-35.

Zhuang, Y. and A. M. Weiner (1989). "A compensatory base change in human U2 snRNA can suppress a branch site mutation." Genes Dev **3**(10): 1545-52.

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

APPENDIX

The work described in this Thesis has been published in the following scientific journal:

Goina, E.; Skoko, N. and Pagani F.

Binding of DAZAP1 and hnRNP A1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant.

Mol Cell Biol, Jun;28(11):3850-60.

MOLECULAR AND CELLULAR BIOLOGY, June 2008, p. 3850–3860 Vol. 28, No. 11
0270-7306/08/$08.00+0 doi:10.1128/MCB.02253-07

Binding of DAZAP1 and hnRNPA1/A2 to an Exonic Splicing Silencer in a Natural BRCA1 Exon 18 Mutant[▽]

Elisa Goina, Natasa Skoko, and Franco Pagani*

International Center for Genetic Engineering and Biotechnology, Padriciano 99, Trieste 34012, Italy

Received 20 December 2007/Returned for modification 30 January 2008/Accepted 26 March 2008

A disease-causing G-to-T transversion at position +6 of BRCA1 exon 18 induces exclusion of the exon from the mRNA and, as has been suggested by in silico analysis, disrupts an ASF/SF2-dependent splicing enhancer. We show here using a pulldown assay with an internal standard that wild-type (WT) and mutant T6 sequences displayed similar ASF/SF2 binding efficiencies, which were significantly lower than that of a typical exonic splicing enhancer derived from the extra domain A exon of fibronectin. Overexpression or small interfering RNA (siRNA)-mediated depletion of ASF/SF2 did not affect the splicing of a WT BRCA1 minigene but resulted in an increase and decrease of T6 exon 18 inclusion, respectively. Furthermore, extensive mutation analysis using hybrid minigenes indicated that the T6 mutant creates a sequence with a prevalently inhibitory function. Indeed, RNA-protein interaction and siRNA experiments showed that the skipping of T6 *BRCA1* exon 18 is due to the creation of a splicing factor-dependent silencer. This sequence specifically binds to the known repressor protein hnRNPA1/A2 and to DAZAP1, the involvement of which in splicing inhibition we have demonstrated. Our results indicate that the binding of the splicing factors hnRNPA1/A2 and DAZAP1 is the primary determinant of T6 BRCA1 exon 18 exclusion.

Pre-mRNA splicing is an essential step in the gene expression process, connecting DNA transcription to protein translation. The splicing process is catalyzed by the spliceosome, a dynamic complex of five small nuclear ribonucleoproteins (snRNPs) and a large number of additional proteins (19, 30). During pre-mRNA splicing, exon coding sequences must be precisely distinguished from noncoding intronic sequences and joined together to form mature mRNA (2). To achieve this, the splicing machinery requires essential signals located around the splice site junctions: 5′ and 3′ splice sites, polypyrimidine tracts, and branch site sequences (2). However, these splice signals alone provide only a part of the information required by the spliceosome machinery for an efficient splicing process (23, 39). In fact, correct exon definition requires positive and negative accessory elements broadly referred to as exonic splicing enhancers (ESEs) and silencers (ESSs). Most ESEs are known to interact with members of the serine/arginine-rich family, which are splicing factors essential for promoting spliceosome assembly at the correct splice sites (3, 17, 36). In contrast, ESS sequences have often been found to bind specific *trans*-acting factors belonging to the heterogeneous nuclear ribonucleoprotein (hnRNP) family, which plays a role in splicing repression (13, 22, 35). In addition, enhancers or silencers may in some cases coexist in composite exonic splicing regulatory elements, giving rise to complex effects of natural and site-directed mutants on splicing (32, 34). Mutations in exonic splicing regulatory elements that result in splicing alterations are a common event in human pathology (6, 7, 16, 31). The analysis of several gene systems, such as the ATM (40), NF1 (1), CFTR (34), SMN (26), and BRCA1 (27) systems, has shown that exonic mutations can affect, independently from their effect on the amino acid sequence, splicing regulatory sequences and induce different splicing defects, including exon skipping.

A clear identification of the nature and location of exonic splicing regulatory elements is fundamental to understand the effect of genomic variants on splicing and consequently develop appropriate therapeutic strategies. To identify the composition of exonic splicing regulatory sequences and thus predict the splicing phenotype of exonic mutations in human genes, several in silico programs have been developed (10, 14, 15, 41, 42). However, this approach does not always correlate with the splicing phenotype observed, presenting severe limitations concerning its use in clinical genetics (12, 32, 34).

Notwithstanding the relatively high occurrence of defective splicing due to exonic mutations in several disease-related genes, the underlying mechanism is still controversial. Two different models have been proposed to explain exon skipping due to exonic mutations. The enhancer loss model suggests the disruption of sequences with enhancer properties, with ASF/SF2 being the main binding factor involved. On the contrary, the silencer gain-of-function model implies the creation of a silencer sequence by the mutation, which is mostly recognized by members of the hnRNP family. These two models have been extensively explored in the SMN1/2 systems, in which a critical C-to-T synonymous substitution in position 6 induces exon 7 skipping (8, 9, 20, 21, 37). The natural G-to-T mutation at position +6 of BRCA1 exon 18 (E1694X) presents common features with the SMN system. This mutation occurs at the same exonic position of SMN2; it has been previously reported to be associated with exon skipping (27), and both computer-assisted analysis and in vitro splicing analysis of a limited number of mutants have suggested that it disrupts a putative binding site for ASF/SF2 (25). However, this binding has never

* Corresponding author. Mailing address: Human Molecular Genetics, International Center for Genetic Engineering and Biotechnology, Padriciano 99, 34012 Trieste, Italy. Phone: 0039-040-375-7342. Fax: 0039-040-226-555. E-mail: pagani@icgeb.org.

▽ Published ahead of print on 7 April 2008.

been experimentally proven, although the BRCA1 exon 18 system has been used for the in vitro characterization of new ASF/SF2-specific exonic splicing enhancers by systematic evolution of ligands by exponential enrichment (SELEX) (38). Small interfering RNA (siRNA) of ASF/SF2 has been recently shown not to have any functional effect on wild-type (WT) BRCA1 exon 18 splicing, whereas the BRCA1 E1694X mutant was found to bind hnRNPA1/A2 (21). However, as depletion of both hnRNPA1 and hnRNPA2 had no effect on mutant BRCA1 splicing, it was suggested that the disruption of enhancer binding to an unknown splicing factor contributes to mutant exon skipping (21).

Our work evaluates extensively the composition of the splicing-regulatory element affected by the natural G-to-T transversion at position +6 in BRCA1 exon 18. The analysis we performed includes systematic site-directed mutagenesis in hybrid minigenes, in vitro binding experiments, and siRNA treatment of the splicing factors involved. We find no preferential binding of ASF/SF2 to the WT sequence compared with the T6 mutant. On the contrary, the BRCA1 exon 18 T6 substitution creates a sequence with silencer properties that binds to hnRNPA1/A2, a known repressor protein, and deleted in azoospermia-associated protein 1 (DAZAP1). Altogether, our data support the gain-of-function model for defective splicing due to the +6 G-to-T exonic mutation in BRCA1 exon 18.

MATERIALS AND METHODS

Hybrid minigene constructs. The BRCA1 genomic region containing the last 85 bp of exon 17 and first 1,408 bp of intron 17 was amplified from normal DNA using BRC90BstEII Dir (5′-CTGGTACCAAGTTTGCCAGAAAACACCACA TCACTTTAACTAATC-3′) and BRC1566 Rev (5′-AACACCCAGAGGTCTC CTGTATTACACAAG-3′) primers and blunt ligated in the SmaI site of pBluescript KS to generate pBS BRA17. Similarly, the last 334 bp of intron 18 along with 58 bp at the beginning of exon 19 was amplified with BRC735 Dir (5′-TA GCAATGTAGCATATGAGCTAGGGATTTA-3′) and BRC4451 Rev (5′-AA CATCAAGTACTTACCTCATTCAGC-3′) primers and blunt ligated in the SmaI site of pBluscript KS to generate pBS BRA19. Subsequently, pBS BRA19 was digested with NdeI and KpnI and cloned in the corresponding sites of pBS BRA17 to generate pBS BRA17-19, with NdeI as a unique site. The BstEII-BstEII fragment contained in pBS BRA17-19 was then inserted in the unique BstEII site of a modified pTB α-globin minigene to obtain pBRAint. The BRCA1 genomic region encompassing exon 18 and adjacent intronic sequences (289 bp of intron 17, full-length exon 18, and 255 bp of intron 18) was amplified with BRC138 Dir (5′-GGCATATGGAGATCTATAGCTAGCCTTGGCGTCT AGAAGATGG-3′) and BRC760 Rev (5′-AATCCCTAGCTCATATGCTAAC ATTGCTAGG-3′) primers and cloned in the unique NdeI site of pBRAint. On this resulting construct, an EcoRI site was inserted in intron 18 (49 bases downstream of the 5′ splice site of exon 18) by PCR-mediated site-directed mutagenesis, creating the final hybrid minigene, named pBRAwt. The natural PstI and the artificial EcoRI unique sites create a cassette system, which facilitates subsequent cloning procedures. The final minigene lacks the central part of intron 17 (1,960 bp), and the α-globin and BRCA1 17, 18, and 19 exons are in frame. pBRA minigene variants were generated by PCR site-directed mutagenesis replacing the pBRAwt PstI-EcoRI cassette with the mutated sequence. The constructs containing the serial deletions (~10 bp) within exon 18 were created by overlapping PCR mutagenesis using suitable primers on pBRA WT or T6 mutant minigenes as templates. The identity of each minigene was verified by sequence analysis. The pSMN2 construct has been previously reported (8).

Cell culture, transfections, and reverse transcription-PCR (RT-PCR) analysis. Human cell lines (Hep3B and HeLa) were cultured in Dulbecco's modified Eagle's medium with Glutamax (Invitrogen), in standard conditions. The DNA used for transfection was purified using JetStar columns (Genomed). Hep3B cells (3×10^5) were transfected with 1.5 μg of each reported plasmid construct, employing the DOTAP (1,2-dioleoyl-3-trimethylammonium-propane) liposomal transfection reagent (Alexis Biochemicals) according to the manufacturer's instructions. After 12 h, the transfection medium was replaced with fresh medium, and 24 h later the total RNA was extracted using TRIreagent solution (Ambion). Retrotranscription was performed with random primers and Moloney murine leukemia virus enzyme (Invitrogen). To amplify only the messengers derived from the transfected plasmid, PCR was performed with BRC90BsteII Dir and Glo800Rev (5′-GCTCACAGAAGCCAGGAACTTGTCCAGG-3′), which hybridize to BRCA1 exon 17 and α-globin exon 3 sequences, respectively. The conditions used for the PCRs were the following: 94°C for 3 min for the initial denaturation; 94°C for 45 s, 54°C for 45 s, and 72°C for 45 s for 30 cycles; and 72°C for 10 min for the final extension. PCRs were optimized to be in the exponential phase of amplification, and products were routinely fractionated in 2% (wt/vol) agarose gels. For protein overexpression experiments, 1.5 μg of the minigene was cotransfected with 1 μg of the ASF/SF2 coding sequence cloned into a pCG vector (a kind gift from J. Caceres). Amplification of the pSMN2 minigene was performed with the previously reported oligonucleotides (8). The results of all the transfections are representative of at least three independent experiments. ImageJ 1.38 software (http://rsb.info.nih.gov/ij/) was used in order to quantify the proportion of exon 18 skipping.

Affinity purification of RNA binding proteins. Two synthetic RNA oligonucleotides, exon 18 WT (UGCAGAUGCUGAGUUUGUGU) and U6 (UGCA GAUGCUUAGUUUGUGU), were generated by Integrated DNA Technologies and used as targets for pulldown assays. Twelve micrograms of target BRCA1 RNA oligonucleotides was placed in 400 μl of reaction mixture (100 mM sodium acetate [NaOAC], pH 5.0, and 5 mM sodium *m*-periodate Σ), incubated for 1 h in the dark at room temperature, ethanol precipitated, and finally resuspended in 100 μl of 100 mM NaOAc (pH 5.0). Approximately 400 μl of adipic acid dehydrazide-agarose beads (50% slurry; Sigma) previously equilibrated with 100 mM NaOAC (pH 5.2) was added to each periodate-treated RNA, and the mixture was incubated for 12 h at 4°C on a rotator.

The beads with the bound RNA were then washed two times with 1 ml of 2 M NaCl and equilibrated in 1× washing buffer (5.2 mM HEPES, pH 7.5, 1 mM $MgCl_2$, 0.8 mM Mg acetate). Then the beads were incubated, in a final volume of 500 μl, with 0.5 mg of HeLa cell nuclear extract (C4; Biotech), 1× binding buffer (5.2 mM HEPES, pH 7.9, 1 mM $MgCl_2$, 0.8 mM Mg acetate, 0.52 mM dithiothreitol, 3.8% glycerol, 0.75 mM ATP, 1 mM GTP), and heparin (final concentration, 1 μg/μl) for 30 min on a rotator at room temperature. The beads were then washed four times with 1.5 ml of washing buffer before addition of sodium dodecyl sulfate (SDS) sample buffer and loading on SDS–10% polyacrylamide gels.

Proteins were visualized by Coomassie brilliant blue staining. Protein sequence analysis of the bands excised from the gel was performed using an electrospray ionization mass spectrometer (LCQ DECA XP; ThermoFinnigan). Protein bands were digested with trypsin, and the resulting peptides were extracted with water and 60% acetonitrile–1% trifluoroacetic acid. Fragments were then analyzed by mass spectrometry, and proteins were identified by analysis of the peptide tandem mass spectrometry data with Turbo SEQUEST (ThermoFinnigan) and MASCOT (Matrix Science).

In vitro transcription and Western blot analysis. Plasmids for in vitro transcription were generated by annealing the sense (5′-CTGTGTGTGTGTGTTTTT TGCAGATGCTKAGTTTGTGTG-3′) and the antisense (5′-GATCCACACAAA CTMAGCATCTGCAAAAAACACACACACACAGAGCT-3′) oligonucleotides, purchased from Sigma, for both WT and T6 mutant exon 18 sequences. This was followed by direct cloning into SacI-BamHI-digested pBluescript vector, under the control of the T7 RNA promoter.

Part of the fibronectin extra domain A (EDA) sequence (5′-CTGTGTGTGT GTGTTTTTTGCACCTGATGGTGAAGAAGACACTGCAGAGC-3′) was also cloned by following the same procedure and used for pulldown analyses, acting as internal control for ASF/SF2 binding.

Plasmids were first linearized by HindIII digestion and then in vitro transcribed with T7 RNA polymerase (Promega) according to standard procedures. Approximately 12 to 15 μg of transcribed and purified RNA was placed in a 400-μl reaction mixture of 100 mM NaOAc, pH 5.0, and 5 mM sodium *m*-periodate (Sigma) and processed as described above for pulldown assays. After the final centrifugation, 60 μl of SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer was added to the samples, followed by heating for 5 min at 90°C before loading onto a 12% SDS-PAGE gel. The gel was electroblotted on a polyvinylidene difluoride membrane according to standard protocols (Amersham Biosciences), and the membrane was blocked with phosphate-buffered saline–5% skimmed milk. Proteins were identified using different antibodies, and Western blot signals were detected with a chemiluminescence kit (ECL; Pierce Biotechnology).

Rabbit anti-hnRNPA1 serum was generously provided by R. Klima, and polyclonal antibodies against hnRNPA2 and TDP43 were provided by E. Buratti (ICGEB, Trieste, Italy). Purified glutathione *S*-transferase (GST)–DAZAP1

protein was used to immunize a rabbit (New Zealand strain) according to standard protocols to obtain polyclonal anti-DAZAP1 antibodies. Anti-ASF/SF2 monoclonal antibody 96 was purchased from Zymed Laboratories Inc., and antitubulin monoclonal antibody was kindly provided by F. Porro (ICGEB, Trieste, Italy).

Electromobility shift assay (EMSA). RNA synthetic oligonucleotides (200 ng) were labeled by phosphorylation with [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs) for 1 h at 37°C, precipitated, and resuspended in 200 μl of water. Each binding reaction was made by mixing the purified protein with the labeled RNA oligonucleotide in 1× binding buffer (5.2 mM HEPES, pH 7.9, 1 mM $MgCl_2$, 0.8 mM Mg acetate, 0.52 mM dithiothreitol, 3.8% glycerol, 0.75 mM ATP, 1 mM GTP) to a 20-μl final volume. The reaction mixture was left at room temperature for 15 to 20 min before loading the sample on a native polyacrylamide gel (5%), which was run at 100 to 120 V at 4°C. The protein-nucleic acid complexes were visualized using Biomax MS films (Kodak). The cDNA of each target protein was amplified and cloned in pGEX-3X plasmid (Pharmacia) and then expressed in BL21(DE3) bacteria (Novagen) under the induction of 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside). The resulting recombinant proteins were purified with glutathione *S*-Sepharose 4B beads (Pharmacia) according to the manufacturer's instructions using imidazole buffers for protein elution.

siRNA transfections. siRNA transfections were performed in HeLa cells using Oligofectamine reagent (Invitrogen). The sense strands of RNA interference oligonucleotides (Dharmacon) used for silencing the different target proteins were the following: human hnRNPA1, CAGCUGAGGAAGCUCUUCA, and human hnRNPA2, GGAACAGUUCCGUAAGCUC (20); human DAZAP1, GAGACUCUGCGCAGCUACU; human ASF/SF2, ACGAUUGCCGCAUCU ACGU (8); and luciferase no. 2 gene control, GCCAUUCUAUCCUCUAGA GGAUG.

HeLa cells were plated at 2.5×10^5 cells per well in 60-mm plates to achieve 40 to 50% confluence. The next day, 6 μl Oligofectamine was combined with 24 μl of Opti-MEM medium (Invitrogen) and 5 to 10 μl of 40 μM siRNA duplex oligonucleotides was diluted in a final volume of 400 μl of Opti-MEM medium. The two mixtures were combined and left for 20 min at room temperature. Finally, this mixture was added to the cells, which were maintained in 1.6 ml of Opti-MEM. After 24 h a second round of siRNA transfection was performed as described above. Six to 8 h later Opti-MEM was exchanged with Dulbecco's modified Eagle medium and the cells were transfected with the minigene of interest (1 μg) using Qiagen Effectene transfection reagents. On the third day, HeLa cells were harvested and divided in two parts for protein and RNA extractions. RT-PCR from total RNA was performed as for the transfection protocol described above. Whole-protein extracts were obtained by cell sonication in lysis buffer (15 mM HEPES, pH 7.5, 250 mM NaCl, 0.5% NP-40, 10% glycerol, and 1 mM phenylmethylsulfonyl fluoride) and analyzed for hnRNPA1 and -A2, DAZAP1, and ASF/SF2 endogenous protein expression by immunoblotting using the antibodies described above. Tubulin was used as protein loading control. Each siRNA treatment experiment was repeated at least three times.

RESULTS

ASF/SF2 affects exon 18 splicing independently from the T6 mutation. To evaluate the role of ASF/SF2 in the regulation of BRCA1 splicing, we set up a modified pulldown assay where exon 18 WT or U6 RNAs transcribed in vitro were linked through a $(U)_5$ spacer to a sequence composed by six UG dinucleotide repeats (Fig. 1A, top). The UG repeats, which specifically bind TDP43 (4), served as an internal standard to normalize the amount of RNAs pulled down. An ASF/SF2 purine-rich enhancer, already described regarding the alternative splicing of the fibronectin EDA exon (29), was used as a positive control to estimate the binding efficiency of the BRCA1 sequences. The pulldown assay was followed by Western blotting and probing with ASF/SF2 and TDP34 antibodies. Relative to the binding of TDP43, used as a normalization reference, the WT and U6 mutant RNAs bind to ASF/SF2 with a similar, low efficiencies in comparison with the binding of ASF/SF2 to the strong EDA enhancer (Fig. 1A, bottom, lanes 2, 3, and 1, respectively). The low binding efficiencies of both WT and mutant sequences suggest a common, minor splicing enhancing effect of ASF/SF2. These results were confirmed in classical pulldown experiments using RNA oligonucleotides without the UG tail (data not shown).

To better evaluate the effect of ASF/SF2 on human BRCA1 exon 18 splicing, we prepared a pBRA hybrid minigene. The pBRA minigene system contains most of the genomic sequences of the BRCA1 gene from exon 17 to 19, cloned in frame within the α-globin pTB minigene (Fig. 1B) (33). We tested the in vivo splicing pattern of the pBRA WT and T6 constructs by transient transfection of HeLa cells followed by RT-PCR analysis of the total RNA. The two minigenes spliced in different ways: while WT exon 18 was efficiently included (Fig. 1C, lane 1), the T6 mutated exon was predominantly skipped, with only 20% of it being included (Fig. 1C, lane 3). Cotransfection experiments using a plasmid coding for ASF/SF2 showed an increase in the percentage of exon 18 inclusion for the T6 mutant minigene (Fig. 1C, lane 4), whereas the enhancing effect was not detectable for the WT minigene since exon 18 is already fully included in the absence of ASF/SF2 overexpression (Fig. 1C, lane 2).

We further investigated the in vivo role of ASF/SF2 by siRNA treatment of HeLa cells. In this case the WT and T6 pBRA constructs were transfected in siRNA-treated cells. After 24 h cells were collected and analyzed for the level of endogenous ASF/SF2 by Western blotting. We observed a strong reduction in ASF/SF2 expression in cells transfected with siRNA but not in luciferase siRNA-treated cells, used as control (Fig. 1D, left). RT-PCR analysis of the total RNA showed that ASF/SF2 depletion induced a nearly complete exclusion of T6 exon 18 from the mature mRNA (Fig. 1D, right, lane 3). In contrast, WT BRCA1 exon 18 did not respond to siRNA treatment (Fig. 1D, right, lane 1). Overall, these results do not support a direct role for ASF/SF2 in the aberrant splicing of T6 mutant exon 18.

The natural G-to-T transition in BRCA1 exon 18 creates a sequence with silencer properties. To understand the composition of the splicing-regulatory element and the effect of the G6T natural mutation, we performed systematic site-directed mutagenesis. A series of single point mutations from position +4 to position +11 were introduced in the pBRA hybrid minigene and evaluated by means of the functional-splicing assay. Interestingly, out of 21 substitutions only the natural T6 mutation resulted in a very low level of exon inclusion (20%) (Fig. 2A, lane 10). Two additional mutations, in positions 4 (T4) and 6 (A6), partially affected the splicing pattern, causing 85% and 70% exon inclusion, respectively (Fig. 2A, lanes 4 and 8). As the widely spanning mutagenesis analysis seemed more compatible with the creation of a negative element by the T6 variant, we decided to test this putative silencer by introducing in the context of the T6 substitution consecutive point mutations from position +4 to +9. As shown in Fig. 2B, the majority of the double site-directed mutations flanking position 6 increased the splicing efficiency compared with that for the single T6 variant. In particular, all of the substitutions in positions 7 and 8 restored almost completely exon inclusion (Fig. 2B, lanes 9 to 14), and three mutations in position 5 and two in position 9 (T6A9 and T6C9) increased significantly the splicing efficiency (Fig. 2B, lanes 6 to 8, 15, and 16), whereas T6G9 had no effect (Fig. 2B, lane 17). In contrast, mutations at position +4

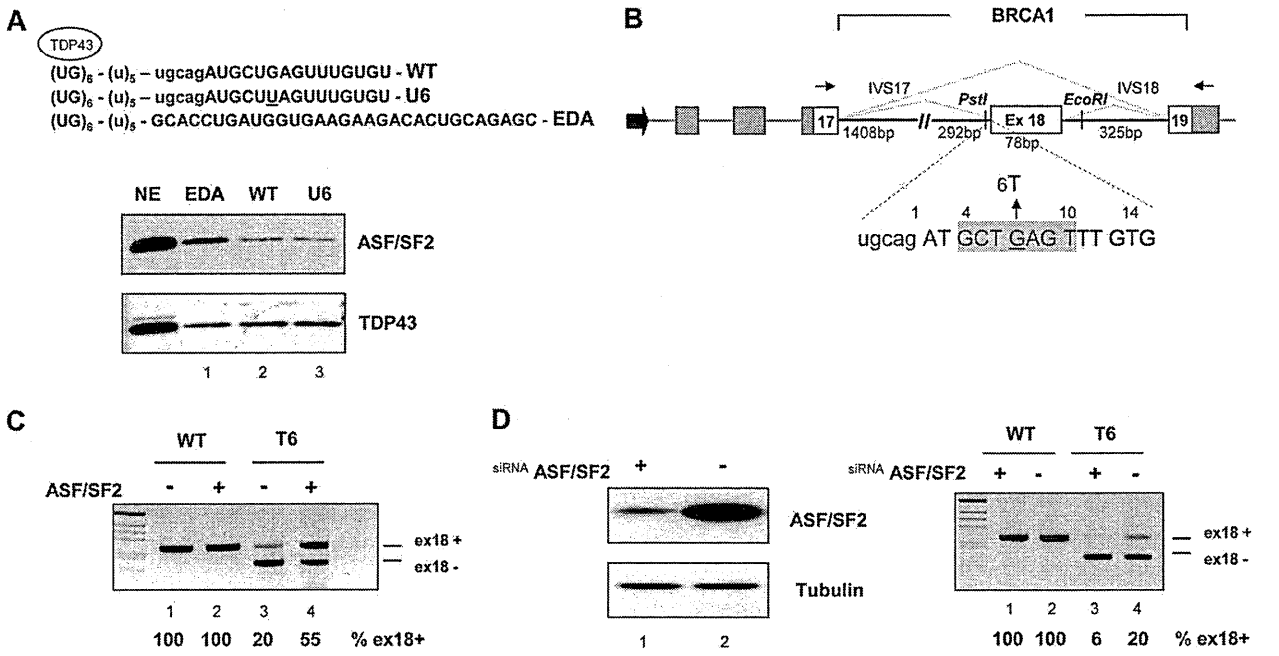

FIG. 1. ASF/SF2 affects exon 18 splicing independently from the T6 mutation. (A) Western blot after pulldown analysis of WT and U6 sequences. Shown is a schematic representation of the RNA sequence transcribed in vitro (top) that contains $(UG)_6$ repeats, a $(U)_5$ spacer, part of intron 17 (lowercase), and the first part of WT and T6 exon 18 (uppercase). The mutant nucleotide U6 is underlined. The bottom lane shows the EDA sequence used as a control for ASF/SF2 binding. TDP43 was used to normalize the assay for the amount of RNA. The nuclear extract sample (NE) corresponds to 1/20 of the amount used for the pulldown assay. (B) Schematic representation of the pBRA hybrid minigene system. Dark gray and white boxes represent α-globin and BRCA1 exons, respectively. Thick lines are introns (IVS), and the large black arrow at the 5′ end indicates the simian virus 40 promoter. The lengths of BRCA1 exon 18 and flanking introns are shown. The primers used for RT-PCR are indicated by thin arrows, and the two alternative splicing possibilities, exon 18 inclusion and exclusion, are indicated as dotted lines. The positions of the unique PstI and EcoRI restriction sites used for subsequent cloning are shown. At the bottom, a few bases from intron 17 (lowercase) and exon 18 (from position +1 to +14) are shown, divided in codon triplets (uppercase). The putative ASF/SF2 binding site predicted by the ESE finder program (25) is boxed in light gray. The G-to-T natural mutation (E1694X) is indicated (27). (C) In vivo overexpression of ASF/SF2 promotes exon 18 T6 inclusions. HeLa cells were transfected with the WT or T6 pBRA minigene as indicated, along with the ASF/SF2-coding plasmid (+) or with the empty vector (−). Splicing patterns were analyzed by separating the RT-PCR products on 2% agarose gel and staining with ethidium bromide. The identity of the transcripts, i.e., including exon 18 (ex18+) and lacking exon 18 (ex18−), is indicated at the right side of the gel. The percentages of exon 18 inclusion, obtained from three independent transfections, are indicated below the lane numbers. (D) In vivo depletion of ASF/SF2 induces complete T6 exon 18 skipping. The left panel shows the Western blot analysis of extracts from HeLa cells treated with either ASF/SF2 siRNA (+) or a control luciferase siRNA (−). The amount of the protein extract was normalized using an antibody against tubulin. The right panel shows the RT-PCR results after transfection of the WT or T6 pBRA plasmid minigenes into ASF/SF2 siRNA-treated (lanes 1 and 3) and luciferase control-treated (lanes 2 and 4) HeLa cells, and the percentages of BRCA1 exon 18 inclusion are indicated.

negatively affected splicing, resulting in complete exon 18 skipping (Fig. 2B, lanes 3 to 5). Thus, double substitutions that restore normal exon 18 inclusion disrupt the mutant "TAG" sequence at the 5′ end of exon 18. The "TAG" sequence therefore appears to be the core silencer element involved.

On the basis of this assumption, we evaluated splicing-regulatory sequences derived from the SMN1 and SMN2 genes (20) and a SELEX A1 silencer sequence (5) in the pBRA minigene context. These sequences include or do not include the "TAG" core element (Fig. 2C, left). Exon 18 splicing was affected by the insertion of the SELEX A1 sequence at the same level as with the T6 mutant (Fig. 2C, lane 5). The SMN2 sequence in the pBRA minigene showed a lower but still negative effect, with 50% exon 18 inclusion, whereas SMN1, which does not contain TAG, did not affect mRNA processing (Fig. 2C, lanes 4 and 3, respectively). Finally, a deletion from position +4 to +9 in the pBRA Δ4-9 construct did not change the splicing pattern in comparison with WT pBRA, further ruling out the presence of a strong enhancer at the 5′ end of exon 18 (Fig. 2C, lane 6).

An indirect effect of nonsense-mediated decay on splicing in this system was excluded, as two nonsense-coding minigenes, T6A8 and 10A11G, and the out-of-frame Δ4-9 mutant showed, unlike the T6 variant (E1694X), complete exon inclusion (Fig. 2B, lanes 12 and 18, and C, lane 6).

Identification of nuclear proteins binding to the splicing-regulatory element of the BRCA1 exon 18 T6 mutant. To identify the *trans*-acting factor(s) differentially binding to BRCA1 exon 18 WT and mutant sequences, we performed a pulldown analysis using two synthetic RNA oligonucleotides containing either the WT or the natural U6 mutated sequence (Fig. 3A). Agarose bead-linked RNAs were incubated with HeLa nuclear extracts, and the associated proteins were analyzed by SDS-PAGE followed by Coomassie blue staining. Comparison of the patterns of binding proteins showed two specific bands of about 50 and 38 kDa that were specifically

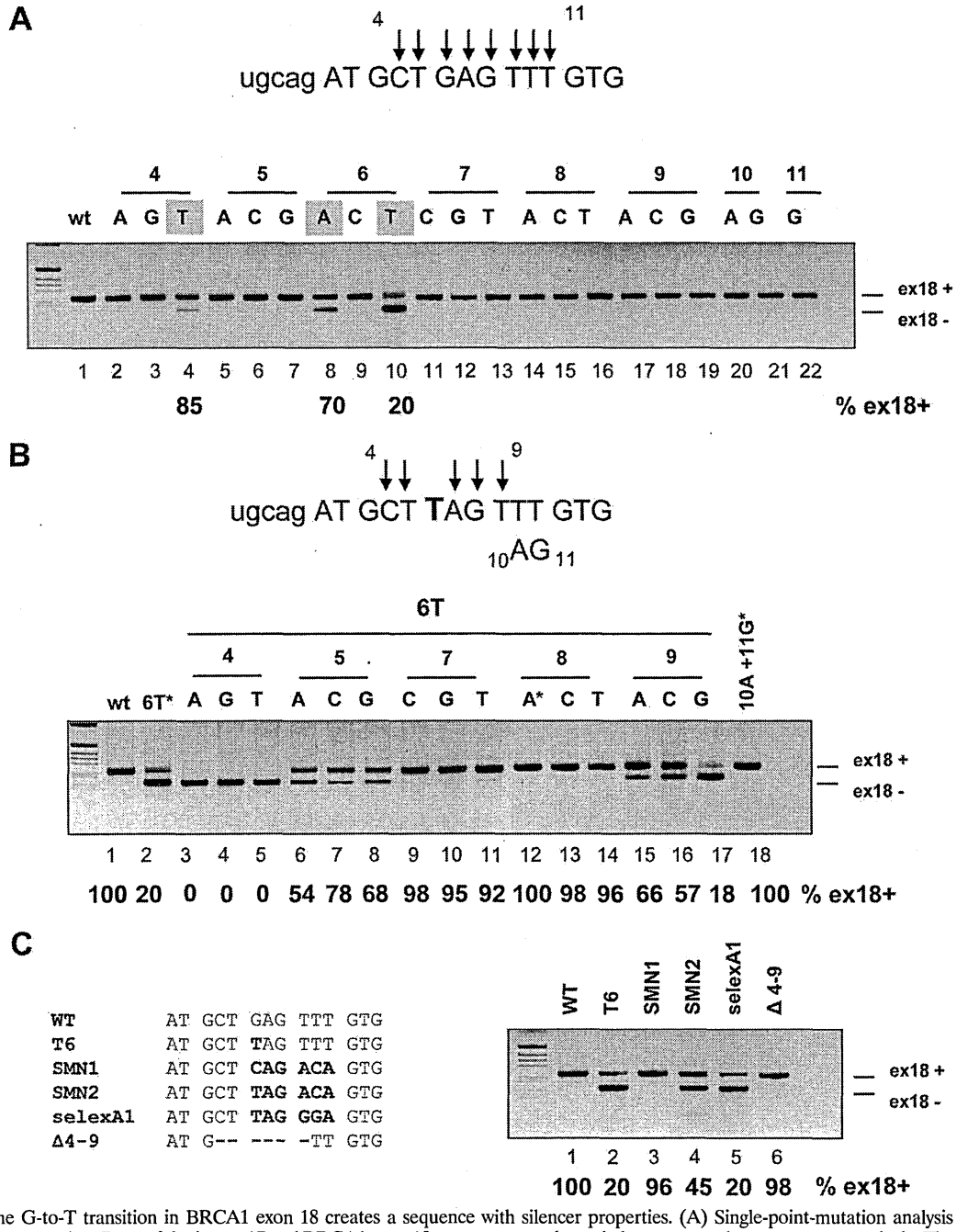

FIG. 2. The G-to-T transition in BRCA1 exon 18 creates a sequence with silencer properties. (A) Single-point-mutation analysis of the exon 18 putative regulatory region. Parts of the intron 17 and BRCA1 exon 18 sequences are shown in lowercase and uppercase, respectively. The arrows indicate the nucleotide positions where all three possible single point mutations were substituted, from nucleotide 4 to 11. The constructs were transfected into Hep3B cells, and the splicing pattern was analyzed by RT-PCR. The splicing products corresponding to exon 18 inclusion (ex18+) and exclusion (ex18−) are indicated. The gray boxes indicate the mutants that showed exon skipping (lanes 4, 8, and 10). Their percentages of exon 18 inclusion are indicated below the lanes. (B) Double-point-mutation analysis. The upper panel shows part of the sequence of intron 17 and exon 18. The positions of single point mutations, introduced in the context of the 6T variant, are indicated by arrows. The position of the $_{10}AG_{11}$ double mutant is indicated. The substitutions that create a stop codon are highlighted by an asterisk (lanes 2, 12, and 18). The percentages of BRCA1 exon 18 inclusion, calculated as for panel A, are indicated below the lanes. (C) Analysis of silencer sequences in the BRCA1 exon 18 context. The left panel shows WT and mutant sequences of BRCA1 exon 18 from position +1 to +14. The mutant elements are in boldface, and the deletion is represented as a dashed line (Δ4-9). The splicing profile was analyzed as described for panel A. The percentages of BRCA1 exon 18 inclusion are indicated.

pulled down by the U6 but not by the WT RNA (Fig. 3B). Sequence analysis of the excised bands and a database search revealed that the band of 50 kDa is DAZAP1, also named proline-rich RNA-binding protein, and that the band of 35 kDa corresponds to both hnRNPA1 and its close homologue hnRNPA2 factor.

To investigate in more detail the differential binding efficiency of the splicing-regulatory sequences of BRCA1 exon 18,

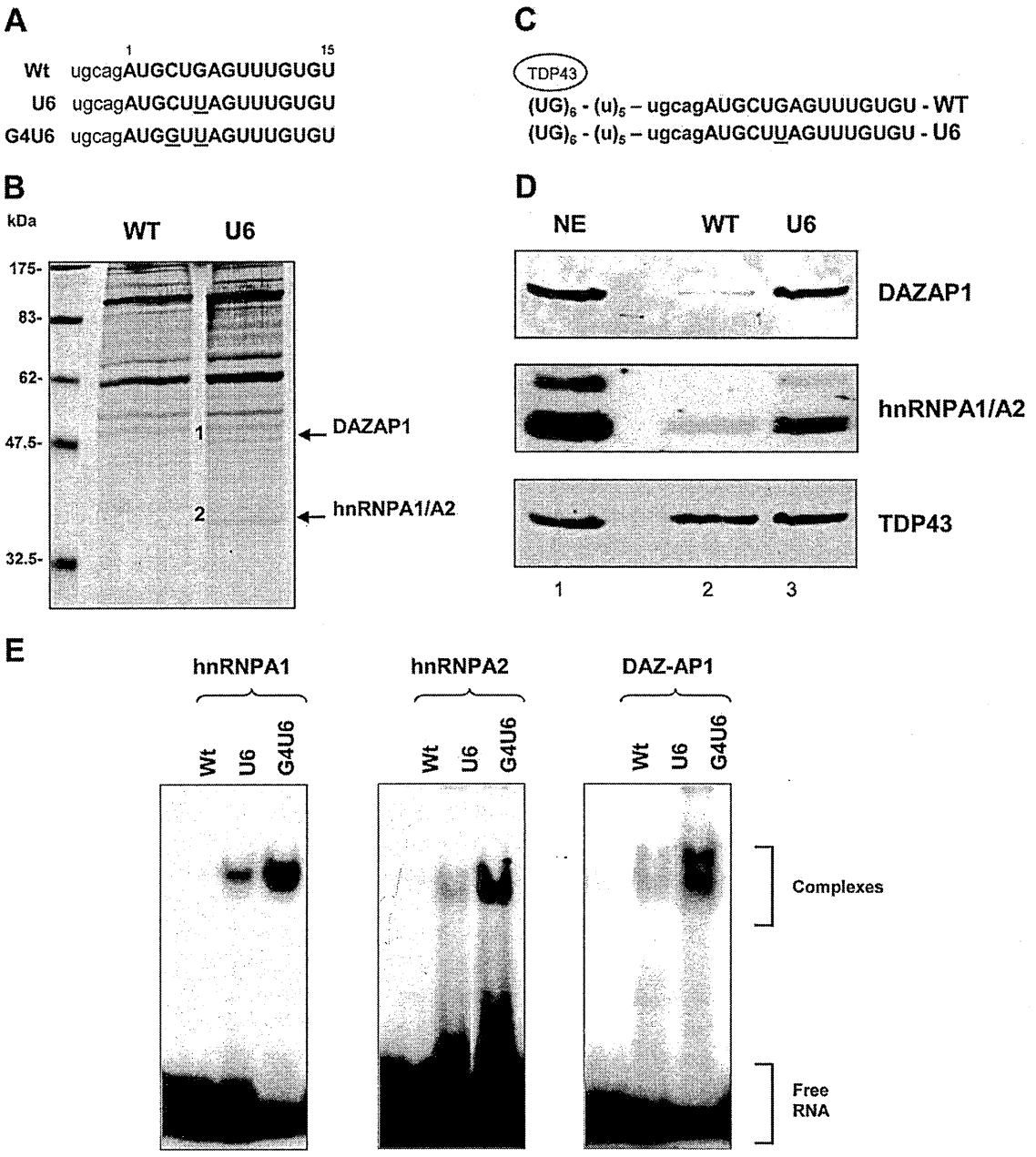

FIG. 3. Identification of nuclear proteins binding to the splicing-regulatory element of the BRCA1 exon 18 T6 mutant. (A) Sequences of the BRCA1 RNA oligonucleotides. Intronic and exonic sequences are shown in lowercase and uppercase, respectively, and the mutated nucleotides are underlined. (B) Coomassie-stained gel of a pulldown analysis where HeLa nuclear extracts were incubated with adipic dehydrazide beads derivatized with the target RNAs. The arrows indicate two bands of about 50 and 38 kDa, present only in the U6 lane. The proteins in these bands were sequenced and identified as DAZAP1 (1) and hnRNPA1/A2 (2), respectively. (C) Schematic representation of the RNA sequence transcribed in vitro that contains $(UG)_6$ repeats, a $(U)_5$ spacer, and the target BRCA1 sequences. The mutagenized nucleotide is underlined. (D) Binding analysis of hnRNPA1 and -A2 and DAZAP1 performed by Western blotting after pulldown assays. After in vitro transcription, RNA was pulled down and the bound proteins were analyzed by Western blotting with anti-hnRNPA1, -hnRNPA2, -DAZAP1, and -TDP43 polyclonal antibodies. The nuclear extract sample corresponds to 1/20 of the amount used for the pulldown assay. (E) Interaction of purified human GST-hnRNPA1 and -A2 and DAZAP1 proteins with WT, U6, and G4U6 RNA oligonucleotides monitored by EMSA. Labeled RNAs were incubated with different GST fusion proteins as indicated, and the complexes were resolved on a 5% polyacrylamide gel. The free RNA and the formed complex are indicated on the right.

we prepared in vitro T7-transcribed RNAs that contained a fixed $(UG)_6$ repeat sequence linked through a poly(U) linker to WT and U6 exon 18 (Fig. 3C). After pulldown and SDS-PAGE, the proteins were blotted onto polyvinylidene difluoride filters and evaluated using specific antibodies. In the presence of comparable amounts of pulled down TDP43, the U6 mutant but not the WT bound very efficiently to DAZAP1 and hnRNPA1/A2 (Fig. 3D). These results were also confirmed by

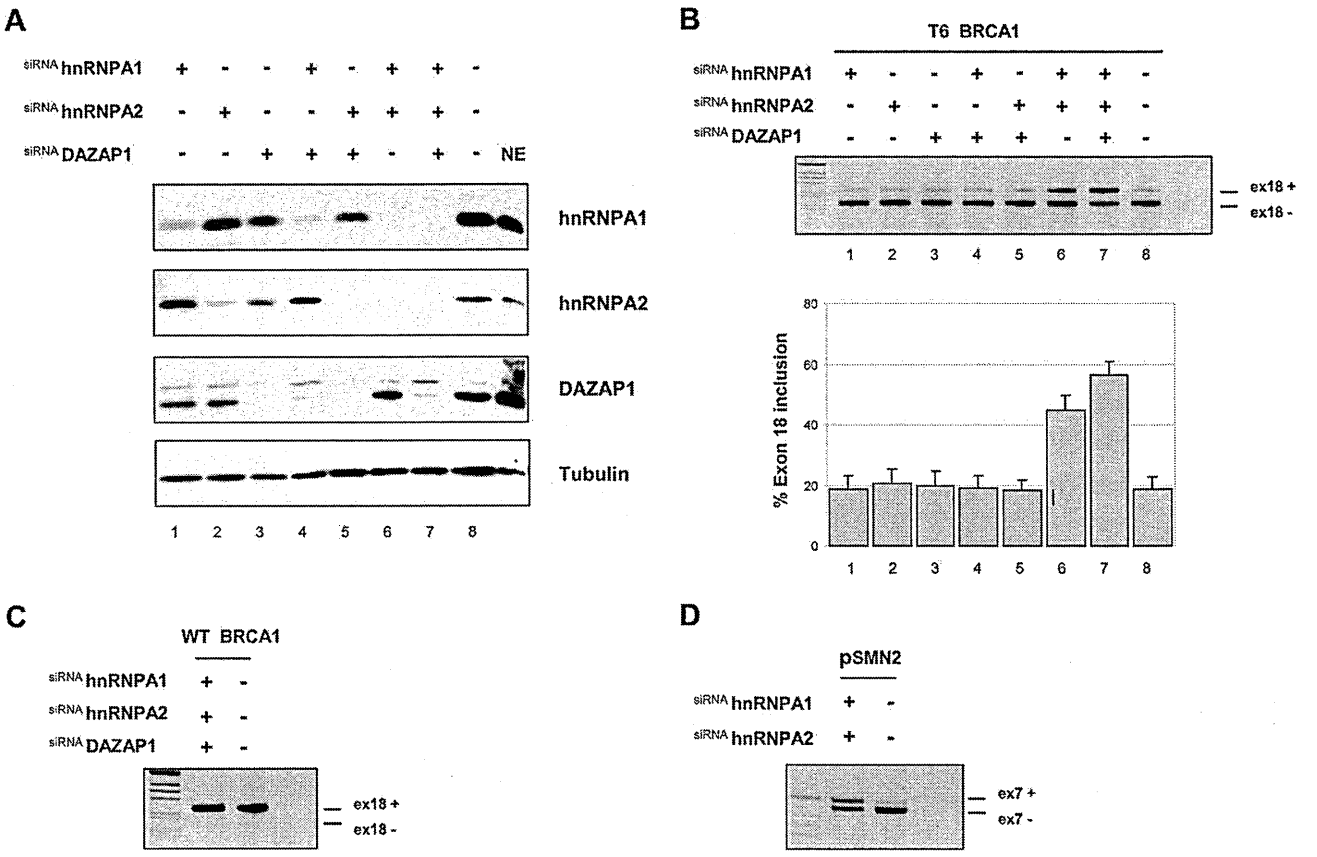

FIG. 4. In vivo depletion of hnRNPA1 and -A2 and DAZAP1 by siRNA treatment partially rescues BRCA1 exon 18 inclusion. (A) Western blot analysis of HeLa cells treated (+) with the indicated different siRNAs. NE, nuclear extract. Lane 8 corresponds to the control luciferase siRNA. (B) siRNA-treated (+) and untreated (−) cells were transfected with the pBRA T6 minigene plasmid. Total RNA was prepared and analyzed by RT-PCR. The splicing products obtained were separated on 2% agarose gels and stained with ethidium bromide. The identities of the transcripts including exon 18 (ex18+) and lacking exon 18 (ex18−) are indicated. The graph is the quantification of three independent experiments expressed as means ± standard deviations. (C) Effect of siRNA against hnRNPA1, hnRNPA2, and DAZAP1 on the WT pBRCA1 minigene. (D) Effect of double siRNA against hnRNPA1 and hnRNPA2 on the splicing pattern of the control pSMN2 minigene. pSMN2 and pBRCA1 T6 minigenes were cotransfected in the siRNA hnRNPA1- and hnRNPA2-treated cells and analyzed with specific primers for pSMN2. The identity of the transcript with inclusion or exclusion of the SMN2 exon 7 is indicated.

classical pulldown experiments using synthetic RNA oligonucleotides (data not shown).

In order to test the specificity of the three different splicing factors for binding to BRCA1, we performed EMSA with purified hnRNPA1, hnRNPA2, and DAZAP1 GST-tagged proteins. For the EMSA experiments we selected the WT and the U6 RNAs and an additional double mutant, G4U6, which induced complete exon skipping in a splicing assay (Fig. 2B, lane 4). Labeled WT and mutant RNAs were incubated with recombinant proteins, and the RNA-protein complexes were resolved on a native acrylamide gel. A band of shifted material with the RNAs carrying the T6 and G4T6 mutations is present, whereas no complexes are formed by the WT oligonucleotide (Fig. 3E). This difference in binding efficiency is visible for all the three proteins analyzed, and it is in agreement with the percentage of exon skipping detected in the minigene splicing assay.

siRNA against hnRNPA1 and -A2 and DAZAP1 affects BRCA1 T6 exon 18 splicing. hnRNPA1 and the related hnRNPA2 protein are well-known splicing-inhibitory factors, whereas no functional role for DAZAP1 in splicing regulation has been reported. In order to evaluate the inhibitory role of hnRNPA1 and -A2 and DAZAP1 in the in vivo splicing of BRCA1 exon 18, we used single or combined siRNA against these splicing factors. We transiently transfected HeLa cells, and the silencing effect was analyzed by Western blotting using antibodies against each target protein. At the same time, the splicing pattern was evaluated by RT-PCR using total RNA. HeLa cells transfected with the corresponding siRNAs showed a strong reduction in hnRNPA1 and -A2 and DAZAP1 levels (Fig. 4A), whereas no effect was detected in the luciferase control protein extract (Fig. 4A, lane 8). Unexpectedly, DAZAP1 showed a small reduction in the amount of hnRNPA1 and hnRNPA2 (Fig. 4A, lane 3), possibly related to an in vivo physical interaction (24) or to a regulatory effect of DAZAP1 on hnRNPA1/A2 splicing and/or stability. In the T6 mutant, the combined siRNA-mediated knockdown of the three splicing-inhibitory factors increased the percentage of BRCA1 exon 18 inclusion from 20 to about 55% (Fig. 4B, lane 7). A slightly smaller effect was observed for the siRNA-me-

diated knockdown of hnRNPA1/A2 (Fig. 4B lane 6), with about 45% exon inclusion, whereas the individual and double (hnRNPA1/DAZAP1 and hnRNPA2/DAZAP1) siRNAs did not change the pattern of splicing (Fig. 4B, lanes 1 to 5). As a control for hnRNPA1/A2 silencing efficiency we cotransfected the pBRCA1T6 mutant with the pSMN2 minigene in the double-siRNA-treated cells. siRNA-mediated silencing of hnRNPA1 and hnRNPA2 showed, as previously reported (8, 21), significant rescue of the SMN2 splicing defect (Fig. 4D). Exon 18 inclusion in the WT was not affected by siRNA-mediated depletion of any of the three splicing-inhibitory factors (Fig. 4C).

BRCA1 exon 18 contains additional splicing-regulatory sequences that are functional only in the context of the T6 mutant. To identify additional exonic splicing-regulatory elements that modulate T6 aberrant skipping, we performed several consecutive deletions within BRCA1 exon 18. Those deletions analyzed in the context of the WT pBRA had no effect on normal exon 18 inclusion, indicating the absence of strong enhancer sequences in the natural exon (Fig. 5B, lanes 3 to 8). On the other hand, some deletions analyzed in the context of the T6 mutant resulted in changes in the pattern on splicing. In comparison to the T6 mutant, T6Δ35-44 and T6Δ64-74, which disrupt "TAG" elements, increased the percentage of exon inclusion up to 90% and 100%, respectively (Fig. 5C, lanes 5 and 7), indicating the presence of two downstream silencers at positions 35 to 44 and 64 to 74. On the other hand, deletion of the sequences between positions 23 and 32 in the T6Δ23-32 minigene induced complete exon skipping (Fig. 5C, lane 4), indicating the presence of a weak enhancer, which is functionally relevant only in the context of the T6 mutant. In a similar manner we evaluated also the combined effect of the two deletions in positions 4 to 9 and 23 to 32. These deletions alone had no effect on splicing but, when combined (Δ4-9 plus Δ23-32), are able to induce partial exon skipping (Fig. 5C, lane 8). These data suggest the presence of two nearby weak enhancer sequences, which may contribute in a synergic manner to the definition of the WT exon. However, in the context of the Δ23-32 deletion the extent of splicing inhibition mediated by the single T6 substitution is more severe than the splicing inhibition mediated by the Δ4-9 deletion. In fact, the T6Δ23-32 minigene induced complete exon exclusion, whereas Δ4-9-Δ23-32 showed only partial exon skipping, with 50% exon inclusion. This result is consistent with the strong inhibitory effect mediated by the binding of hnRNPA1/A2 and DAZAP1 on the T6 mutant.

To identify potential splicing-regulatory sequences that could mediate the ASF/SF2 observed enhancing effect on the T6 mutant (Fig. 1C), we also evaluated in cotransfection experiments the effect of this splicing factor on the different minigenes with single deletions. Overexpression of ASF/SF2 increased the percentage of exon inclusion in the single T6 mutant and in three deletion mutants, T6Δ12-21, T6Δ35-44, and T6Δ47-54 (Fig. 5D, lanes 4, 8, and 10), and had no effect on the already fully included T6Δ64-74 minigene (Fig. 5D, lane 12). On the other hand, overexpression of this splicing factor did not induce exon inclusion in the T6Δ23-32 minigene (Fig. 5D, lane 6). Thus, the ESE located between nucleotides 23 and 32 in exon 18 stimulates BRCA1 exon 18 splicing through ASF/SF2 only in the context of the T6 mutant.

DISCUSSION

In this paper we provide biochemical and functional evidence that the natural G-to-T transversion at position +6 of BRCA1 exon 18, previously suggested to disrupt an ASF/SF2-dependent splicing enhancer, induces defective splicing through the creation of a sequence with splicing-inhibitory function. We show, in a series of systematic site-directed mutagenesis and pulldown RNA and EMSA experiments, that the mutation creates a splicing silencer element that interacts specifically with three splicing-inhibitory factors: hnRNPA1, hnRNPA2, and DAZAP1. By means of siRNA experiments, we found that hnRNPA1/A2 and DAZAP1 are involved in the generation of a defective splicing of T6 BRCA1 exon 18. In addition, we report a novel role for DAZAP1 in splicing regulation.

Computer-assisted analysis originally suggested that the natural G6T transversion in BRCA1 exon 18 disrupts a putative binding site for ASF/SF2 and accordingly affects an exonic splicing enhancer. Based on this prediction and the functional evaluation of a limited number of site-directed mutants in in vitro splicing assays, a general model of defective splicing due to exonic mutation was proposed consistent with the disruption of an ESE-dependent ASF/SF2 binding site (25). Using a modified pulldown technique followed by Western blot analysis, we unexpectedly found that both WT and T6 mutants bind equally to ASF/SF2 (Fig. 1A, lanes 2 and 3). Most importantly, compared with the classical GAA-rich fibronectin EDA enhancer (Fig. 1A, lane 1), the WT and mutant BRCA1 sequences show a diminished interaction with ASF/SF2, which is unlikely to be functional. The efficiency of this binding was carefully evaluated by linking BRCA1 RNA targets to a (UG) tail which specifically binds TDP43 and allows an accurate normalization of the amount of proteins pulled down. Consistent with the absence of a strong ASF/SF2-dependent enhancer at the 5′ end of the exon disrupted by the mutation, and in accordance with recently published data (21), siRNA-mediated depletion of ASF/SF2 did not induce exon exclusion in the WT minigene (Fig. 1D, lane 1), indicating that ASF/SF2 is not essential for BRCA1 exon 18 splicing. Similarly, the lack of a strong enhancer at the 5′ end of BRCA1 is also indicated by site-directed mutagenesis analysis. In fact, most point substitutions along the 5′ end of BRCA1 exon 18 (Fig. 2A) and the complete deletion of the putative enhancer sequence (Δ4-9 pBRA construct) have no effect on the splicing pattern (Fig. 2C, lane 6).

On the other hand, we found that the T6 mutant responds to siRNA-mediated depletion (Fig. 1D, right, lane 3) or overexpression of ASF/SF2 (Fig. 1C, lane 4) by reducing or increasing, respectively, the percentage of exon inclusion. A small negative effect by depletion of ASF/SF2 has been also recently observed (21). In contrast to what was found for the BRCA1 exon T6 mutant, the defective splicing of the related SMN2 exon 7 gene was not affected by either overexpression or depletion of ASF/SF2 (20, 21), suggesting different roles for this splicing factor in the two systems. Indeed, in a series of short exonic deletions in the WT and T6 BRCA1 minigene contexts we identify a series of splicing-regulatory sequences, which are functionally relevant only in the presence of the T6 mutation (Fig. 5). In particular, the enhancer sequence in position 23 to 32 mediated ASF/SF2-dependent splicing activation (Fig. 5D).

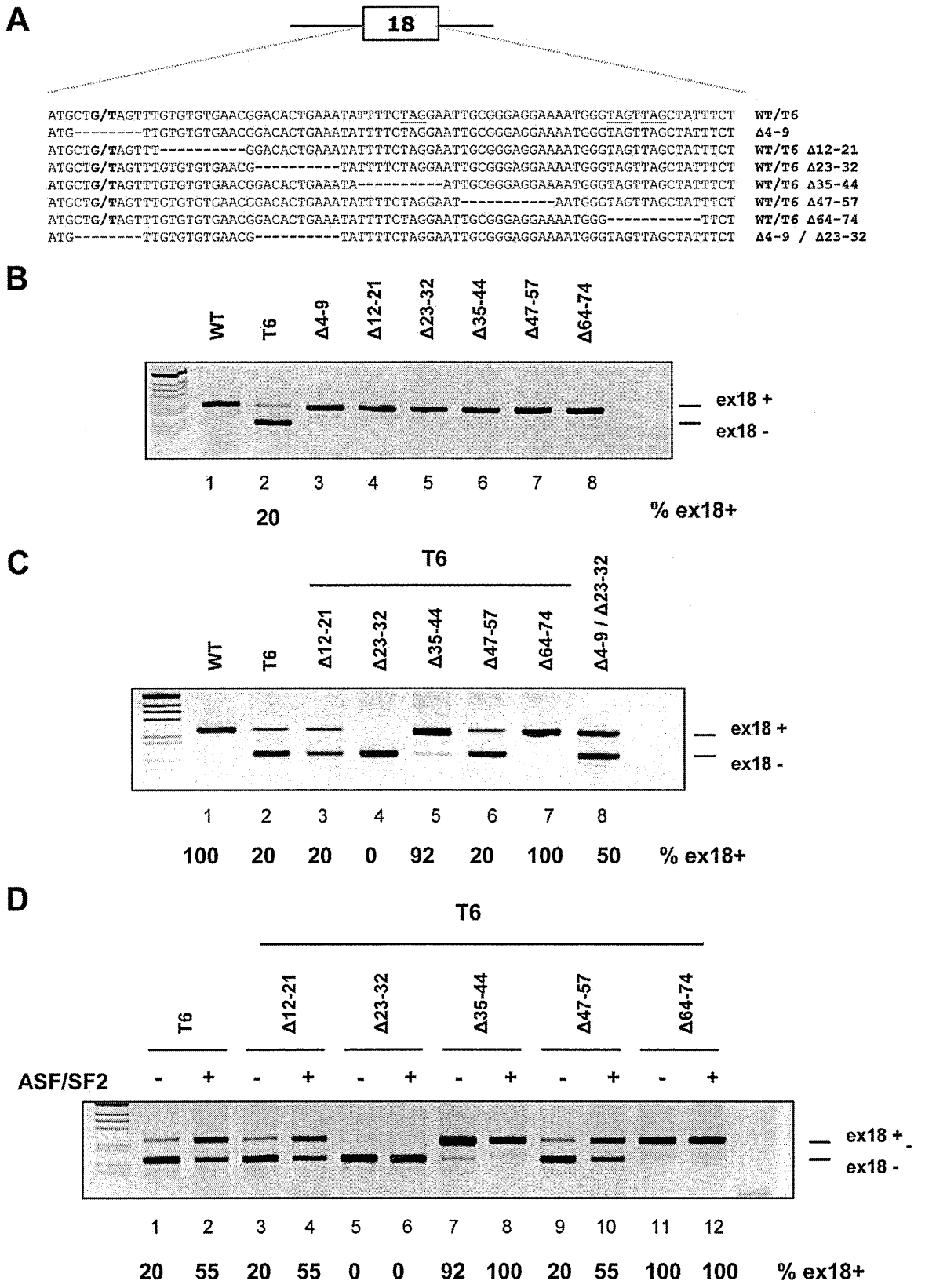

FIG. 5. Mapping the exonic regulatory elements within the BRCA1 exon 18 sequence. (A) Schematic representation of the BRCA1 exon 18 deletion mutants. The G or T substitution at position +6 and the deleted nucleotides are indicated, while the TAG motives are underlined. (B and C) Exon 18 WT (B) and T6 (C) serial deletion analysis. The indicated minigenes were transfected into Hep3B cells, and the splicing pattern was analyzed by RT-PCR. The two resulting splicing products corresponding to exon 18 inclusion (ex18+) and exclusion (ex18−) are indicated, and the percentage of exon inclusion is reported. (D) Overexpression of ASF/SF2 in T6 deletion mutant minigenes. The indicated constructs were cotransfected with the ASF/SF2 coding plasmid (+) or with the empty vector (−).

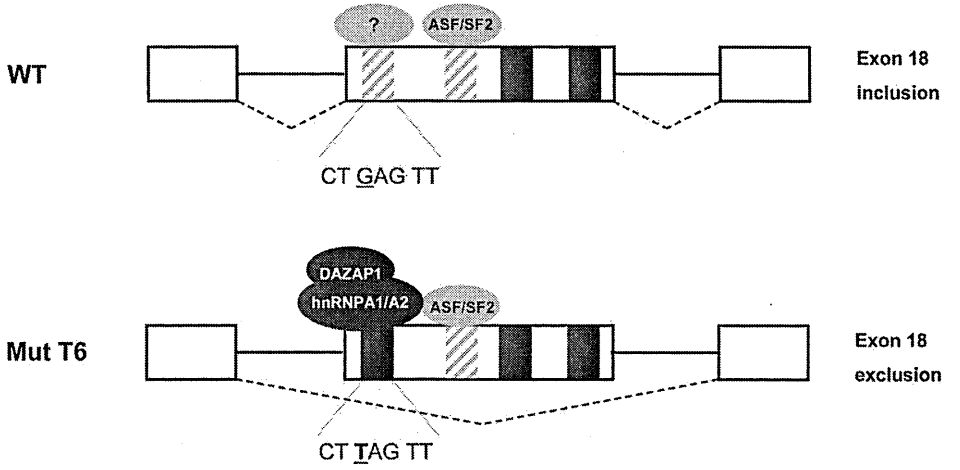

FIG. 6. Splicing-regulatory elements in BRCA1 exon 18 and creation of an ESS by the T6 mutant. WT exon 18 contains two weak ESEs (striped box) and two ESSs (black boxes), and its inclusion is largely dependent on 3′ and 5′ splice site definition. The G-to-T mutation at position +6 creates a strong silencer element at the 5′ end of the exon binding to hnRNPA1/A2 and DAZAP1 and may also disrupt a weak splicing enhancer. The ASF/SF2 ESE located in the middle of exon 18 stimulates BRCA1 exon 18 splicing only in the context of the T6 mutant.

This evidence suggests that the T6 BRCA1-specific response to ASF/SF2 in overexpression and siRNA experiments may be due to the presence of a binding site for this splicing factor at position 23 to 32. This ASF/SF2 ESE stimulates BRCA1 exon 18 splicing, inducing exon inclusion only in the context of the T6 mutant (Fig. 6).

Several lines of evidence indicate that the natural mutation studied here creates an exonic splicing silencer. The double-site-directed mutation experiments indicate that the majority of point mutations affecting the "TAG" core sequence between positions +6 and +8 completely restore normal splicing (Fig. 2B). The mutant sequence, different from the WT, binds to three different proteins in pulldown assays and EMSA: hnRNPA1 and the related hnRNPA2 factor and DAZAP1. hnRNPA1 and hnRNPA2 are well-characterized splicing-inhibitory factors involved in a variety of cellular and viral systems (11) and have been identified as splicing repressors in the regulation of SMN2 exon 7 (20, 21). In contrast to hnRNPA1/A2, few data are available regarding the functional properties of DAZAP1, which has been mainly implicated in mRNA transport and stability (24, 28). Interestingly, SELEX experiments have shown two conserved sequences for the mouse homologue, AAUAG and $GU_{1-3}AG$ (18), which contain the same UAG splicing sequence as the BRCA1 T6 mutant. By means of siRNA experiments, we provide functional evidence that DAZAP1 and hnRNPA1/A2 are involved and probably have a redundant function in the splicing regulation of the defective BRCA1 exon 18. Treatment with hnRNPA1/A2 siRNA resulted in a significant recovery of the defective splicing, which was further enhanced by DAZAP1 depletion (Fig. 4B). The requirement of multiple siRNAs for an efficient rescue of defective T6 splicing and the EMSA experiments with purified proteins suggest that hnRNPA1/A2 and DAZAP1 specifically recognize the inhibitory sequence. This is further supported by the fact that DAZAP1 colocalizes and physically interacts with the splicing repressor proteins hnRNPA1 and hnRNPC1 (24). However, triple treatment with siRNA for hnRNPA1/A2 and DAZAP1 did not completely restore BRCA1 exon inclusion. This can be explained by the presence of residual amounts of the splicing factors, not completely depleted by the siRNAs, or to the binding of an additional splicing-inhibitory factor(s) at the silencer element not revealed by the pulldown analysis. Even if our results clearly indicate that the T6 mutant creates a strong silencer binding to hnRNA1/A2 and DAZAP1, we cannot exclude the simultaneous disruption of a weak enhancer, as recently suggested (21). In fact, we observed that the combined deletion of the 5′ end of BRCA1 exon 18 along with the downstream enhancer (Δ4-9-Δ23-32) induced partial exon skipping. Thus, with the limitation provided by the gross deletion analysis, which can interfere with RNA secondary structure and acceptor site accessibility or reduce the exon definition below a critical threshold, the 5′ end on BRCA1 exon 18 might contain a weak enhancer, which becomes evident only when the downstream enhancer is deleted. This weak enhancer in the WT sequence might bind to an unidentified splicing factor(s) or regulate RNA secondary structure constraints.

In agreement with our findings, the BRCA1 T6 mutant was recently found to bind to hnRNPA1, but its aberrant splicing was not corrected by hnRNPA1 and hnRNPA2 silencing (21). We do not have a clear explanation of these different findings, but culture cell lines may show variable concentrations of splicing-inhibitory factors like DAZAP1. These factors can efficiently replace hnRNAPA1/A2 at the T6 mutant silencer or can act on other splicing-regulatory elements modulating siRNA rescue splicing efficiency. Consistent with this hypothesis, we observed a different cell-type-dependent rescue efficiency of the splicing pattern in siRNA experiments (data not shown).

In summary, our results indicate that the primary determinant of the T6 BRCA1 exon 18 exclusion is not the disruption of an ASF/SF2-dependent enhancer but the creation of a new silencer element recognized specifically by the splicing-inhibitory factors hnRNPA1/A2 and DAZAP1 (Fig. 6). These act as repressors that bind to the silencer and inhibit exon recognition during the splicing process. We have clearly established

that the basic mechanism affecting splicing due to mutations is not always obvious and cannot be simply derived from in silico predictions. The creation of exonic splicing silencers may be a common event in human pathology and has to be taken into account in order to develop therapeutic strategies aimed at correcting splicing defects.

ACKNOWLEDGMENTS

This work was supported by the Associazione Italiana Ricerca sul Cancro and Italian Cystic Fibrosis Foundation to F.P. and E.G.

We thank Cristiana Stuani and Erica Bussani for technical assistance, Adrian Krainer for the pSMN2 minigene, Rodolfo Garcia for helpful discussion, and F. E. Baralle for continuous support and suggestions.

REFERENCES

1. **Ars, E., E. Serra, J. Garcia, H. Kruyer, A. Gaona, C. Lazaro, and X. Estivill.** 2000. Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. Hum. Mol. Genet. **9:**237–247.
2. **Black, D. L.** 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. **72:**291–336.
3. **Bourgeois, C. F., F. Lejeune, and J. Stevenin.** 2004. Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. Prog. Nucleic Acid Res. Mol. Biol. **78:**37–88.
4. **Buratti, E., T. Dork, E. Zuccato, F. Pagani, M. Romano, and F. E. Baralle.** 2001. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. EMBO J. **20:**1774–1784.
5. **Burd, C. G., and G. Dreyfuss.** 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. EMBO J. **13:**1197–1204.
6. **Caceres, J. F., and A. R. Kornblihtt.** 2002. Alternative splicing: multiple control mechanisms and involvement in human disease. Trends Genet. **18:** 186–193.
7. **Cartegni, L., S. L. Chew, and A. R. Krainer.** 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat. Rev. Genet. **3:**285–298.
8. **Cartegni, L., M. L. Hastings, J. A. Calarco, E. de Stanchina, and A. R. Krainer.** 2006. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. Am. J. Hum. Genet. **78:**63–77.
9. **Cartegni, L., and A. R. Krainer.** 2002. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat. Genet. **30:**377–384.
10. **Cartegni, L., J. Wang, Z. Zhu, M. Q. Zhang, and A. R. Krainer.** 2003. ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acids Res. **31:**3568–3571.
11. **Chabot, B., C. LeBel, S. Hutchison, F. H. Nasim, and M. J. Simard.** 2003. Heterogeneous nuclear ribonucleoprotein particle A/B proteins and the control of alternative splicing of the mammalian heterogeneous nuclear ribonucleoprotein particle A1 pre-mRNA. Prog. Mol. Subcell. Biol. **31:** 59–88.
12. **Disset, A., C. F. Bourgeois, N. Benmalek, M. Claustres, J. Stevenin, and S. Tuffery-Giraud.** 2006. An exon skipping-associated nonsense mutation in the dystrophin gene uncovers a complex interplay between multiple antagonistic splicing elements. Hum. Mol. Genet. **15:**999–1013.
13. **Dreyfuss, G., V. N. Kim, and N. Kataoka.** 2002. Messenger-RNA-binding proteins and the messages they carry. Nat. Rev. Mol. Cell Biol. **3:**195–205.
14. **Fairbrother, W. G., R. F. Yeh, P. A. Sharp, and C. B. Burge.** 2002. Predictive identification of exonic splicing enhancers in human genes. Science **297:** 1007–1013.
15. **Fairbrother, W. G., G. W. Yeo, R. Yeh, P. Goldstein, M. Mawson, P. A. Sharp, and C. B. Burge.** 2004. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. Nucleic Acids Res. **32:**W187–W190.
16. **Faustino, N. A., and T. A. Cooper.** 2003. Pre-mRNA splicing and human disease. Genes Dev. **17:**419–437.
17. **Graveley, B. R.** 2000. Sorting out the complexity of SR protein functions. RNA **6:**1197–1211.
18. **Hori, T., Y. Taguchi, S. Uesugi, and Y. Kurihara.** 2005. The RNA ligands for mouse proline-rich RNA-binding protein (mouse Prrp) contain two consensus sequences in separate loop structure. Nucleic Acids Res. **33:**190–200.
19. **Jurica, M. S., and M. J. Moore.** 2003. Pre-mRNA splicing: awash in a sea of proteins. Mol. Cell **12:**5–14.
20. **Kashima, T., and J. L. Manley.** 2003. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. Nat. Genet. **34:**460–463.
21. **Kashima, T., N. Rao, C. J. David, and J. L. Manley.** 2007. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. Hum. Mol. Genet. **16:**3149–3159.
22. **Krecic, A. M., and M. S. Swanson.** 1999. hnRNP complexes: composition, structure, and function. Curr. Opin. Cell Biol. **11:**363–371.
23. **Lim, L. P., and C. B. Burge.** 2001. A computational analysis of sequence features involved in recognition of short introns. Proc. Natl. Acad. Sci. USA **98:**11193–11198.
24. **Lin, Y. T., and P. H. Yen.** 2006. A novel nucleocytoplasmic shuttling sequence of DAZAP1, a testis-abundant RNA-binding protein. RNA **12:**1486–1493.
25. **Liu, H. X., L. Cartegni, M. Q. Zhang, and A. R. Krainer.** 2001. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nat. Genet. **27:**55–58.
26. **Lorson, C. L., and E. J. Androphy.** 2000. An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. Hum. Mol. Genet. **9:**259–265.
27. **Mazoyer, S., N. Puget, L. Perrin-Vidoz, H. T. Lynch, O. M. Serova-Sinilnikova, and G. M. Lenoir.** 1998. A BRCA1 nonsense mutation causes exon skipping. Am. J. Hum. Genet. **62:**713–715.
28. **Morton, S., H. T. Yang, N. Moleleki, D. G. Campbell, P. Cohen, and S. Rousseau.** 2006. Phosphorylation of the ARE-binding protein DAZAP1 by ERK2 induces its dissociation from DAZ. Biochem. J. **399:**265–273.
29. **Muro, A. F., M. Caputi, R. Pariyarath, F. Pagani, E. Buratti, and F. E. Baralle.** 1999. Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. Mol. Cell. Biol. **19:**2657–2671.
30. **Nilsen, T. W.** 2003. The spliceosome: the most complex macromolecular machine in the cell? Bioessays **25:**1147–1149.
31. **Pagani, F., and F. E. Baralle.** 2004. Genomic variants in exons and introns: identifying the splicing spoilers. Nat. Rev. Genet. **5:**389–396.
32. **Pagani, F., E. Buratti, C. Stuani, and F. E. Baralle.** 2003. Missense, nonsense, and neutral mutations define juxtaposed regulatory elements of splicing in cystic fibrosis transmembrane regulator exon 9. J. Biol. Chem. **278:** 26580–26588.
33. **Pagani, F., E. Buratti, C. Stuani, M. Romano, E. Zuccato, M. Niksic, L. Giglio, D. Faraguna, and F. E. Baralle.** 2000. Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element. J. Biol. Chem. **275:**21041–21047.
34. **Pagani, F., C. Stuani, M. Tzetis, E. Kanavakis, A. Efthymiadou, S. Doudounakis, T. Casals, and F. E. Baralle.** 2003. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. Hum. Mol. Genet. **12:**1111–1120.
35. **Pozzoli, U., and M. Sironi.** 2005. Silencers regulate both constitutive and alternative splicing events in mammals. Cell. Mol. Life Sci. **62:**1579–1604.
36. **Shen, H., and M. R. Green.** 2004. A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. Mol. Cell **16:**363–373.
37. **Singh, N. N., E. J. Androphy, and R. N. Singh.** 2004. An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy. Biochem. Biophys. Res. Commun. **315:**381–388.
38. **Smith, P. J., C. Zhang, J. Wang, S. L. Chew, M. Q. Zhang, and A. R. Krainer.** 2006. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum. Mol. Genet. **15:**2490–2508.
39. **Sun, H., and L. A. Chasin.** 2000. Multiple splicing defects in an intronic false exon. Mol. Cell. Biol. **20:**6414–6425.
40. **Teraoka, S. N., M. Telatar, S. Becker-Catania, T. Liang, S. Onengut, A. Tolun, L. Chessa, O. Sanal, E. Bernatowska, R. A. Gatti, and P. Concannon.** 1999. Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. Am. J. Hum. Genet. **64:**1617–1631.
41. **Zhang, X. H., and L. A. Chasin.** 2004. Computational definition of sequence motifs governing constitutive exon splicing. Genes Dev. **18:**1241–1250.
42. **Zhang, X. H., T. Kangsamaksin, M. S. Chao, J. K. Banerjee, and L. A. Chasin.** 2005. Exon inclusion is dependent on predictable exonic splicing enhancers. Mol. Cell. Biol. **25:**7323–7332.